

# A MEMETICS COMPENDIUM

Compiled By

Dr. Robert Finkelstein

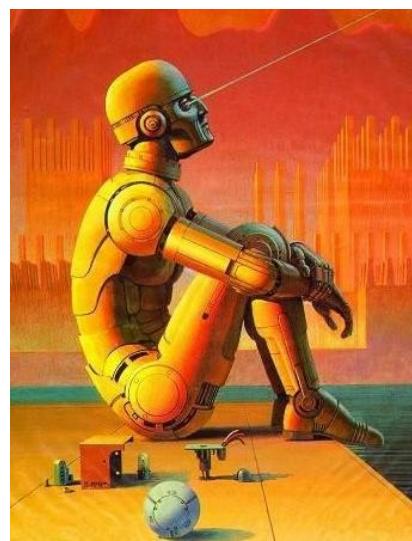
President, Robotic Technology Inc.

Collegiate Professor, University Of Maryland University College

[BobF@RoboticTechnologyInc.com](mailto:BobF@RoboticTechnologyInc.com)

[www.RoboticTechnologyInc.com](http://www.RoboticTechnologyInc.com)

May 2008



## CONTENTS

(Note: Some of the papers have been edited for grammar, spelling, format, and clarity)

<b>COMPENDIUM INTRODUCTION: MILITARY MEMETICS OVERVIEW</b>	<b>Robert Finkelstein</b>	<b>12</b>
<b>COMPENDIUM EXCERPTS</b>	<b>Robert Finkelstein</b>	<b>21</b>
<b>CRITIQUE AND DEFENSE OF MEMESIS</b>	<b>Onar Aam</b>	<b>58</b>
<b>MEMETICS AND MONTAG</b>	<b>B. Antomarini, et. al.</b>	<b>63</b>
<b>DO MEMES ACCOUNT FOR CULTURE?</b>	<b>Robert Aunger</b>	<b>69</b>
<b>CULTURAL EVOLUTION OF “GUIDING CRITERIA”</b>	<b>Gianluca Baldassarre</b>	<b>78</b>
<b>MEMESIS CRITIQUE</b>	<b>Richard Barbrook</b>	<b>95</b>
<b>MEMETIC MEANINGS</b>	<b>Francis A. Beer</b>	<b>100</b>
<b>MODELS FOR INTERACTING POPULATIONS OF MEMES</b>	<b>Michael L. Best</b>	<b>104</b>
<b>A CRITIQUE OF MEMETIC MODELS</b>	<b>Michael L. Best</b>	<b>122</b>
<b>THE LIFECYCLE OF MEMES</b>	<b>Henrik Bjarneskans, et. al.</b>	<b>127</b>
<b>MEMES, MINDS, AND SELVES</b>	<b>Susan Blackmore</b>	<b>148</b>
<b>WAKING FROM THE MEME DREAM</b>	<b>Susan Blackmore</b>	<b>160</b>
<b>IMITATION AND THE DEFINITION OF A MEME</b>	<b>Susan Blackmore</b>	<b>170</b>
<b>THE FORGET-MEME-NOT THEORY</b>	<b>Susan Blackmore</b>	<b>182</b>
<b>THE POWER OF THE MEME MEME</b>	<b>Susan Blackmore</b>	<b>186</b>
<b>THE EVOLUTIONARY ECOLOGY OF SCIENCE</b>	<b>Marion Blute</b>	<b>201</b>
<b>MEMETICS AND EVOLUTIONARY SOCIAL SCIENCE</b>	<b>Marion Blute</b>	<b>223</b>
<b>CONSTRUCTION OF IGNORANCE, EVOLUTION OF KNOWLEDGE</b>	<b>P. Bouissac</b>	<b>227</b>
<b>EDITORIAL: MEMES MATTER</b>	<b>Paul Bouissac</b>	<b>241</b>
<b>WHY DO MEMES DIE?</b>	<b>Paul Bouissac</b>	<b>248</b>
<b>THE HUMAN AGENCY OF MEME MACHINES</b>	<b>Gary Boyd</b>	<b>256</b>
<b>INTRODUCTION TO THE CRISIS OF THE MIND</b>	<b>Richard Brodie</b>	<b>264</b>
<b>THE MEME HUNTER</b>	<b>Andrew Brown</b>	<b>271</b>
<b>MINIMAL REQUIREMENTS FOR DARWINIAN BOOTSTRAPING</b>	<b>William H. Calvin</b>	<b>275</b>
<b>TOWARDS A COGNITIVE MEMETICS</b>	<b>Cristiano Castelfranchi</b>	<b>286</b>

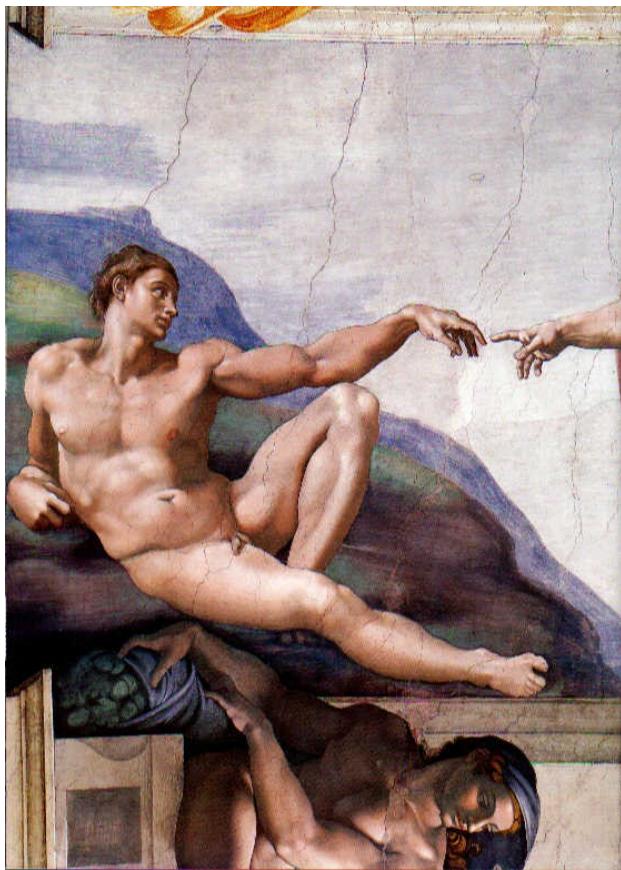
<b>HOW (UN) REALISTIC ARE EVOLUTIONARY ALGORITHMS Edmund Chattoe</b>	<b>311</b>
<b>VIRTUAL URBAN LEGENDS Edmund Chattoe</b>	<b>355</b>
<b>PARASITE ECOLOGY AND THE EVOLUTION OF RELIGION Ben Cullen</b>	<b>372</b>
<b>CHAPTER 11 FROM “THE SELFISH GENE” Richard Dawkins</b>	<b>378</b>
<b>SURVIVAL OF THE INSTITUTIONALLY FITTEST CONCEPTS Martin de Jong</b>	<b>394</b>
<b>MEMES AND THE EXPLOITATION OF IMAGINATION Daniel Dennett</b>	<b>414</b>
<b>MEMES: MYTHS, MISUNDERSTANDINGS, AND MISGIVINGS Daniel C. Dennett</b>	<b>427</b>
<b>THE EVOLUTION OF CULTURE Daniel C. Dennett</b>	<b>438</b>
<b>COMPETING MEMES ANALYSES David K. Dirlam</b>	<b>451</b>
<b>USING MEMETICS TO GROW MEMETICS David K. Dirlam</b>	<b>461</b>
<b>SIMULATING COLLECTIVE MISBELIEF Jim Doran</b>	<b>469</b>
<b>THE REVEALED POVERTY OF THE GENE-MEME ANALOGY Bruce Edmonds</b>	<b>483</b>
<b>THE EVOLUTION OF LANGUAGE AND SCIENCE Börje Ekstig</b>	<b>486</b>
<b>A JUSTIFICATION OF SOCIETAL ALTRUISM John R. Evers</b>	<b>500</b>
<b>STROLLING THROUGH THE MEMETIC MINE FIELD Leon Felkins</b>	<b>506</b>
<b>VIRAL MARKETING: SPREAD A COLD, CATCH A CUSTOMER Frank Fiore</b>	<b>519</b>
<b>CULTURAL R/K SELECTION Agner Fog</b>	<b>522</b>
<b>APPLYING MEMETICS TO FINANCIAL MARKETS Joshua Frank</b>	<b>538</b>
<b>A COMPUTATIONAL MODEL OF CULTURAL EVOLUTION Liane M. Gabora</b>	<b>550</b>
<b>A DAY IN THE LIFE OF A MEME Liane Gabora</b>	<b>558</b>
<b>MEMES: THE CREATIVE SPARK Liane M. Gabora</b>	<b>590</b>
<b>THE ORIGIN AND EVOLUTION OF CULTURE AND CREATIVITY Liane Gabora</b>	<b>592</b>
<b>A TENTATIVE SCENARIO FOR THE ORIGIN OF CULTURE Liane Gabora</b>	<b>623</b>
<b>REVIEW OF THE MEME MACHINE Liane Gabora</b>	<b>649</b>
<b>MACROMEMETICS Derek Gatherer</b>	<b>655</b>
<b>“THOUGHT CONTAGION” METAPHOR IS RETARDING PROGRESS Derek Gatherer</b>	<b>673</b>
<b>THE CASE FOR COMMENTARY Derek Gatherer</b>	<b>698</b>
<b>A PLEA FOR METHODOLOGICAL DARWINISM Derek Gatherer</b>	<b>700</b>

<b>REPLY TO COMMENTARIES ON MY PAPER</b> Derek Gatherer	<b>703</b>
<b>MODELING MEMETIC TABOOS ON GENETIC HOMOSEXUALITY</b> Derek Gatherer	<b>709</b>
<b>THE SPREAD OF IRRATIONAL BEHAVIORS BY CONTAGION</b> Derek Gatherer	<b>727</b>
<b>FINDING A NICHE FOR MEMETICS IN THE 21<sup>ST</sup> CENTURY</b> Derek Gatherer	<b>748</b>
<b>THE INHERENT INSTABILITY OF MEMETIC SYSTEMS</b> Derek Gatherer	<b>754</b>
<b>MEMES AND ORGANIZATIONAL STRUCTURES</b> Hélène Giroux et. al.	<b>776</b>
<b>THEISTIC MEMES IN RELIGIOUS CANONS</b> John D. Gottsch	<b>783</b>
<b>MEMES: INTRODUCTION</b> Glenn Grant	<b>818</b>
<b>MEMETIC LEXICON</b> Glenn Grant	<b>820</b>
<b>MEMES AND THE CREATION OF DANCE</b> Christine Greiner	<b>826</b>
<b>MEMES AND GRINNING IDIOT PRESS</b> Grinning Idiot Press 1993	<b>829</b>
<b>THE ADOPTION MEME</b> Dave Gross	<b>835</b>
<b>APPLYING NATURAL SELECTION THINKING TO URBAN LEGENDS</b> Dave Gross	<b>839</b>
<b>REFLECTIONS ON CREATION VERSUS EVOLUTION OF MEMES</b> Dave Gross	<b>847</b>
<b>MEMETICS: A SYSTEMS METABIOLOGY</b> Ron Hale-Evans	<b>849</b>
<b>EXPERIMENTS IN THE META-NOOSPHERE</b> David Hales	<b>867</b>
<b>BELIEF HAS UTILITY – AN INTENTIONAL STANCE</b> David Hales	<b>897</b>
<b>MEMETICS</b> F. Heylighen	<b>900</b>
<b>COMPETITION BETWEEN MEMES AND GENES</b> F. Heylighen	<b>903</b>
<b>STRUCTURE OF MEMES</b> F. Heylighen	<b>905</b>
<b>MEMETIC SELECTION CRITERIA</b> F. Heylighen	<b>908</b>
<b>EVOLUTION OF MEMES ON THE NETWORK: GLOBAL BRAIN</b> Francis Heylighen	<b>911</b>
<b>WHAT MAKES A MEME SUCCESSFUL?</b> Francis Heylighen	<b>917</b>
<b>REBUTTAL OF BEHAVIORIST APPROACH TO MEMETICS</b> Francis Heylighen	<b>923</b>
<b>STRATEGIES IN MEME THEORY</b> David L. Hull	<b>929</b>
<b>MEMETICS AS A METHODOLOGICAL TOOL</b> Rogan Jacobson	<b>932</b>
<b>GREAT MEN, GREAT THOUGHTS, AND THE ENVIRONMENT</b> William James	<b>939</b>
<b>REPLICATING SONORITIES: TOWARDS A MEMETICS OF MUSIC</b> Steven Jan	<b>955</b>

<b>MATHEMATICAL MODELS FOR MEMETICS</b> Jeremy R. Kendal and K. Laland	<b>1004</b>
<b>DIFFERENT TYPES OF MEMES</b> John Z. Langrish	<b>1016</b>
<b>A NOTE ON THE ORIGIN OF “MEMES/MNEMES”</b> John Laurent	<b>1031</b>
<b>THE REDEFINITION OF MEMES: MEANING TO AN EMPTY CLICHÉ</b> Michael R. Lissack	<b>1033</b>
<b>UNITS, EVENTS, AND DYNAMICS IN MEMETIC EVOLUTION</b> Aaron Lynch	<b>1048</b>
<b>RATIONALES FOR MATHEMATICAL MODELING</b> Aaron Lynch	<b>1081</b>
<b>MISLEADING MIX OF RELIGION AND SCIENCE</b> Aaron Lynch	<b>1087</b>
<b>A RESPONSE TO PAUL MARDEN</b> Aaron Lynch	<b>1091</b>
<b>BIRDSONG: MEMETIC EVOLUTION OF SONGS OR GRAMMARS</b> William Majoros	<b>1097</b>
<b>DAWKINS BAD IDEA: MEMES, GENES, METAPHORS</b> Arthur J. Marr	<b>1117</b>
<b>CRASH CONTAGION</b> Paul Marsden	<b>1123</b>
<b>OPERATIONALIZING MEMETICS: SUICIDE, WERTHER EFFECT</b> Paul Marsden	<b>1129</b>
<b>MEMETICS ON THE EDGE OF CHAOS</b> Paul Marsden	<b>1139</b>
<b>MEMETICS AND SOCIAL CONTAGION</b> Paul Marsden	<b>1145</b>
<b>A STRATEGY FOR MEMETICS: MEMES AS STRATEGIES</b> Paul Marsden	<b>1161</b>
<b>REVIEW: THOUGHT CONTAGION</b> Paul Marsden	<b>1166</b>
<b>FOREFATHERS OF MEMETICS</b> Paul Marsden	<b>1176</b>
<b>IS SUICIDE CONTAGIOUS? A CASE STUDY IN APPLIED MEMETICS</b> Paul Marsden	<b>1181</b>
<b>THE INTERNET AND MEMETICS</b> Garry Marshall	<b>1190</b>
<b>THOUGHTS AS TOOLS: THE MEME IN DANIEL DENNETT'S WORK</b> Kelby Mason	<b>1196</b>
<b>THE FUTURE OF MAN</b> Sir Peter Medawar	<b>1203</b>
<b>EVOLUTIONARY PARADIGM FOR GLOBAL POLITICS</b> George Modelska	<b>1211</b>
<b>AN EVOLUTIONARY THEORY OF CULTURE?</b> George Modelska	<b>1232</b>
<b>AN APPROACH TO GENDER AND CONSUMER BEHAVIOR: MEMETICS</b> TJ Olney	<b>1234</b>
<b>CREATIVITY, EVOLUTION, AND MENTAL ILLNESS</b> Antonio Preti and Paola Miotto	<b>1246</b>
<b>THE LEARNING ORGANISATION MEME</b> If Price and Ray Shaw	<b>1254</b>
<b>PUNCTUATING ORGANISATIONAL EQUILIBRIUM</b> If Price	<b>1267</b>
<b>THE AUTHORS' REPLY TO PAUL MARDEN</b> If Price and Ray Shaw	<b>1272</b>

<b>STEPS TOWARD THE MEMETIC SELF If Price</b>	<b>1277</b>
<b>RECONSTRUCTION OF ORGANIZATIONAL PHYLOGENY If Price and Andrew Lord</b>	<b>1283</b>
<b>THE SELFISH TEXT: THE BIBLE AND MEMETICS Hugh S. Pyper</b>	<b>1294</b>
<b>DO ANIMALS HAVE MEMES? Simon M. Reader &amp; Kevin N. Laland</b>	<b>1309</b>
<b>TRANSMISSION OF CULTURAL TRAITS BY EMULATION Robert G. Reynolds, et. al.</b>	<b>1318</b>
<b>LANGUAGE CHANGE AS EVOLUTION: LINGUISTIC “GENES” Nikolaus Ritt</b>	<b>1350</b>
<b>CONTROVERSIES IN MEME THEORY Nick Rose</b>	<b>1361</b>
<b>WHO WOULD ESCAPE THE TYRANNY OF THE REPLICATORS? Nick Rose</b>	<b>1373</b>
<b>DARWINIAN SOUP W.G. Runciman</b>	<b>1376</b>
<b>DARWINIAN PROCESSES AND MEMES IN ARCHITECTURE Nikos Salingaros</b>	<b>1382</b>
<b>GERMS AND IDEAS André Siegfried</b>	<b>1403</b>
<b>ARE ANIMAL MINDS SUITABLE HABITATS FOR MEMES? Brent Silby</b>	<b>1410</b>
<b>WHAT IS A MEME? Brent Silby</b>	<b>1417</b>
<b>EVOLUTION OF TECHNOLOGY: EXPOSING CREATIVE DESIGN MYTH Brent Silby</b>	<b>1421</b>
<b>A SHORT COMMENT FROM A BIOLOGIST Hans-Cees Speel</b>	<b>1431</b>
<b>MEMES ARE ALSO INTERACTORS H.C.A.M. Speel</b>	<b>1437</b>
<b>ON MEMETICS AND MEMES AS BRAIN-ENTITIES Hans-Cees Speel</b>	<b>1443</b>
<b>A MEMETIC ANALYSIS OF POLICY MAKING Hans-Cees Speel</b>	<b>1446</b>
<b>EXPLAINING CULTURE: A NATURALISTIC APPROACH Dan Sperber</b>	<b>1474</b>
<b>AN OBJECTION TO THE MEMETIC APPROACH TO CULTURE Dan Sperber</b>	<b>1480</b>
<b>THE MEMETIC ORIGIN OF LANGUAGE Mario Vaneechoutte</b>	<b>1489</b>
<b>THE REPLICATOR: A MISNOMER Mario Vaneechoutte</b>	<b>1525</b>
<b>MERGERS AND TAKEOVERS; A MEMETIC APPROACH Ed Vos and Ben Kelleher</b>	<b>1533</b>
<b>A REVIEW OF: THE NEW EVOLUTIONARY MICROECONOMICS Jack J. Vromen</b>	<b>1547</b>
<b>HOW TO MAKE MEMES AND INFLUENCE PEOPLE Adam Westoby</b>	<b>1549</b>
<b>WHAT'S A MEME: REFLECTIONS FROM BIOLOGY John S. Wilkins</b>	<b>1634</b>
<b>MEMES AIN'T (JUST) IN THE HEAD John Wilkins</b>	<b>1670</b>
<b>ON CHOOSING TO EVOLVE John Wilkins</b>	<b>1678</b>

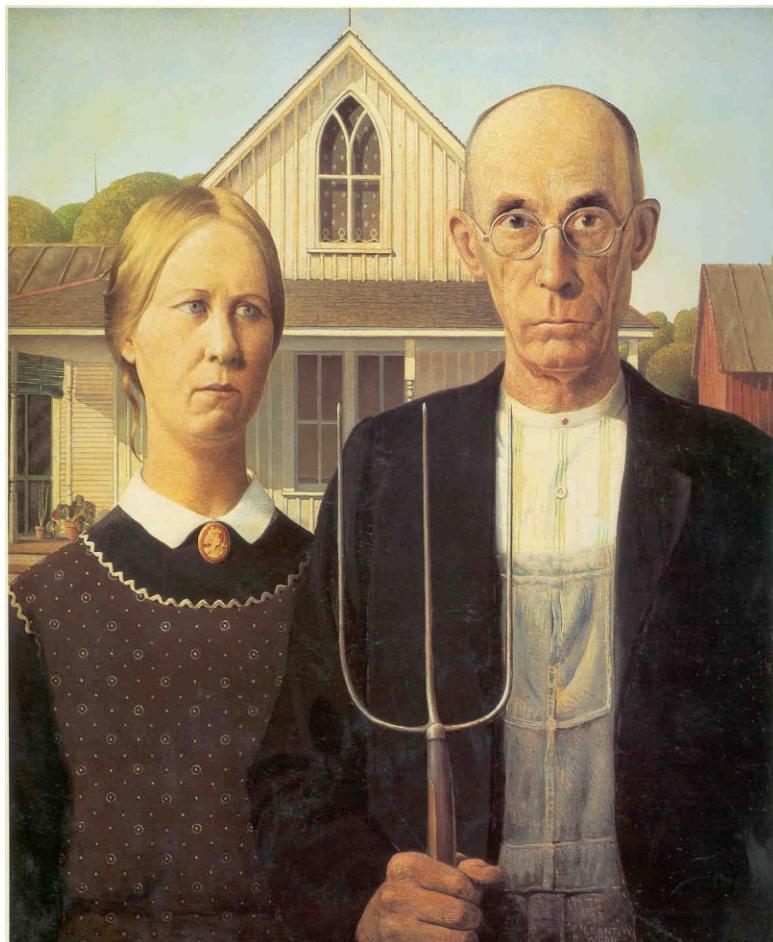
(This Page Intentionally Blank)

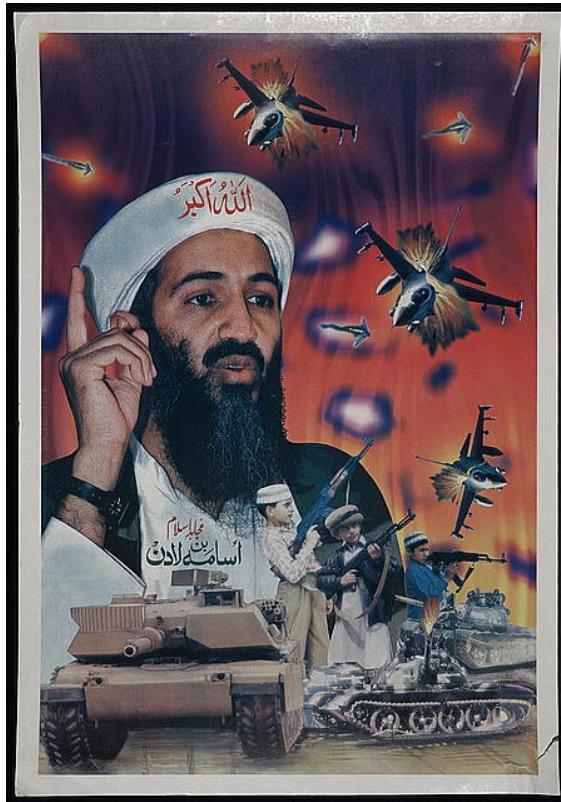






**MEMES  
AHEAD**





# **INTRODUCTION to the COMPENDIUM and a MILITARY MEMETICS OVERVIEW**

**By**  
**Robert Finkelstein**  
**Robotic Technology Inc.**

## **PURPOSE**

The purpose of this Compendium is to provide a broad overview of memetics based on a variety of papers by various authors. These digital papers were gathered from the Internet during 2006 to 2008, assembled, and, in some cases, edited for grammar, spelling, and clarity; all of the papers were transformed into a common format (font, font size, appearance) for coherence to ease reading. The compendium was developed as part of memetics research conducted for an Agency of the Department of Defense (DOD).

In the more than 30 years since the meme was identified as an entity of interest, there has been almost no research to determine whether memetics can be placed on a scientific basis so that its phenomena and effects can be quantified, predicted, and controlled. Most of the research of memetics, such as it is, has been conducted in scattered efforts, often as personal projects on the part of dedicated academics, mostly outside the U.S. There have been no significant, coherent efforts until now, where there is an effort to explore whether a scientific framework can be established for memetics.

The papers were selected subjectively, based on their perceived relevance to the subject and interest to the reader. For the ease of locating papers in the Compendium, they are arranged alphabetically by author, rather than topic.

Following this Introduction is a section of Compendium Excerpts, with passages I subjectively extracted from the papers based on what I thought might be of interest to the memetics project. The excerpts are alphabetical, by author, so the original complete text can be found easily in the Compendium by readers interested in a particular passage.

A separate volume includes articles and papers which are relevant to memetics but are not directly about memetics.

The purpose of the following overview is to provide an indication of the prospective value of memetics to the U.S. military for conventional and asymmetric operations, including counter-terrorism.

The attempt to establish a scientific basis for memetics is critically important. For example, within a suitable memetics framework could be the means to prevent irrational conflict and promote rational solutions to endemic national and international problems. Of course, without safeguards memetics can become a double-edged sword.

## THE MEME

The word “meme” is a neologism coined by Richard Dawkins in *The Selfish Gene* (1976), (although it may have had earlier roots) and defined as a self-reproducing and propagating information structure analogous to a gene in biology. Dawkins focused on the meme as a replicator, analogous to the gene, able to affect human evolution through the evolutionary algorithm of variation, replication, and differential fitness. But for near-term military or national security applications, the relevant characteristics of the meme are that it consists of information which persists, propagates, and influences human behavior.

### Meme Definitions

There are a plethora of definitions for the meme, with most being variations of Dawkins original notion of a unit (whatever that means) of cultural transmission, where culture may be defined as the total pattern of behavior (and its products) of a population of agents, embodied in thought, behavior, and artifacts, dependent upon the capacity for learning and transmitting knowledge to succeeding generations. None of extant definitions of a meme is sufficient to allow a meme to be clearly recognized, measured, or provide the basis for scientific research. And without the establishment of a scientific basis and the ability to explain, predict, and control phenomena, memetics will remain a functionally useless pseudo-science. A few of the many definitions extracted from the literature include:

- A self-reproducing and propagating information structure analogous to a gene in biology
- A unit of cultural transmission (or a unit of imitation) that is a replicator that propagates in the meme pool leaping from brain to brain via (in a broad sense) imitation; examples: “tunes, ideas, catch-phrases, clothes fashions, ways of making pots or of building arches”
- Ideas that program for their own retransmission or propagation
- Actively contagious ideas or thoughts
- Shared elements of a culture learned through imitation from others – with culture being defined rather broadly to include ideas, behaviors, and physical objects
- An element of a culture that may be considered to be passed on by non-genetic means, especially imitation
- Information patterns infecting human minds
- While the internal meme is equivalent to the genotype, its expression in behavior (or the way it affects things in its environment) is its phenotype
- Any information that is copied from person to person or between books, computers, or other storage devices. Many mental contents are not memes because they are not

acquired by imitation or copying, including perceptions, visual memories, and emotional feelings. Skills or knowledge acquired by ordinary learning are not memes

- A (cognitive) information-structure able to replicate using human hosts and to influence their behavior to promote replication
- Cultural information units that are the smallest elements that replicate themselves with reliability and fecundity
- A rule of behavior, encoded by functional neuronal groups or pathways. [Behavior is action, whether mental or physical. Ideas such as tying shoe-laces or opening a door represent rules of physical action, i.e., rules of patterned neural-muscular interaction. Concepts such as apple, seven, or causality, represent rules of mental action, or rules of cognition, i.e., rules of patterned neural-neural interaction. Hence, physical movement is governed by memes which represent rules of physical action and thought is governed by memes which represent rules of mental action]
- Any kind, amount, and configuration of information in culture that shows both variation and coherent transmission
- A pattern of information (a state within a space of possible states)]
- A unit of cultural information as it is represented in the brain
- An observable cultural phenomenon, such as a behavior, artifact or an objective piece of information, which is copied, imitated or learned, and thus may replicate within a cultural system. Objective information includes instructions, norms, rules, institutions and social practices provided they are observable
- A pattern of information, one that happens to have evolved a form which induces people to repeat that pattern
- A contagious information pattern that replicates by parasitically infecting human minds and altering their behavior, causing them to propagate the pattern. Individual slogans, catch-phrases, melodies, icons, inventions, and fashions are typical memes. An idea or information pattern is not a meme until it causes someone to replicate it, to repeat it to someone else. All transmitted knowledge is memetic
- The smallest idea that can copy itself while remaining self contained and intact – essentially sets of instructions that can be followed to produce behavior

While Dawkins focused on the meme as a replicator, analogous to the gene, able to affect human evolution through the evolutionary algorithm of variation, replication, and differential fitness, for near-term practical applications the relevant key characteristics of the meme are that it consists of information which persists, propagates, and influences human behavior.

We developed a pragmatic definition to distinguish a meme from other sorts of information (such as common daily utterances):

*A meme is information which propagates, persists, and has impact.*

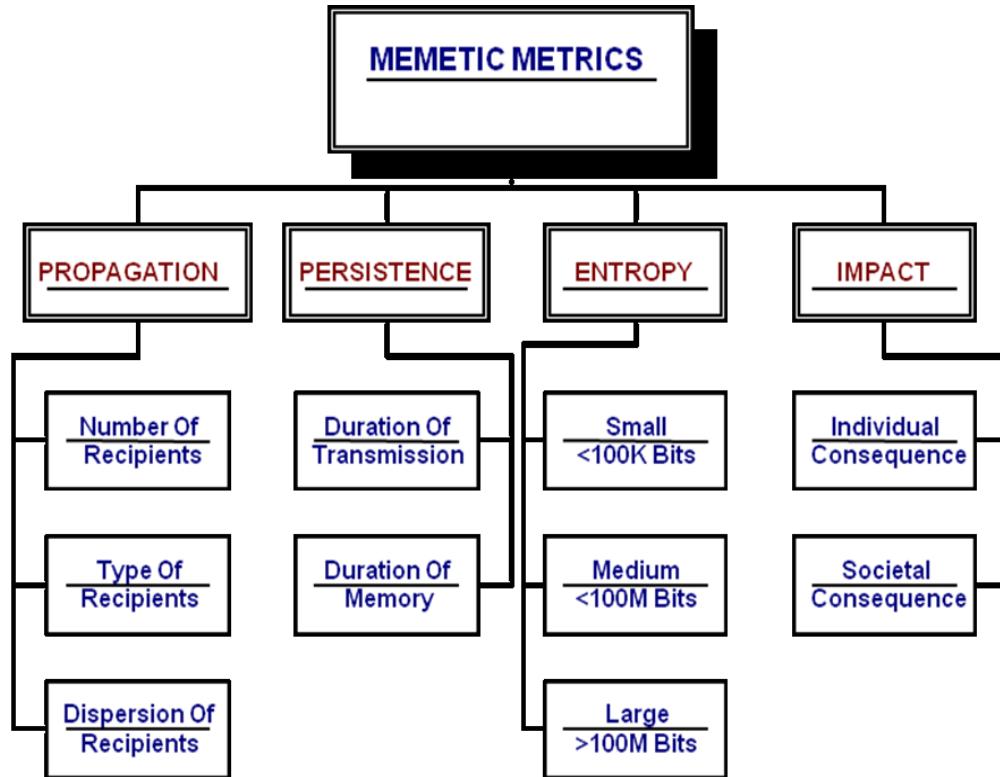
To distinguish memes from other kinds of information, an elaboration of the definition invokes a threshold for propagation and persistence and employs Shannon's definition of *information* as *that which reduces uncertainty*. Impact might take place in the individual, in terms of neurophysiological changes (e.g., as detected by a brain scan) or manifestations of behavior; or it might take place in society (groups of individuals) as manifested by behavior. Metrics are defined (each with submetrics) to facilitate the measurement and comparison of memes. Additional metrics may be defined based on the results of future experiments with tools such as functional magnetic resonance imaging (fMRI). Other experimental approaches include the modeling and simulation of social networks and information propagation through social networks, as well as the tools and techniques of neuroeconomics.

As a practical approach, the metrics we initially defined for evaluating memes, as shown in Figure 1, include: propagation, persistence, impact, and entropy. The submetrics for the *propagation* metric include the number, type, and dispersion of recipients of the meme. Depending on the problem under consideration, the type of recipients might be characterized or categorized by their economic, social, or educational class, ethnicity or culture, religion, gender, age, tribe, politics, etc., while the dispersion of recipients might categorized as local, tribal, familial, regional, national, global, etc. The submetrics for the *persistence* metric distinguish between the duration of transmission of the meme and the duration of the meme in memory or storage. The submetrics for the *entropy* metric distinguish among small, medium, and large memes, which (using an order of magnitude rule) are characterized as less than or equal to 100K bits, less than or equal to 100M bits, and greater than 100M bits. Submetrics for the *impact* metric distinguish between the impact (or potential impact) of the meme on the individual (i.e., individual consequence) and its impact (or potential impact) on society as a whole (i.e., societal consequence). For the initial exercises, we used a multivariate decision tool – the Analytic Hierarchy Process (AHP) – to weight the metrics and submetrics. This definition of the meme may be modified as the research progresses. A paper, not included in this Compendium, discusses our definition and its application in detail. In the paper we also develop an equation for a characteristic Memetic Fitness. Our definition of meme and derivation of memetic fitness are intended to elicit comment and suggestions for improvement.

## Meme Transmission

As shown in Figure 2 (which replicates Claude Shannon's iconic schematic of a general communications system), a meme is transmitted after either being created in the mind of an individual or re-transmitted after being received by an individual from elsewhere. Arriving at a new potential host, the meme is received and decoded. The potential host becomes an actual host if the meme satisfies certain selection and fitness criteria. The new host replicates and transmits the meme (perhaps with a different vector, such as a text message instead of speech). Because the number of memes at any given time typically exceeds the number of recipients able to absorb them, fitness criteria determine which memes will survive, propagate, persist, and have

impact. The selection and fitness criteria include such human motivators as fear (e.g., of going to hell or failing in business) and reward (e.g., of going to heaven or succeeding in business); or the meme might be beneficial in a practical way (such as instruction on how to make a hard-boiled egg or an improvised explosive device); or the meme might be entertaining to the recipient, such as a joke (“why did the terrorist cross the road?”) or a song (“you can’t always get what you want”), or consist of appreciative direct feedback to the recipient (such as providing emotional satisfaction, e.g., reinforcement and pride in membership in a nation, tribe, religion, ethnic group, or ideology).

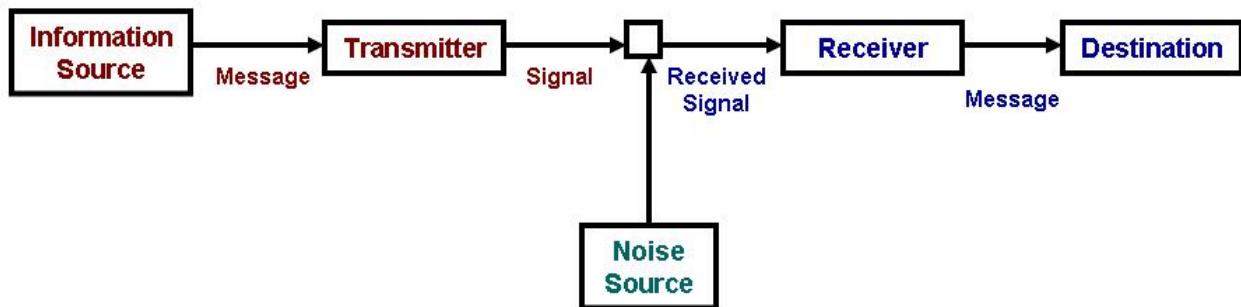


**Figure 1: Memetic Metrics and Submetrics**

To be readily acceptable to the host, the meme should fit existing constructs or belief systems of the host, or be a paradigm to which the host is receptive. Memes also aggregate and reinforce in complexes (memeplexes) so that a suitable existing framework in the mind of the host is especially susceptible to a new meme which fits the framework (such as a new precept by a religious leader being absorbed by a follower of that religion, whereas it would be ignored or escape notice by a non-follower). Suitable storage capacity, in memory or media, is necessary for the meme to persist, along with enduring vectors (e.g., the meme is literally chiseled in stone (such as the Rosetta stone) or reproduced in many, widely distributed copies of books or electronic media).

New research projects could provide a scientific and quantitative basis for memetics and an exploration of its prospective applicability and value, possibly discovering whether brief memes

such as “Death to America” or “Glory to the Martyrs” or “Winston tastes good like a cigarette should” are, in fact, cognitively and functionally different from non-memes such as “I like your hair,” or “please pass the salt”.



**Figure 2: Meme Transmission Replicates Claude Shannon’s Iconic Schematic of a General Communications System**

## Quantitative Basis

Since Dawkins’ revelation about memes, the concept has attracted a coterie of proponents, skeptics, and opponents. In more than thirty years there has been no significant research of the concept to establish a scientific basis for it – but neither has there been a definitive refutation. To progress as a discipline with useful applications, memetics needs a general theory – a theoretical foundation for development of a scientific discipline of memetics. It needs a narrowly-focused, pragmatically-useful definition, and, ultimately, the ability to make testable predictions and falsifiable hypotheses. The discrete meme must be defined, identified, and distinguished in the near-continuum of information, just as the discrete gene can be identified (more or less) in long string of DNA nucleotides (albeit, with current technology a gene may not be clearly identifiable). A quantitative basis for memes must be established, using, for example, such tools as information theory and entropy; genetic, memetic, and evolutionary algorithms; neuroeconomics tools such as functional magnetic resonance imaging (fMRI) and biochemical analyses; and modeling and simulation of social networks and information propagation and impact.

## Military Worth

If memetics can be established as a scientific discipline, its potential military worth includes applications involving information operations to counter adversarial memes and reduce the number of prospective adversaries while reducing antagonism in the adversary’s military and civilian culture, i.e., it could have the ability to reduce the probability of war or defeat while increasing the probability of peace or victory.

In the context of asymmetric or irregular warfare, information operations (IO) are of increasing importance for achieving victory (or avoiding defeat). Information operations consist of the integrated employment of the core capabilities of: electronic warfare, computer network operations, psychological operations, military deception, public affairs, and operations security. In concert with specified supporting and related capabilities, IO are deployed to influence, disrupt, corrupt or usurp adversarial human and automated decision making while protecting one's own. Potentially, memetics can have a major effect on psychological operations, military deception, and public affairs.

Psychological operations (PSYOP) are intended to induce or reinforce foreign attitudes and behavior favorable to the originator's objectives and to convey selected information and indicators to foreign audiences to influence their emotions, motives, objective reasoning, and ultimately the behavior of foreign governments, organizations, groups, and individuals. PSYOP focuses on the cognitive domain of the battlespace and targets the mind of the adversary. It seeks to induce, influence, or reinforce the perceptions, attitudes, reasoning, and behavior of foreign leaders, groups, and organizations in a manner favorable to friendly national and military objectives. It exploits the psychological vulnerabilities of hostile forces to create fear, confusion, and paralysis, thus undermining their morale and fighting spirit.

There are strategic, operational, and tactical PSYOP, as described in Joint Publication 3.53, *Doctrine for Joint Psychological Operations* (5 September 2003). Strategic PSYOP consists of international activities conducted by US Government agencies primarily outside the military arena but which may use DOD assets. Operational PSYOP is conducted across the range of military operations, including during peacetime, in a defined operational area to promote the effectiveness of the joint force commander's campaigns and strategies. Tactical PSYOP is conducted in the area assigned to a tactical commander, for a range of military operations, to support the tactical mission.

Psychological operations may occur across the spectrum of peace to conflict to war, integral to diplomacy, economic warfare, and military action, from negotiations and humanitarian assistance to counterterrorism. But according to *Information Operations Roadmap*, DOD (30 Oct. 03), despite PSYOP being a low-density, high-demand asset which is particularly valued in the war on terrorism, PSYOP capabilities have deteriorated. Well-documented PSYOP limitations include: an inability to rapidly generate and immediately disseminate sophisticated, commercial-quality products targeted against diverse audiences, as well as a limited ability to disseminate PSYOP products into denied areas. Remedial action is urgently required – and memetics may be a solution to recent PSYOP difficulties.

Military deception (MILDEC) involves actions executed to deliberately mislead the adversary's military decision makers about friendly military capabilities, intentions, and operations, thereby causing the adversary to take specific actions (or inactions) that will contribute to the accomplishment of the friendly force's mission. According to *Information Operations Roadmap*, DOD (30 Oct. 03), military deception should be one of the five core capabilities of IO and the value of military deception is intuitive. Counter-propaganda includes activities to identify and counter adversary propaganda and expose adversary attempts to influence friendly populations and military forces situational understanding. It focuses on efforts to negate,

neutralize, diminish the effects of, or gain an advantage from foreign psychological operations or propaganda efforts.

Public Affairs Operations (PA) assess the information environment in areas such as public opinion and attempt to recognize political, social, and cultural shifts. PA is a key component of information-based flexible deterrent options, intended to build the commanders' predictive awareness of the international public information environment and the means to use information to take offensive and preemptive defensive actions. PA is considered to be a lead activity and the first line of defense against adversary propaganda and disinformation (with the caveat that it must never be used to mislead the public, national leaders or the media). According to Field Manual FM 46-1, *Public Affairs Operations* (HQ, Department of the Army, Washington, DC, 30 May 1997), public affairs operations are combat multipliers in that they keep soldiers informed, maintain public support for the soldier in the field, and mitigate the impact of misinformation and propaganda.

Memetics also has potential military worth in supporting the military culture by enhancing recruitment and training. Recruitment may be improved with memetics by influencing the motivation of prospective recruits, enhancing the image of the military, increasing service awareness ("branding"), providing a national perspective and global situational context for serving one's country. Likewise, training can be improved by increasing trainee motivation, providing better explanations for the training, easing comprehension of the training components, enhancing retention of what is learned during training, and solidifying military culture for the trainees (traditions, customs, and mores).

## Prospects

Memetics can mitigate unfavorable consequences from an adversary's culture and help to deter conflict and reduce animosity through cultural education and generating acceptable solutions to endemic problems. In the case of combat, memetics can enhance tactics, strategy and doctrine by making an otherwise adversarial situation more acceptable to non-combatants, helping to minimize collateral damage by improving the ability to discern combatants from non-combatants, and encouraging civilians to identify insurgents and terrorists. At the conclusion of combat, memetics can bolster peacekeeping, occupation, and nation-building by making these operations more palatable to civilians and facilitating patience for the "long-haul." During post-combat with continuing insurgency, memetics can help minimize collateral damage by enhancing the ability to discern combatants from civilians and identify insurgents and terrorists.

## BIBLIOGRAPHY: MEMETICS-RELATED BOOKS

- Asher, Ely, *Disinfect Your Mind: Defend Yourself with Memetics*, Galiel.Net, 2005
- Aunger, Robert, ed., *Darwinizing Culture: The Status of Memetics as a Science*, Oxford University Press, 2000
- Aunger, Robert, *The Electric Meme: A New Theory of How We Think*, The Free Press, 2002
- Blackmore, Susan, *The Meme Machine*, Oxford University Press, 1999 (Paperback 2000)

- Brodie, Richard, *Virus of the Mind: The New Science of the Meme*, Integral Press, 1996 (Paperback 2004)
- Dawkins, Richard, *The Selfish Gene*, Oxford University Press, 1976 (Paperback 1999)
- Farrell, Winslow, *How Hits Happen: Forecasting Predictability in a Chaotic Marketplace*, HarperBusiness, 1997
- Gladwell, Malcom, *The Tipping Point: How Little Things Can Make a Big Difference*, Little, Brown, & Co., 2002
- Heath, Chip and Dan, *Made To Stick: Why Some Ideas Survive and Others Die*, Random House, 2007
- Lynch, Aaron, *Thought Contagion: The New Science of Memes*, Basic Books, 1996
- Poole, Steven, *Unspeak: How Words Become Weapons*, Grove Press, 2006

# COMPENDIUM EXCERPTS

Subjectively Excerpted By  
Robert Finkelstein

*Based on What Might Be of Interest to the Project*  
May 2008

## CRITIQUE AND DEFENSE OF MEMESIS, By Onar Aam, May 1996

- As a scientist who have been working a lot with these ideas I feel competent to outline a sparse list of fields which memetics should be developed upon: behaviorist sciences, phenomenology, communication theory, systems theory, Artificial Sociality (ASoc) and cognitive sciences. This list is far from complete, but it illustrates the diversity that is required to turn memetics into a science
- If you ask a geneticist what a gene is he can give you a quite elaborate technical definition. If you however ask a memeticist what a meme is she can give you a lot of examples, but no coherent definition.
- Information -- or rather culturally transmitted information -- is therefore the substance which memes are made of. This is how far traditional memetic ontology extends.
- Now, if the meme-concept is to have any scientific meaning whatsoever there must exist a consistent method of measurement.
- A taxonomy for memes does not exist.
- Another serious problem is what role consciousness plays in memetics.
- Hence measuring ideas is virtually impossible with current behaviorist methodology.
- Thus, the problem that memetics has to deal with greatly resembles that of nurture vs. nature. I.e. is the behavior of people environmentally encoded (i.e. cultural transmitted) or is it biological (i.e. emergent)? The answer is yes! (i.e., both).
- A culture with a superior technology also has the power to take over other cultures. The other cultures have two options: 1) surrender and perish, or 2) adapt to the technology of the enemy and survive. Since history is written by the survivors history is full of arms races, and in the long run the only winner is technology. As such we are undeniably slaves of technology whether we like it or not.
- As of today memetics is very far from fulfilling the requirements of a science. But I think it is just a matter of time before most of these problems have been resolved.

## A REPORT ON THE CONFERENCE: "DO MEMES ACCOUNT FOR CULTURE?" By Robert Aunger

- The possibility that it is behaviors rather than ideas which replicate was not specifically considered, although that is a position adopted by others within the memetics community (e.g., Derek Gatherer, William Benson). David Hull forcefully argued, however, that definitional uncertainties should not stand in the way of progress; it certainly did not in the case of genetics during the first part of this century.
- It was argued that memes can certainly influence genetic evolution. One need only think of lactose intolerance (an example discussed by Kevin Laland), where cultural practices (such as drinking milk) influence gene frequencies in a group. Likewise, genes may determine meme frequencies, if indirectly, through psychological biases against the adoption of some kinds of beliefs or values. The general consensus was that both replicators should be seen as coevolving in potentially complex ways.
- As parts of the brain, memes are intimately involved in highly distributed, highly structured and interconnected networks of neurons. Therefore, the means by which they are stored and evoked are highly contextualized. So, memes spend most of their time as parts of streams of thought which are continuously renewing, revising, and reassessing the constituent memes. It is the analysis of these more holistic streams of information which should be our primary focus. This can be accomplished.
- Initially, Richard Dawkins argued it was important to consider the evolution of memplexes such as religious doctrines. What happens if some component of the complex is missing? Does its ability to function degrade significantly?
- Another major set of issues concerns the psychology of memes. The question which dominated discussion in this workshop is whether memetics can proceed without a clear idea of what kinds of transformations memes undergo during storage and retrieval by brains. Can memetics leave the brain as a black box, and deal only with social transmission aspects?
- It was also remarked that there is a problem of circularity in the way memetics is generally conducted. Memeticists only study things which seem likely to follow a memetic process, like fashions and fads (e.g., the infamous backward baseball cap). The perceived success of such empirical adventures leads memeticists to self-congratulation. But many aspects of culture aren't small, isolatable bits of information or practices that readily diffuse in observable time. Take the example of language, which permeates every aspect of culture. How does memetics expect to explain these more fundamental components of culture?
- The word "meme" itself has problems. Its close parallel to "gene" may lead memetics astray, if in fact memes are not the same kind of thing.

- A general disappointment was the lack of discussion about what might be called "applied memetics." More time certainly needs to be devoted in future to thinking of ways to do memetics.

### **CULTURAL EVOLUTION OF “GUIDING CRITERIA” AND BEHAVIOR IN A POPULATION OF NEURAL NETWORK AGENTS, By Gianluca Baldassarre**

- An important form of cultural evolution involves individual learning of behavior by the members of a population of agents and cultural transmission of learned behavior to the following generations. The selection of behaviors generated in the process of individual learning requires some "guiding criteria".
- Culture can be defined as the total pattern of behavior (and its products) of a population of agents, embodied in thought, action and artifacts, and dependent upon the capacity for learning and transmitting knowledge to succeeding generations.
- "Direct bias" implies that the descendants directly test (some of) the cultural traits present in the population, and adopt the best ones. "Indirect bias" implies that the descendants adopt the cultural traits of the most successful individuals, while "frequency bias" implies that the descendants adopt the cultural traits with the highest frequency within the population. Finally "guided variation" implies that the behavior is transmitted by one generation to the other ("cultural transmission") and that the agents in the new generation go through a process of individual learning that modifies the acquired behavior before transmitting it to the next generation. Guided variation leads to the evolutionary emergence of new behaviors in the population ("cultural evolution") in a Lamarckian-like fashion.

### **MEMETIC MEANINGS – A COMMENTARY ON ROSE’S PAPER: CONTROVERSIES IN MEME THEORY, By Francis A. Beer**

- Memes may be like genes in certain ways, but they cannot be identical if only because of the differences between physical and cultural phenomena.
- The Darwinian theory of evolution was itself a memetic adaptation.
- 'Mind' has different meanings in biological, psychological, linguistic, and philosophical contexts.

### **LETTER ON: MEMES ON MEMES – A CRITIQUE OF MEMETIC MODELS, By Michael L. Best**

- I hold building a science of memetics as an admirable project. Slow progress towards such a science has been made as evidenced by much of the cited literature.

- It is best to test formal models on real empirical data. In this way predictions can be verified, fit can be measured, hypotheses can be tested. Sadly, empirical data is not always available. Alternatively, computational simulation and analysis can act as a surrogate for empirical data.

### **THE LIFECYCLE OF MEMES, By Henrik Bjarneskans, Bjarne Grønnevik and Anders Sandberg**

- Memes, self reproducing mental information structures analogous to genes in biology, can be seen as the basis for an explanatory model of cultural and psychological behavior.
- Memes were originally described by Richard Dawkins in his book *The Selfish Gene* (1976): *A unit of cultural transmission or a unit of imitation.*
- What makes the meme concept so powerful is its close analogies to the theory of natural selection. Natural selection occurs whenever the following conditions exist (Dennet 1990):
  - Variation: a continuing abundance of different elements.
  - Heredity or replication: the elements have the capacity of creating copies or replicas of themselves.
  - Differential "fitness": the number of copies of an element that are created in a given time varies, depending on the interaction between the features of that element (whatever it is that makes it different from other elements) and features of the environment in which it persists.
- Although people often make the decision to spread a meme or not consciously, this process is influenced by the meme. Some memes are viewed as important, and hence spread to others after a conscious and sometimes rational evaluation; some memes exploit aspects of cognition or emotion to bias their hosts to spread them. Natural selection favors memes that are good at reproducing, which suggests that in time there will exist many memes that are very efficient replicators. Their accuracy is irrelevant for their survival, only their ability to replicate and find new hosts; memes that interest people and encourage them to spread the meme will thrive at the expense of less attractive versions.
- In memetics, ideas are viewed as almost independent creatures in a symbiotic relationship with human minds and cultures.
- Memes do not only influence behavior to promote replication, but many of the most successful memes have other side-effects (for example, being able to invoke various emotions) or promote their replication by being useful or through other features (like parasiting on other memes, e.g. parodies and imitations); using a biological analogy one could say symbiotic memes spread mainly using their usefulness, while parasitic memes compel the host to spread them. This compulsion can be more or less subtle, ranging from explicit orders like in chain letters ("Send ten copies of this letter to your friends") to implicit influences that link with our attitudes like the "Save the whales" meme.

- Host: A host must be able to possess at least *the potential capacity* to elaborate on the meme and to perform those cognitive tasks connected to the meme that we normally refer to as "understanding". This means that only humans can be hosts (animals can perhaps become hosts for simpler memes, but we will not discuss this here), at least until the development of artificial intelligences reaches further.
- Vector: A vector is anything that transports the meme between hosts without the capacity to reflect on the meme. Examples are a wall, a voice, an email-program, or a picture. Can a human be a vector? Yes she can, if she lacks the cognitive capacity (or interest) to elaborate on a specific meme. Then she is just a non-reflective carrier of the meme, much the same as a book. Note though that the human vector is still a potential host - or *inactive* host (Grant, 1990) - for the meme, should she suddenly choose to analyze the meme (in its widest sense) or achieve the contextual understanding which would make this possible.
- Dawkins (1976) suggests three qualities of a meme that gives it a high survival value: longevity, fecundity and copying fidelity.
- Religions represent some of the most powerful and elaborate meme complexes in existence today; they have evolved over millennia into countless variants and co- evolved with cultures.
- Religions are often better than other meme complexes (such as science) at explaining how the world works on an emotional level. They provide answers to existential questions that are emotionally appealing, creating a satisfying world model (which then becomes intellectually satisfying regardless of its consistency due to cognitive dissonance). Because religions seldom try to empirically prove themselves they cannot be disproved, which further aids their stability. A religion can spread regardless of the truth or falsity of its claims.
- It is our intention that by now you, by reading this text, have been infected with one of the strongest memes on the planet: The Meta-Meme, e.g. the meme about the theory of memes. It is our sincere hope that you will tell your friends about this (yes, transmission and further infection) or maybe even let them read this paper.

## **MEMES, MINDS, AND SELVES, By Susan Blackmore**

- Memetic transmission is in some sense Lamarckian... Whether we see the process as Lamarckian also depends in part on what we consider the equivalent of the phenotype to be. If we follow some authors in persisting in seeing the organism as the phenotype then obviously the process is Lamarckian. If we follow Dawkins in treating meme products such as words, music, gestures, skills and fashions, as the equivalent of the phenotype, then the process still appears Lamarckian because these are the very things that are copied. However, we might follow his other suggestion that memes are still drifting clumsily about in their primeval soup. We can see the brain as the replicating machinery

for behaviors which have not yet created clear phenotypes. In this case the process is not so obviously Lamarckian.

- For many purposes we may proceed with a memetic analysis without specifying the size of the unit.
- Altruism may turn out to be yet another of the meme tricks that religions (those most powerful of meme-complexes) have purloined. Almost all of them thrive on making their members believe they are doing good.

### **WAKING FROM THE MEME DREAM, By Susan Blackmore**

- Religions are special because they use just about every meme-trick in the book (which is presumably why they last so long and infect so many brains). Think of it this way. The idea of hell is initially useful because the fear of hell reinforces socially desirable behavior. Now add the idea that unbelievers go to hell, and the meme and any companions are well protected. The idea of God is a natural companion meme, assuaging fear and providing (spurious) comfort. The spread of the meme-complex is aided by exhortations to convert others and by tricks such as the celibate priesthood. Celibacy is a disaster for genes, but will help spread memes since a celibate priest has more time to spend promoting his faith.
- Another trick is to value faith and suppress the doubt that leads every child to ask difficult questions like “where is hell?” and “If God is so good, why did those people get tortured?” Note that science (and some forms of Buddhism) does the opposite and encourages doubt.
- Finally, once you’ve been infected with these meme-complexes they are hard to get rid of. If you try to throw them out, some even protect themselves with last-ditch threats of death, ex-communication, or burning in hell-fire for eternity.
- The point I want to make is that these religious memes have not survived for centuries because they are true, because they are useful to the genes, or because they make us happy. In fact I think they are false and are responsible for the worst miseries in human history. No - they have survived because they are selfish memes and are good at surviving - they need no other reason.
- Once you start to think this way a truly frightening prospect opens up. We have all become used to thinking of our bodies as biological organisms created by evolution. Yet we still like to think of ourselves as something more. We are in charge of our bodies, we run the show, we decide which ideas to believe in and which to reject. But do we really? If you begin to think about selfish memes it becomes clear that our ideas are in our heads because they are successful memes. American philosopher Dan Dennett (1995) concludes that a “person” is a particular sort of animal infested with memes. In other words you and I and all our friends are the products of two blind replicators, the genes and the memes.

- Like religions, astrology is a successful meme-complex. The idea that Leos get on well with Aquarians is unlikely to survive on its own, but as part of astrology is easy to remember and pass on. Astrology has obvious appeal that gets it into your brain in the first place; it provides a nice (though spurious) explanation for human differences and a comforting (though false) sense of predictability. It is easily expandable (you can go on adding new ideas forever!) and is highly resistant to being overturned by evidence. In fact the results of hundreds of experiments show that the claims of astrology are false but this has apparently not reduced belief in astrology one bit (Dean, Mather and Kelly, 1996). Clearly, once you believe in astrology it is hard work to root out all the beliefs and find alternatives. It may not be worth the effort. Thus we all become unwitting hosts to an enormous baggage of useless and even harmful meme-complexes.
- There are two systems I know of that are capable of dismantling meme-complexes (though I am sure there are others). Of course these systems are memes themselves but they are, if you like, meme-disinfectants, meme-eating memes, or “meme-complex destroying meme-complexes”. These two are science and Zen. ... Science works this way because of its ideals of truth and seeking evidence. It doesn’t always live up to these ideals, but in principle it is capable of destroying any untruthful meme-complex by putting it to the test, by demanding evidence, or by devising an experiment. ... Zen does this too, though the methods are completely different. In Zen training every concept is held up to scrutiny, nothing is left uninvestigated, even the self who is doing the investigation is to be held up to the light and questioned. “Who are you?”

## **IMITATION AND THE DEFINITION OF A MEME, By Susan Blackmore**

- There are many ways of defining the meme but there are two that we should perhaps take particularly seriously. First, Dawkins, who coined the term meme, described memes as units of cultural transmission which "*propagate themselves in the meme pool by ... a process which, in the broad sense, can be called imitation*" (Dawkins, 1976 p 192). Second, the Oxford English Dictionary defines a meme as follows: "*meme (mi:m), n. Biol. (shortened from mimeme ... that which is imitated, after GENE n.). An element of a culture that may be considered to be passed on by non-genetic means, esp. imitation*". Both these definitions include the critical point that memes are cultural information that is copied, and that it is copied by imitation. The OED is arguably the most important dictionary of the English language and is, as far as I know, the first to include the word 'meme'. It would be unfortunate if memeticists began to use definitions of the meme that were incompatible with the dictionary definition, unless there were good reasons for doing so.
- I have no doubt that in the future memetics will become involved in discovering just what the cognitive processes are that are involved in teaching, learning, instruction, and any other kinds of cultural transmission, but for the moment I suggest that we may consider all of them as in some way being forms of imitation, or based on the ability to imitate.

- My argument has been that the definition of the meme depends on, and should depend on, the concept of imitation. Therefore, only those things that can be passed on by imitation should count as memes. ... This means we can immediately exclude many things that a few authors have confusingly included as memes, such as perceptions, emotional states, cognitive maps, experiences in general, or "*anything that can be the subject of an instant of experience*". Furthermore we can build on the long history of research in animal behavior to distinguish imitation from contagion, and from individual and social learning, and so to eliminate from memetics the catching of yawns or all the many things we each learn for ourselves, by ourselves.

## **THE FORGET-MEME-NOT THEORY, By Susan Blackmore**

### **➤ What the Experts Say about Memetics**

- "I think memes provide a neat way of explaining some of the paradoxes of cultural transmission and, if they help with explaining something that important, there had better be a science of them." Nicholas Humphrey, senior research fellow, London School of Economics
- "The idea of memes is a meaningless metaphor". Stephen Jay Gould, professor of zoology, Harvard University
- "A surprising number of people these days even talk about 'memes' ... Absurd as it may seem to imagine the seamless web of culture being disaggregated and transmitted between minds by gene-like replication, even some sociologists and philosophers have become captivated by the metaphor." Steven Rose, biologist, Open University
- "It is an empty and misleading metaphor to call religion, science, or any other human activity a 'virus' or 'parasite'". Memes are a "useless and essentially superstitious notion." Mary Midgley, philosopher, Newcastle.
- "Memetics needs to come up with supported, unique predictions and/or an existence proof to become valuable. The challenge to our speakers is to provide some support---either theoretical or empirical---for the meme hypothesis." Robert Aunger, anthropologist, Cambridge.

## **EDITORIAL: MEMES MATTER, By P. Bouissac**

- Although the *meme* hypothesis has been around for quite some time - it was first proposed by Richard Dawkins in 1976 - it does not seem to have had any decisive impact yet on the speculations and investigations of semioticians.

- As has been often the case in the history of scientific knowledge, an idea may inconsequentially float around for some time until it connects with other concepts or methods, usually apparently unrelated, to eventually precipitate, so to speak, and drastically change the way in which humans perceive their environment and themselves. Is the *meme* hypothesis of such a kind?
- Memes can become extinct for the same reasons as parasites. An examination of the life conditions of parasites thus sheds some heuristic light on memes' modalities of existence.
- Another important factor to be considered is that memes cannot be confused with evolved, wired-in, adaptive behavior. As any parasitic organisms, their biological dynamic exclusively concerns their own replication - in this respect they are strictly "selfish" to use the metaphor that Dawkins applied to genes - and consequently are neither detrimental nor beneficial to their hosts by necessity.
- As any organism, memes cannot be conceptualized independently of their environment. Adaptation by selection upon both genetic and memetic variations operates with respect to environmental factors. Catastrophic changes in memes' environment, that is the brain or a particular anatomical, physiological or chemical component of it, may eradicate a meme population, or trap it in a dead end without possibilities of further replications.

### **INTRODUCTION TO THE CRISIS OF THE MIND, By Richard Brodie**

- This is the most surprising and most profound insight from the science of memetics: your thoughts are not always your own original ideas. You catch thoughts—you get infected with them, both directly from other people and indirectly from viruses of the mind. People don't seem to like the idea that they aren't in control of their thoughts. The reluctance of people to even consider this notion is probably the main reason the scientific work done so far is not better known. As we'll see, ideas people don't like have a hard time catching on.
- Once created, a virus of the mind gains a life *independent* of its creator and evolves quickly to infect as many people as possible.

### **TOWARDS A COGNITIVE MEMETICS: SOCIO-COGNITIVE MECHANISMS FOR MEMES SELECTION AND SPREADING, By Cristiano Castelfranchi**

- The individual mind with its cognitive processes is the fundamental selection environment for memes. For the great majority of memes, it holds that: *if memes do not win within the mind of the individual, they cannot spread around in the population.*
- To be believed, something should be 'plausible'. Often we resist or reject a new item just because it is not 'plausible'. What is the basis of this kind of evaluation? Why is it a basis for rejection? ... Plausibility is *the credibility value assigned* to the coming item from

*inside*. It is evaluated just on the basis of previous knowledge: the same knowledge it has to be integrated with.

## CHAPTER 11 FROM “THE SELFISH GENE”, By Richard Dawkins

- Examples of memes are tunes, ideas, catch-phrases, clothes fashions, ways of making pots or of building arches. Just as genes propagate themselves in the gene pool by leaping from body to body via sperms or eggs, so memes propagate themselves in the meme pool by leaping from brain to brain via a process which, in the broad sense, can be called imitation. If a scientist hears, or reads about, a good idea, he passes it on to his colleagues and students. He mentions it in his articles and his lectures. If the idea catches on, it can be said to propagate itself, spreading from brain to brain. ... Memes should be regarded as living structures, not just metaphorically but technically. When you plant a fertile meme in my mind you literally parasitize my brain, turning it into a vehicle for the meme's propagation in just the way that a virus may parasitize the genetic mechanism of a host cell.
- Cultural transmission is analogous to genetic transmission in that, although basically conservative, it can give rise to a form of evolution.
- I have been a bit negative about memes, but they have their cheerful side as well. When we die there are two things we can leave behind us: genes and memes. We were built as gene machines, created to pass on our genes. But that aspect of us will be forgotten in three generations. Your child, even your grandchild, may bear a resemblance to you, perhaps in facial features, in a talent for music, in the color of her hair. But as each generation passes, the contribution of your genes is halved. It does not take long to reach negligible proportions. Our genes may be immortal but the *collection* of genes that is any one of us is bound to crumble away. ... But if you contribute to the world's culture, if you have a good idea, compose a tune, invent a sparking plug, write a poem, it may live on, intact, long after your genes have dissolved in the common pool. Socrates may or may not have a gene or two alive in the world today ... but who cares? The meme-complexes of Socrates, Leonardo, Copernicus and Marconi are still going strong.
- Faith cannot move mountains (though generations of children are solemnly told the contrary and believe it). But it is capable of driving people to such dangerous folly that faith seems to me to qualify as a kind of mental illness. It leads people to believe in whatever it is so strongly that in extreme cases they are prepared to kill and die for it without the need for further justification. Keith Henson has coined the name 'memeoids' for 'victims that have been taken over by a meme to the extent that their own survival becomes inconsequential ... You see lots of these people on the evening news ... Faith is powerful enough to immunize people against all appeals to pity, to forgiveness, to decent human feelings. It even immunizes them against fear, if they honestly believe that a martyr's death will send them straight to heaven. What a weapon! Religious faith deserves a chapter to itself in the annals of war technology, on an even footing with the longbow, the warhorse, the tank, and the hydrogen bomb.

## **SURVIVAL OF THE INSTITUTIONALLY FITTEST CONCEPTS, By Martin de Jong**

- Memetics, on the other hand, being a theoretically rich and promising framework badly needs empirical filling.

## **USING MEMETICS TO GROW MEMETICS, By David K. Dirlam**

- From an empiricist's point of view, empirical research in memetics is currently difficult to reproduce. No one has yet written the methodological recipe book for memetics research.

## **MEMES AND THE EXPLOITATION OF IMAGINATION, By Daniel Dennett**

- Evolution occurs whenever the following conditions exist:
  - Variation: a continuing abundance of different elements
  - Heredity or replication: the elements have the capacity to create copies or replicas of themselves
  - Differential "fitness": the number of copies of an element that are created in a given time varies, depending on interactions between the features of that element (whatever it is that makes it different from other elements) and features of the environment in which it persists.
- This is a new way of thinking about ideas. It is also, I hope to show, a good way, but at the outset the perspective it provides is distinctly unsettling, even appalling. We can sum it up with a slogan: A scholar is just a library's way of making another library.
- I don't know about you, but I am not initially attracted by the idea of my brain as a sort of dung heap in which the larvae of other people's ideas renew themselves, before sending out copies of themselves in an informational Diaspora. It does seem to rob my mind of its importance as both author and critic. Who is in charge, according to this vision--we or our memes?
- The first rule of memes, as it is for genes, is that replication is not necessarily for the good of anything; replicators flourish that are good at ... replicating! -- for whatever reason.
- A meme that made its bodies run over cliffs would have a fate like that of a gene for making bodies run over cliffs. It would tend to be eliminated from the meme-pool. . . . But this does not mean that the ultimate criterion for success in meme selection is gene survival. . . . Obviously a meme that causes individuals bearing it to kill themselves has a grave disadvantage, but not necessarily a fatal one. . . . A suicidal meme can spread, as

when a dramatic and well-publicized martyrdom inspires others to die for a deeply loved cause, and this in turn inspires others to die, and so on.

- The important point is that there is no *necessary* connection between a meme's replicative power, its "fitness" from *its* point of view, and its contribution to *our* fitness (by whatever standard we judge that). The situation is not totally desperate. While some memes definitely manipulate us into collaborating on their replication *in spite of* our judging them useless or ugly or even dangerous to our health and welfare, many--most, if we are lucky--of the memes that replicate themselves do so not just *with* our blessings, but *because of* our esteem for them. I think there can be little controversy that the following memes are, all things considered, good *from our perspective*, and not just from their own perspective as selfish self-replicators.
- Memes now spread around the world at the speed of light, and replicate at rates that make even fruit flies and yeast cells look glacial in comparison. They leap promiscuously from vehicle to vehicle, and from medium to medium, and are proving to be virtually unquarantinable. Memes, like genes, are *potentially* immortal, but, like genes, they depend on the existence of a continuous chain of physical vehicles, persisting in the face of the Second Law of Thermodynamics. Books are relatively permanent, and inscriptions on monuments even more permanent, but unless these are under the protection of human conservators, they tend to dissolve in time. As with genes, immortality is more a matter of replication than of the longevity of individual vehicles. The preservation of the Platonic memes, via a series of copies of copies, is a particularly striking case of this. Although some papyrus fragments of Plato's texts roughly contemporaneous with him have been recently discovered, the survival of the memes owes almost nothing to such long-range persistence. Today's libraries contain thousands if not millions of physical copies (and translations) of the *Meno*, and the key ancestors in the transmission of this text turned to dust centuries ago.
- Minds are in limited supply, and each mind has a limited capacity for memes, and hence there is a considerable competition among memes for entry into as many minds as possible. This competition is the major selective force in the infosphere, and, just as in the biosphere, the challenge has been met with great ingenuity. For instance, whatever virtues (from our perspective) the following memes have, they have in common the property of having phenotypic expressions that tend to make their own replication more likely by disabling or pre-empting the environmental forces that would tend to extinguish them: the meme for *faith*, which discourages the exercise of the sort of critical judgment that might decide that the idea of faith was all things considered a dangerous idea; the meme for *tolerance* or *free speech*, the meme of including in a chain letter a warning about the terrible fates of those who have broken the chain in the past, the *conspiracy theory* meme, which has a built-in response to the objection that there is no good evidence of the conspiracy: "Of course not--that's how powerful the conspiracy is!" Some of these memes are "good" perhaps and others "bad"; what they have in common is a phenotypic effect that systematically tends to disable the selective forces arrayed against them. Other things being equal, population memetics predicts that conspiracy theory memes will persist quite independently of their truth, and the meme for faith is apt to

secure its own survival, and that of the religious memes that ride piggyback on it, in even the most rationalistic environments. Indeed, the meme for faith exhibits *frequency-dependent fitness*: it flourishes particularly in the company of rationalistic memes.

- *Homo sapiens* have been around for half a million years. The first serious invasion of memes began with spoken language only tens of thousands of years ago, and the second great wave, riding on the meme for writing, is considerably less than ten thousand years in progress--a brief moment in biological time. Since memetic evolution occurs on a time scale thousands of times faster than genetic evolution, however, in the period since there have been memes--only tens of thousands of years--the contributing effects of meme-structures on our constitution--on human phenotypes--vastly outweigh the effects of genetic evolution during that period.

### **MEMES: MYTHS, MISUNDERSTANDINGS, AND MISGIVINGS, By Daniel C. Dennett**

- Some memes are like domesticated animals; they are prized for their benefits, and their replication is closely fostered and relatively well understood by their human owners. Some memes are more like rats; they thrive in the human environment in spite of being positively selected against--ineffectually--by their unwilling hosts. And some are more like bacteria or other viruses, commandeering aspects of human behavior (provoking sneezing, for instance) in their "efforts" to propagate from host to host. There is artificial selection of "good" memes--like the memes of arithmetic and writing, which are carefully taught to each new generation. And there is unconscious selection of memes of all sorts--like the subtle mutations in pronunciation that spread through linguistic groups, presumably with some efficiency advantage, but perhaps just hitchhiking on some quirk of human preference. And there is unconscious selection of memes that are positively a menace, but which prey on flaws in the human decision-making apparatus, as provided for in the genome and enhanced and adjusted by other cultural innovations--such as the abducted-by-aliens meme, which makes perfect sense when its own fitness as a cultural replicator is considered. Only the meme's-eye perspective unites all these possibilities under one view.

### **THE REVEALED POVERTY OF THE GENE-MEME ANALOGY: WHY MEMETICS PER SE HAS FAILED TO PRODUCE SUBSTANTIVE RESULTS, By Bruce Edmonds**

- Here I distinguish between what might be called the "broad" and the "narrow" approaches to memetics. The former, broad, approach involves modeling communication or other social phenomena using approaches which are evolutionary in structure. Work within this approach is often done without appealing to "memes" or "memetics" since it can be easily accommodated within other frameworks. In other words, it does not require an analogy with genetics. The later, narrow, sense involves a closer analogy between genes and memes – not necessarily 100% direct, but sufficiently direct so as to justify the

epithet of "memetics". What has failed is the narrow approach – that of memetics. Work continues within the broad approach, albeit under other names, and in other journals.

- I claim that the underlying reason memetics has failed is that it has not provided any extra explanatory or predictive power beyond that available without the gene-meme analogy. Thus whilst the idea of memes has retained its attractiveness for some in terms of a framework for thinking about phenomena, it has not provided any "added value" in terms of providing new understanding of phenomena. ... The ability to think of some phenomena in a particular way (or describe it using a certain framework), does not mean that the phenomena has those properties in any significant sense.
- Academics who seek to study memetics in serious ways have suffered in the respect that they are often confused with those on the penumbra for whom memetics is a fad. However, this mistake is grounded in an element of truth. The study of memetics has been characterized by theoretical discussion of extreme abstraction and over ambition. Thus for example, before any evidence is available or detailed causal models constructed, attempts have been made to "explain" some immensely complex phenomena such as religion in general or consciousness. This sort of discussion shifts any study of memetics from the realm of science to that of a philosophy...

#### **A JUSTIFICATION OF SOCIETAL ALTRUISM ACCORDING TO THE MEMETIC APPLICATION OF HAMILTON'S RULE, By John R. Evers**

- The essentials of kin selection and inclusive fitness are summarized according to a simple equation, called "Hamilton's Rule," which is expressed:  $C/B < b$ . "This says that the cost C (which is the loss in expected personal reproductive success through the self-sacrificing behavior) divided by the benefit B (the increase in the relatives' expected reproductive success) must be less than  $b$ , the probability that the relatives have the same allele," if the altruist gene is to survive natural selection.
- Any meme can be defined generally as a *rule of behavior*, encoded by functional neuronal groups or pathways. Behavior is action, whether mental or physical. Ideas such as *tying shoe-laces* or *opening a door* represent rules of physical action, i.e., rules of patterned neural-muscular interaction. Concepts such as *apple*, *seven*, or *causality*, represent rules of mental action, or rules of cognition, i.e., rules of patterned neural-neural interaction. Hence, physical movement is governed by memes which represent rules of physical action and thought is governed by memes which represent rules of mental action. ... If memes are units of natural selection (allowing for a memetic application of Hamilton's Rule); and, if memes copy horizontally (allowing the "memetic family" to extend beyond the scope of the genetic family); and, if memes can be directly responsible for altruistic behavior, then memetically driven and inclusively fit altruism can extend to the whole of any given (freely communicating) human population according to the mathematical purity of Hamilton's Rule.

- Hamilton's Rule offers a comprehensive justification for general (intra-societal or intra-cultural) altruism. Meanwhile, it can be said with newfound certainty that purely altruistic behavior is *possible* within any memetically fertile population.

### **STROLLING THROUGH THE MEMETIC MINE FIELD, By Leon Felkins**

- It is obvious, however, that many memes are "planted" in the group by certain individuals for their own reasons. The politicians, the schools, the press, the clergy and others are guilty of sowing a constant stream of memes within the populace and encouraging their growth.
- Can we think without being influenced by memes and genes? While we can't completely eliminate the influence of memes and genes we certainly can diminish their impact. Further, we can adopt the "Scientific Method" which requires that all conclusions be subject to verification. That is, we can keep an open mind for errors in our thinking and try to minimize accepting anything on blind faith.
- Funerals and other ceremonies surrounding a death are powerful and ancient memes that remain unchallenged in the most modern of communities. The United States has spent millions retrieving the bones of service men and others from foreign countries without a thought as to what could be in those bones that make them so valuable.
- Religion has had a great deal of success with memes. In fact, for anyone wanting to do research on the impact of memes, this is probably your best laboratory.

### **VIRAL MARKETING: SPREAD A COLD, CATCH A CUSTOMER, By Frank Fiore**

- They call it "viral marketing." Having someone else spread the word to drive more visitors to your site. ... Let's face it. The best kind of marketing is the kind you don't have to do yourself. Especially if you're a small business on the web with a limited advertising budget. Viral marketing is like it sounds. Call it word-of-mouth, spawning, self-propagation -- organic. ... Great new idea, right? Nope. Viral marketing has been around forever. Spreading the word through word-of-mouth was the world's first form of marketing. But the Internet has taken this organic form of marketing to new heights by making communications better and communities of people tighter -- thus making word-of-mouth even more effective.

### **MEME AND VARIATIONS: A COMPUTATIONAL MODEL OF CULTURAL EVOLUTION, By Liane M. Gabora**

- In order for a pattern of information to evolve, we need a way of generating variations of the pattern, and a way of selectively replicating it.

- Ideas, like genomes, are patterns that evolve; however their evolution is not subject to the same constraints, and employs different mechanisms. The generation of variation is less random in cultural evolution than in biological evolution; it reflects the accumulated knowledge of individuals, and the social structure in which they are embedded. In cultural evolution, selective replication is Lamarckian; an idea can be modified through experience after it has been phenotypically expressed. Genetic information is coded in a physical sequence of nucleotides, and contains the instructions for its own replication, whereas ideas are coded in patterns of neuron activation, and do not contain the instructions for their replication; they rely on human hosts to replicate them. Ideas that satisfy our needs or drives are preferentially learned and implemented.

### **A DAY IN THE LIFE OF A MEME, By Liane Gabora**

- Some believe that looking to biological evolution to gain insight into cultural evolution is a waste of time. As Gould put it: "Biological evolution is a bad analog for cultural change... biological evolution is powered by natural selection, cultural evolution by a different set of principles that I understand but dimly."
- Skeptics may wonder how we can develop a theory of cultural evolution before we understand how memes are instantiated in the brain. This situation has a precedent: Darwin came up with the theory of biological evolution through natural selection before the discovery of genes.
- The world can be viewed as a vast network of computations wherein information is created, transformed, and destroyed. This information often exhibits pattern, or statistical regularity that can be expressed mathematically.
- Today the Earth is embedded with artifacts like computer networks and circuses that cannot be accounted for by appeal to either the properties of matter or biological evolution. That is, biological evolution does not provide us with adequate explanatory power to account for the existence of computers any more than the properties of matter can explain the existence of giraffes. Computers are manifestations of yet another causal principle: the evolution of culture.
- Which memes spread and which ones die off reflects the dynamics of the entire society of individuals hosting them.
- Durham defines a meme as "any kind, amount, and configuration of information in culture that shows both variation and coherent transmission." Problems with this definition arise because it does not distinguish between cultural information as mental representation and cultural information as implemented behavior or artifact.
- What constitutes a meme ... will differ amongst individuals and within an individual over time.

- The Hamming distance between two memes is the number of bits that differ. (So the Hamming distance between 11111 and 11100 is two.) Since each meme has an antipode (for example, the antipode of 11111 is 00000), the space of all possible memes can be visualized as a sphere. The address of a meme is the information pattern that specifies where the meme is stored. ... The smaller the Hamming distance between two memes, the higher the probability that they will be retrieved simultaneously and blended together in the focus
- If  $L$  is the number of possible features in a meme, the number of possible memes is  $2^L$ . Assuming  $L$  is large the size this space is enormous, so the memory is sparse in that it stores only a small fraction of the set of all possible memes.
- Recall that to implement a meme is to express it, so that it crystallizes from the world of ideas into words or body language or objects in the physical world. Memes fool potential hosts into wanting or identifying with them, by aligning themselves with memes we already hold dear (as advertisers are well aware). The more we value ourselves in terms of the memes and implemented artifacts we possess or lack, the more vulnerable we are to ever-more-seductive forms of persuasion and advertising which tie up time, energy, and resources that could be applied toward other goals.
- One way to defend oneself against painful or manipulative memes is to construct ... a meme-immunological system; that is, formulate new memes specifically to deflect memetic antigens. However constructing memetic antibodies of this sort is time-consuming, and like any immunological response it has to be repeated every time the outside agent evolves a counter-response.
- Although the cultural evolution of memes operates through very different mechanisms from those of biology, culture is the only system comparable to biology, because it is the only other system to exhibit the imperative features of evolution adaptive exploration and transformation of an information space through variation, selection, and transmission. All patterns in the information we encounter can be traced to either (1) the physical constraints and self-organizing properties of matter, (2) biological evolution, (3) cultural evolution, or (4) interactions between these causal principles. ... One important difference between the two forms of evolution is that culture is less random--new patterns have a greater-than-chance probability of being more fit than their predecessors.

## **MEMES: THE CREATIVE SPARK, By Liane M. Gabora**

- Memes, as advertisers are well aware, can fool potential hosts into believing they are needed by associating with memes we already identify as necessary.

## WHY THE “THOUGHT CONTAGION” METAPHOR IS RETARDING PROGRESS OF MEMETICS, By Derek Gatherer

- Broadly, my intention is to show that what is probably the most popular, even one might say the 'orthodox', definition of the meme as a 'unit of information residing in a brain' ... presents us with serious philosophical difficulties that may hamper (and indeed are already hampering) the development of memetics as a science. I shall argue that an earlier and broader picture of the meme as a 'unit of cultural transmission, or a unit of imitation' ... is in many respects a better working definition, despite the fact that it is now generally regarded as obsolete.
- I shall focus on two main theses: (1) There can be no population memetics; (2) Memetics cannot be used to study why beliefs spread. In brief, thesis 1 is consequent on the impossibility of quantifying meme frequencies in populations, since this requires a clear meme-host relationship and that is precisely what we cannot obtain in most cases ... Thesis 2 is consequent on the impossibility of accurately identifying when belief has been replicated or transmitted. It is possible to accept thesis 2 while still rejecting thesis 1. However, if thesis 1 is accepted, thesis 2 is a necessary consequence.
- The central flaw in the 'thought contagion' and mind virus' hypotheses may be summed up in a single phrase: these theories require individuals to *have* memes. ... The state of having a meme is taken as synonymous with the state of having a certain informational pattern in the brain.
- Gene frequencies are of course abstractions, pure quantitative values, but they are abstractions which relate directly to a physical reality, i.e., the reality of gene sequences within the bodies of individuals. Thus, providing the required technical methods are available, one may derive an unambiguous estimation of gene frequencies which are comparable between one population and another. ... Memes, on the other hand, are more difficult to pin down. [In one definition] Dawkins included concepts, cultural artifacts and subjective states such as beliefs, whereas [in another definition] Dawkins restricted memes to units of neural information giving rise to behaviours or the production of artifacts ... Dawkins is careful to stipulate that when a meme is transmitted from one brain to another, it is not necessary that exact neural configurations are reproduced. Two individuals, who are exhibiting the same mental state, entertaining the same idea, performing the same behaviour, are taken to have the same meme, for all intents and purposes - even if their neural configurations are not identical. Thus at this most basic physical level, [in Dawkins latter definition] memes are less tangible than genes. They are merely informational specifications within brains.
- Even if one were able to use some method of scanning an individual's brain to discover the internal neural configuration, we would not be able to identify the Dawkins memes [as units of information] of that individual.
- While we *can* often observe the communication of information, we can *never* directly observe transmission of belief. *Information and belief are not the same kind of thing.*

- Nevertheless, behavior, and the artifacts produced by behavior, such as micropipettes, laboratories and cathedrals, *is* more easily quantified than conceptual abstractions. It is possible to measure the increase in the number of Christian churches per capita in the Roman Empire, to measure the number of scientific articles on the subject of Darwinism per total articles published, to measure the percentage of Aka hunters carrying crossbows, to measure the percentage of votes cast for political parties defining themselves as socialist etc. *These* are our meme frequency statistics, or as near as we shall ever get to such a thing. *These* memes are behaviours, or artefacts that are the products of behaviour, and *not* abstract informational instantiations in individual brains. ... And, crucially, individuals do not *have* any of these memes. They build them, say them, do them, make them, assent to them or deny them, but the memes are entirely outside the human beings that generate them. These meme frequency statistics are *not* per capita of human populations, and therefore *do not constitute a body of data which is formally analogous to that of population genetics*. For this reason, there can be no population memetics.
- One who spends so much effort criticizing the meme definitions of others, ought to present a clear target for his own critics. The following is therefore offered: ***Meme: an observable cultural phenomenon, such as a behaviour, artifact or an objective piece of information, which is copied, imitated or learned, and thus may replicate within a cultural system. Objective information includes instructions, norms, rules, institutions and social practices provided they are observable.***

## THE CASE FOR COMMENTARY, By Derek Gatherer

- The article 'Why the Thought Contagion Metaphor is Retarding the Progress of Memetics' argues that the field of memetics took a wrong turning in the early 1980s, with the result that all subsequent memetic theorizing has driven down the same blind alley. The wrong turning was to substitute the earlier, flexible definition of a meme (which in the article I term **Dawkins A**) with a far narrower and more specific definition as a unit of information in a mind or brain (termed **Dawkins B**). ... Since science generally abhors vagueness and thrives on precise definitions, such a move naturally seemed like a step forward. In this case, however, it was counter-productive, since:
  - Direct observation was relegated to a subsidiary role, while inference and speculation took centre stage. The derivation of internal memes from behavior relies on acts of the most naive inference. This is not the same situation as that of a classical geneticist of the early 20<sup>th</sup> century, since genes were not naively inferred from phenotypes, but were indirectly demonstrable as segregation ratios, patterns of independent assortment or linkage. There is no equivalent phenomenology in memetics. Internal memes are not even indirectly demonstrable. Their existence cannot be independently verified in any way.
  - Quantification became impossible. We cannot quantify any internal mental unit, and quantification is as essential as precise definitions (if not more so). While it is possible to quantify behavior, either roughly in the field or more precisely in the

laboratory, such quantification of behaviour does not translate into quantification of internal memes, unless the internal meme and the behaviour exist in a one-to-one correspondence. There is no evidence for this.

- The theory lacks explanatory power. To say that behavior can be explained in terms of something which cannot be observed is unhelpful. In any case, there is also no evidence that replication of behaviour is necessarily correlated with replication of internal mental states. We have no grounds for believing that thoughts are contagious, and such a view trivializes psychology.
  - Lynch's calculus of mnemonic conjugations is incompatible with any present thinking in either cognitive sciences or linguistics. The picture of the mind as a stack of 'awareness of statement x' and 'belief in statement x' would require the mind to store language in a manner reminiscent of a computer RAM. Such possibilities were effectively excluded by Chomsky in the early 1950s.
- I should therefore like to ask the commentators if they consider 'internal' memetics to be still tenable in the light of these criticisms. If so, how do they believe it should be constituted as an experimental science (as opposed to a pseudoscience dealing in postulated unobservable entities, where neither verification nor falsification are possible)? If the commentators agree with me that 'internal' memetics is untenable, do they tend towards my quasi-behaviorist version of 'external' memetics, or have they a third alternative?

#### **A PLEA FOR METHODOLOGICAL DARWINISM: a COMMENTARY ON ROSE'S PAPER – CONTROVERSIES IN MEME THEORY, by Derek Gatherer**

- Dennett ... takes the meme theory a stage further. Now the meme is not seen merely as a replicator within consciousness, but as the essence of consciousness itself. ... The illusion of consciousness is produced through the combined action of myriads of behavioral and abstract memes.

#### **MEMETICS: A SYSTEMS METABIOLOGY (Version 950220), By Ron Hale-Evans**

##### ***Intrinsic factors in memetic reproduction: "hooks":***

1. Intrinsically rewarding factors
  - 1.1. Promises of reward
  - 1.2. Interestingness of meme
    - 1.2.1. Promises of reward
    - 1.2.2. Aesthetic qualities
    - 1.2.3. Humor
    - 1.2.4. Strangeness/novelty
  - 1.3. Feelings of superiority

- 1.4.** Scapegoating
- 1.5.** Explanation of everything
- 1.6.** Usefulness
- 1.7.** Self-directed reward
  - 1.7.1.** Unconditional reward
  - 1.7.2.** For obeying
- 1.8.** Other-directed conditional reward
- 2.** Claiming intrinsically rewarding qualities
- 3.** Intrinsically punishing factors
  - 3.1.** Threats of punishment: fear
  - 3.2.** Guilt
  - 3.3.** Self-directed conditional physical harm
  - 3.4.** Other-directed conditional harm
- 4.** Mechanical factors in reproduction
  - 4.1.** Ease of reproduction
    - 4.1.1.** Simplicity
    - 4.1.2.** Comprehensibility
    - 4.1.3.** Reproductive ability of medium
    - 4.1.4.** Copying fidelity
      - 4.1.4.1.** Cohesiveness
      - 4.1.4.2.** Noteworthiness
  - 4.2.** Redundancy
  - 4.3.** Longevity
  - 4.4.** Plausibility
  - 4.5.** Adaptability
    - 4.5.1.** Syncretism
      - 4.5.1.1.** Subsuming other memes
      - 4.5.1.2.** Consistency with other memes: cooperation
    - 4.5.2.** Intolerance to other memes
  - 4.6.** Precise marketing
- 5.** Claiming mechanical factors
- 6.** General commands to host mind
  - 6.1.** Faith
  - 6.2.** Commands to explore the meme complex
  - 6.3.** Commands to spread the meme
  - 6.4.** Duty

### ***Extrinsic factors in memetic reproduction:***

- 7.** Input transducer
  - 7.1.** Isolation
  - 7.2.** Events consistent with meme
- 8.** Internal transducer
  - 8.1.** Stresses in population increasing suggestibility
- 9.** Channel and net
  - 9.1.** Fitness of channel and net --

- 9.1.1. High information capacity
- 9.1.2. Signal-to-noise ratio
- 9.1.3. Ease of use
- 10. Decoder
  - 10.1. Availability of decoder units for population
  - 10.2. Ability of population to interpret memes
- 11. Associator
  - 11.1. Suggestibility of population
- 12. Memory
  - 12.1. Presence of consistent memes
  - 12.2. Durability of memory
  - 12.3. Ease of storage and recall
- 13. Decider
  - 13.1. Favor of central authority
- 14. Encoder
  - 14.1. Intelligence of those propagating the meme
- 15. Output transducer
  - 15.1. Attractiveness of those propagating the meme

## **AN OPEN MIND IS NOT AN EMPTY MIND: EXPERIMENTS IN THE META-NOOSPHERE, By David Hales**

- This paper builds on previous work using the Minimeme model ... The model is extended to incorporate open-mindedness meta-memes (memes about memes). In the scenarios presented such meta-memes have dramatic effects, increasing the optimality of population distribution and the accuracy of existing beliefs. It is argued that artificial society experimentation offers a potentially fruitful response to the inherent problems of building new meme theory.

## **WHAT MAKES A MEME SUCCESSFUL? SELECTION CRITERIA FOR CULTURAL EVOLUTION, By Francis Heylighen**

- To be replicated, a meme must pass successfully through four subsequent stages: 1) assimilation by an individual, who thereby becomes a *host* of the meme; 2) retention in that individual's memory; 3) expression by the individual in language, behavior or another form that can be perceived by others; 4) transmission of the thus created message or *meme vehicle* to one or more other individuals. This last stage is followed again by stage 1, thus closing the replication loop. At each stage there is selection, meaning that some memes will be eliminated.
- The overall survival rate of a meme  $m$  can be expressed as the meme *fitness*  $F(m)$ , which measures the average number of memes at moment  $t$  divided by the average number of memes at the previous time step or "generation"  $t - 1$ . This fitness can be expressed in a

simplified model as the product of the fitnesses or survival rates for each of the four stages, respectively assimilation  $A$ , retention  $R$ , expression  $E$  and transmission  $T$ :

- $F(m) = A(m) \cdot R(m) \cdot E(m) \cdot T(m)$
- $A$  denotes the proportion of memes vehicles encountered (or memes independently discovered) by the host that are assimilated.  $R$  represents the proportion of these assimilated memes that are retained in memory. Therefore,  $A \leq 1$ ,  $R \leq 1$ .  $E$  is the number of times a retained meme is expressed by the host.  $T$  is the number of copies of an expression that is transmitted to a potential new host. Unlike  $A$  and  $R$ ,  $E$  and  $T$  do not have an upper bound, although  $E$  is likely to be more restricted than  $T$ . Note that  $F$  is zero as soon as one of its components ( $A$ ,  $R$ ,  $E$ ,  $T$ ) is zero. This expresses the fact that a meme must successfully pass through *all* four stages in order to replicate. Also note that for a meme to spread ( $F > 1$ ), you must have  $E > 1$  or  $T > 1$ .

### **MATHEMATICAL MODELS FOR MEMETICS, By Jeremy R. Kendal and Kevin N. Laland**

- The science of memetics aims to understand the evolution of socially transmitted cultural traits. Recently attention has focused on the interaction between memetic and genetic evolution, a phenomenon described as meme-gene co-evolution.

### **DIFFERENT TYPES OF MEMES: RECIPEMES, SELECTEMES, AND EXPLANEMES, By John Z. Langrish**

- This paper attempts to provide three directions for advancing Dawkins' original description of memes as 'units' of cultural transmission'. These three directions are
  - There could be different types of memes
  - Memes are not 'units'
  - The transmission of memes might not be restricted to something called 'culture'
- A biological perspective allows for different types of memes with different transmission mechanisms. The disease type mechanism (epidemiology) has been over emphasised and is only one such mechanism. Three types of memes are suggested and given the names recipemes, selectemes and explanemes.
- Much of what has been written about memes can be summed up by the title of Aaron Lynch's book (1996), 'Thought Contagion'. This approach has been used to attack both Communism and religion, provoking a counter attack by John Bowker whose 'Is God a Virus?' (1995) attempts to bring free will into the discussion. ... If some extra-galactic intelligence discovered our earth, they might be expected to be interested in carbon based life. However, if after twenty years they had only studied viruses and parasites we might

think they were deficient in curiosity or just weird. Why no interest in ants, elephants, spiders, trees, eagles, coral, sharks or fungi? It is a bit like that with memes; why such little interest in the memes of technology, poetry, design, economics and all those interesting human activities involving ideas. After all, Dawkins' original description of a meme included 'ways of making pots and building arches'.

## UNITS, EVENTS, AND DYNAMICS IN MEMETIC EVOLUTION, By Aaron Lynch

- In *The Extended Phenotype* Dawkins more explicitly clarified that "a meme should be regarded as a unit of information residing in the brain". I have since taken this meaning of the word "meme" and refined it into a more technical definition suitable for symbolic and mathematical analysis.
- ... The further development of strong theoretical treatments, formal definitions of terms, mathematical analyses, discussions of empirical methodology, and criteria for falsifiability remain subjects of interest among hard core scientists.
- Defining the word meme concisely but technically without reference to the other neologisms, we have: MEME: A memory item, or portion of an organism's neurally-stored information, identified using the abstraction system of the observer, whose instantiation depended critically on causation by prior instantiation of the same memory item in one or more other organisms' nervous systems. ("Sameness" of memory items is determined with respect to the above-mentioned abstraction system of the observer.)
- Some efforts have been made to incorporate the idea of "size" into the definition of the word *meme*. This includes those who favor the "smallest" possible units, as well as those who favor the "largest" possible units. Unlike those who expand the definition of *meme*, these scientists are indeed looking for ways to render the meaning more specific. Yet not all attempts at specificity are equally useful in the study of real phenomena.
- The sort of predictions generated by mathematical models and computer simulations are falsifiable. This provides a general method for attempting to falsify specific memetic hypotheses. In general, propagation parameters can be measured over the duration of the time interval being modeled, as can the initial host populations of the memes under study. If the propagation parameters remain in some specified interval, then the final host population plus or minus a calculable error margin should be measured at the end of that interval. A host population measurement outside those error margins would then falsify the model being used for the specific memes under study. The procedure can be quite elaborate, as it involves conducting surveys and performing difficult computations.
- At present, there is no brain scanning device that tells unambiguously whether someone holds a particular belief or not. Nor is there a memory probe that recovers all the communications that lead someone to adopt a particular idea. This means that measuring host populations and propagation parameters must depend on survey methodology. Questionnaires and interview protocols may thus form functional definitions of the

abstractions used in memetics. ... The rates of occurrence of specific events must also be measured for use in a quantitative model, in order to determine the propagation parameters.

### A RESPONSE TO PAUL MARSDEN OF A REVIEW OF *THOUGHT CONTAGION: HOW BELIEF SPREADS THROUGH SOCIETY*, By Aaron Lynch

- **MEME:** A memory item, or portion of an organism's neurally-stored information, whose occurrence depended critically on causation by prior occurrence of the same memory item in one or more other organisms' nervous systems.
  - This definition identifies the minimum conditions needed to achieve the recursive process (or algorithm) that forms the basis of evolution by natural selection in interpersonally transmitted brain-stored information. It becomes the basis of transmission event diagrams and differential equations ... useful in developing computer simulations of the meme transmission patterns ...

### DAWKINS BAD IDEA: MEMES, GENES, AND THE METAPHORS OF PSYCHOLOGY, By Arthur J. Marr

- In times past, if the devil didn't get you, the vapors would, and if they didn't, humours, poisons, bile or any number of fanciful entities would do you in. With Pasteur and 19<sup>th</sup> century biology, these agents of illness were replaced with microscopic organisms, and the invention of the disease model made it easy to attribute your aches and pains to malicious bacteria, viruses, or other little microscopic buggers. Of course, then as now, few people understand the actual biological processes that are responsible for disease, but the *metaphors* for disease do just fine, and have been duly incorporated into the common vernacular. Some may say that these metaphors have worked a bit too well, since they have made a Procrustean stretch to cover all sorts of behaviors, from alcoholism to gambling.
- The definition of a meme as an independent conceptual object is ultimately not simple, but simplistic, since it does not denote the web of informative relationships between behavior and the environment that is denoted by consciously and nonconsciously by the mind and body proper. For example, the concept of the sport of football is a well traveled meme to be sure.
- Football represents a rather involved information pattern that has infected the minds of young men nationwide, and football games, commentary, and assorted chatter has parasitized not only the minds of people, but the network airwaves, the written media, and many unwilling housewives. But is a football game an indivisible meme like entity, or is it somewhat different than the sum of its parts? Actually, the 'meme' of football is not a singular information pattern that replicates like a strand of DNA, but rather emerges from a web of separate patterns of information that are mediated not only by consciously perceived information but by neural and somatic activating processes that we otherwise

call emotion. The meme of football is not just a compendium of rules, but comprises the memory of the somatic responses that occur while watching (excitement, depression), the natural feeling of elation that occurs with a high state of alertness, the virtual extension of control over all those partisans of the losing team, the constantly changing and stimulating prediction error that occurs as one play after another unfolds, the smell and taste of hot dogs and beer, the camaraderie of friends, and so on. The meme of football is in other words a web of perceptual relationships that is volatile and constantly changing.

## **OPERATIONALIZING MEMETICS: SUICIDE, THE WERTHER EFFECT, AND THE WORK OF DAVID PHILLIPS, By Paul Marsden**

- One of the major challenges currently facing memetics is the issue of how to successfully *operationalise* the emerging paradigm. In other words, how can we exploit the innovative analytical framework of memetics in order to generate a body of theoretically informed empirical research? Whilst there are some important *theoretical* issues that have yet to be resolved, the future success of memetics, qua academic discipline, may depend not so much on elaborate theoretical developments, but on the results of *empirical* research findings. Operationalising memetics will not only involve subjecting memetic theory *itself* to empirical testing, but it will also mean assessing the usefulness of the paradigm in describing, understanding and explaining the sociocultural patterns and phenomena that are the *traditional foci of the social sciences*.
- "No fact is more readily transmissible by contagion than suicide." Emile Durkheim (Le Suicide [1897] 1951:141)
- Summary of empirical research on social contagions conducted by Phillips:

<b>The Werther Effect: Fact or Fiction - Summary of Research by D.P. Phillips</b>	
<b>Key findings and Conclusion</b>	<b>Source</b>
Suicide rates increased significantly after suicide stories were reported newspaper stories. The increase was proportional to the amount of newspaper coverage devoted to the suicide stories	American Sociological Review 1974 Vol. 39:340-54
Car accident fatalities increased following media representations of suicide implying that some car accident fatalities were in fact imitative suicides	Science 1977 196: 1464-65
Car accident fatalities, particularly those resulting from single-car accidents increased significantly three days after a suicide story was publicised in the newspaper press. The increase was proportional to the intensity of the publicity, and the age of the accident victim was positively correlated with the age of the suicide story victim. There was also a corresponding correlation between murder-suicide stories and multiple car crashes involving passenger deaths.	American Journal of Sociology 1979 Vol. 84 No.5: 1150 -1174
Publicised murder-suicides were followed by an increase in airplane crashes (airline and non-commercial). The increase in airplane crashes was proportional to the degree of coverage that these stories received	Social Forces 1980 Vol. 58 (Jun): 1001-1024
Daily US suicide rates increased significantly (for a period of less than ten days) following the appearance of highly publicised suicide stories on television	American Sociological Review

evening news programmes.	1982 Vol. 47: 802-809
Suicide rates, motor vehicle fatality statistics and non-fatal accidents all rose immediately following the transmission of fictional televised suicide stories in 1977	American Journal of Sociology 1982 Vol. 87 No. 6: 1340-1359
Homicides in the US increased following heavyweight championship prize-fights in a relationship that persisted after correction for secular trends, seasonal and other extraneous variables. The increase was found to be largest following heavily publicised fights.	American Sociological Review 1983 Vol. 48 (Aug): 560-568
Between 1973 and 1979 teenage suicides increased significantly following 38 nationally televised stories of suicide. The intensity of publicity devoted to the suicide stories was significantly correlated to its effect on teenage suicide rates	New England Journal of Medicine 1986 Vol. 315 No.11: 685-9

- ... In 1974, the sociologist David Phillips coined the term "the Werther effect" to describe imitative suicidal behaviour transmitted via the mass media. Phillips devised an empirical research programme to establish whether media reporting of suicide stories really did affect suicide rates. In over a decade of research, Phillips produced important evidence that supported the hypothesis that behavioural patterns in society can in fact operate as contagions.
- Phillips' methodological approach could be described as *quasi-experimental* in that his analysis was based on an experimental protocol, but he worked with exclusively historical data sourced in the real, uncontrolled, world. Using newspaper records and official suicide statistics, he identified a number of 'control periods' defined by the absence of front-page newspaper suicide stories. Using the suicide statistics from these control periods, he generated a number of expected suicide rates for a pre-defined selection of 'experimental periods' during which front-page suicide stories were published. Working with a null hypothesis that the front-page newspaper reporting of suicide stories had no effect on aggregate suicide rates, he compared the expected rates with the actual rates. After controlling for seasonal and other spurious effects, Phillips tested for significance between the two values, and by comparing a number of different experimental periods he was able to test for a correlation between suicide rates and the intensity of media representations.
  - Importantly, Phillips was not testing for the relationship between individual behaviour and media representations; rather he was testing for a relationship between suicide rates and media representations.
  - ... It is important to recognise that Phillips was testing for the replication of structural patterns in society rather than investigating the individual process of replication/transmission per se. In this way, his approach is an example of what might usefully be called *macro-memetics*, in contradistinction to the equally valid *micro-memetic* approach that currently dominates our paradigm.

- Following Phillips' research protocol, such a macro-memetic analysis can be broken down into a certain number of stages that together might outline a possible method for investigating the structural epidemiology of memes:
  - 1. Define the phenotypical expression/symptomatology that will be used to measure levels of meme infection (suicide)
  - 2. Measure the prevalence of meme infection over time within a given population (suicide rates)
  - 3. Measure the exposure rate within a population to this meme through a particular (mass) medium over time (circulation/viewing figures)
  - 4. Calculate an index of exposure intensity (exposure level multiplied by the share of total medium content (column length/no. of days on front page))
  - 5. Define a series of control periods where media transmission intensity = 0
  - 6. Define an experimental period where media transmission intensity > 0
  - 7. Regress meme infection levels during control period(s) to generate an expectation for the experimental period based on the null hypothesis that media representations of the meme have no effect on the incidence or prevalence of that meme
  - 8. Test for significance between the expected and actual results
  - 9. If expected and actual results are significantly different, test for further relationships (e.g., host similarity/correlation of intensity and suicide rates)
  - Such a research programme might yield results providing evidence to support the hypothesis that meme exposure partly determines the incidence and prevalence of meme infection. Put differently, Phillips' macro-memetic approach would help determine whether patterns of behaviour in society do in fact operate in a manner similar to that of contagions.
- A memetic approach, operationalised as *the epidemiology of social contagions in society* could build on research such as that conducted by Phillips, and could capitalise on the inability of current paradigms to adequately deal with the phenomenon of replication/imitation.

### **MEMETICS ON THE EDGE OF CHAOS, By Paul Marsden**

- The memetic stance involves taking a 'meme's-eyes view' of the social world, and analysing social phenomena as if human agents were vehicles for these replicating

cultural traits. ... Apply such a memetic stance to corporate culture and management theory, and in doing so develop a broadly memetic understanding of organisations.

- ... Companies are both constructed and constrained by the patterns of beliefs and rules (memes) that produce behaviour, and ultimately determine commercial performance. To change the patterns of performance, it is necessary to change the patterns of beliefs and rules (both explicit and implicit) that determine that performance.
- Rather than adopting the conventional view that individuals are the hosts for memes, this "*unconventional perspective*" holds that companies should be understood as the proper hosts for memes. Embracing 19th Century organicism (Spencer 1969), the authors suggest that organisations can be understood "*as 'living', self-organising, self-maintaining entities*", and these entities are the proper hosts for memes. Specifically, the meme concept is introduced as a heuristic device for understanding the components (sets of beliefs and rules) of the patterns which define the organic company.
- Rather than adopting the conventional view that individuals are the hosts for memes, this "*unconventional perspective*" holds that companies should be understood as the proper hosts for memes. ... Organisations can be understood "*as 'living', self-organising, self-maintaining entities*", and these entities are the proper hosts for memes. Specifically, the meme concept is ... a heuristic device for understanding the components (sets of beliefs and rules) of the patterns which define the organic company.
- [An unclear] definition of the meme: "*An organism is coded via chemical strands of DNA, sections of which form genes, the smallest units capable of being copied. An organisation is coded via 'ideas and images of the mind', abstract strands of thinking, perception and language, the smallest units of which can be thought of as memes which may be interpreted as: the smallest element capable of being exchanged, with an associated sense of meaning and interpretation, to another brain.*"
- ... whilst we are the products of our memetic inheritance, we also have the power to shift the patterns, to rebel against the tyranny of the memetic replicators, in a world where small changes can have major effects.

## **MEMETICS AND SOCIAL CONTAGION: TWO SIDES OF THE SAME COIN? By Paul Marsden**

- It is suggested that social contagion research, currently lacking a conceptual framework or organising principle, may be characterised as a body of evidence without theory. Conversely, it is suggested that memetics, now over two decades old but yet to be operationalised, may be characterised as a body of theory without evidence. [Propose] a memetic theory of social contagion ... social contagion research and memetics are indeed two sides of the same social epidemiological coin ... a call for their synthesis into a comprehensive body of theoretically informed research.

- The failure of mainstream social science to take ... social contagion evidence seriously is certainly in part due to the ... disorganised and incoherent state of the field.
- Not a fully blown theory, the memetic stance is more of a way of looking at the world, a set of guiding principles, and a useful heuristic, based on some hopefully important insight into the nature of the social world. Whether the memetic stance turns out to be an explanatory device in an evolutionary extension of folk psychology, or a proper theory of mind where memes are internally instantiated in the neural networks of our brains, is an issue that will one day have to be resolved empirically.
- So what exactly is the memetic stance? The memetic stance states that human condition is minimally defined by two selective processes operating in two different substrates, the biological and social ... This is because the necessary conditions for the evolutionary loop of replication, variation and selection are present in the two substrates. This is not contentious in itself, what is more contentious is that the memetic stance sees these processes operating at the level of what is being replicated, that is, the gene and the meme. Thus, the memetic stance involves taking a meme's-eye perspective and understanding of the social world, thinking not in terms of selfish genes, but selfish memes. Taking this memetic stance has allowed researchers to explain the spread of non-rational behaviour in terms of the fitness of that behaviour itself. Examples include altruism ... chain letters ... chain e-mail ... religions and cults ... political revolutions and war ... religious scriptures ... management practices ... media representations ... urban legends ... and consumer behaviour.
- Once we take the memetic stance, features of the world that are difficult to explain from the orthodoxy of traditional social science become non-miraculous and eminently explicable. The memetic stance can explain not only apparent design in the social world, but importantly it can also explain phenomena that seem to negate the omnipresence of individual agency in human affairs. Put simply, the memetic stance states that the reason why some social behaviour doesn't seem to make sense from the perspective of the individual is because we are looking at that behaviour at the wrong level. We are taking an anthropocentric or homuncular view of a social world that was created at least in part at a memetic level. Trying to explain the social world from the perspective of the individual is like trying to explain the movements of a car without reference to the driver.
- Taking the memetic stance involves, to use an overused concept, a true Kuhnian paradigm shift; just as evolution in the biological world evolves according to what is better (not best) for the gene in its environment, so too does the social world evolve according to what is better for the meme. The memetic stance involves describing, explaining and understanding social behaviour from this meme's-eye perspective. From the memetic stance "What makes this person want to do x?" becomes "What is it about x that makes people want to do it?" Social contagion can be explained by the memetic stance because culture has an independent evolutionary dynamic that is derived from the genetically evolved human capacity and predisposition to replicate culture. ... We have an evolved predisposition to replicating the behaviour of those around us. Successful social contagions are those elements of culture that operate as both stimulus and

response, and that are adapted to the evolved architecture of the human brain. No homunculus need be invoked, only evoked imitation.

- ... Memeticists could develop the social contagion research tradition of using the substrate neutral tools of epidemiology to assist their research programme. These tools could be adapted to provide useful information about differential incidence and prevalence of evolving cultural traits, as well as structure of endemic and epidemic features of society.

### **A STRATEGY FOR MEMETICS: MEMES AS STRATEGIES, By Paul Marsden**

- Memetics organised around the interactions of thought contagions ... is problematic because that interaction is neither observable nor measurable. This simple fact renders the thought contagion metaphor both conceptually vacuous and empirically redundant. It is certainly not by performing arbitrary mathematical manipulations and endowing the objects of introspection with contagious properties that will improve the status of a partial and misleading metaphor. ... The thought contagion metaphor persists largely because of what could charitably be called a 'disregard' for linguistics and cognitive science.
- Dawkins has recently declared that he is "alarmed" that his readers have taken his memetic speculations at face value as a theory of culture ... and a number of leading evolutionary psychologists ... have also made known their scepticism as to the scientific viability of the whole memetic enterprise. Of course, the fact that the founding fathers of memetics have doubts about memetics, qua scientific enterprise, is no reason in itself to abandon research. But surely we should be addressing their specific concerns, and those of other evolutionary theorists, rather than burying our heads in outmoded pseudoscientific introspectionism.
- ... Meme theory built on the thought contagion metaphor is simply bad strategy for the establishment of memetics as a respectable and respected research enterprise. If memetics wishes to be taken seriously, it must first explicitly address the established theories that already seek to explain the object of memetics' own enquiry and demonstrate *theoretically and empirically* how the new memetic approach is superior. Growing memetics in an intellectual vacuum on a staple diet of partial metaphors is not only foolhardy, it is intellectual suicide.

### **AN ALTERNATIVE APPROACH TO GENDER AND CONSUMER BEHAVIOR: MEMETICS, By TJ Olney**

- **Categories of Memes**

- Distinction Memes
- Association Memes
- Strategy Memes

- Resistance memes
- Observation
- Word of Mouth
- Parental Transmission
- Outside the family
- Mass media (include books)
- Evangelism
- Crisis
- Religious
- Commercial

➤ **Issues in meme spreading**

- Rate of Spread
- Intention
- Indirect meme transmission
- Susceptibility of host
- Preexisting Meme set
- Resistance memes
- Faith
- Skepticism
- Taboo
- Adversarial/Intimidation
- Hot Buttons
  - Primary Hot Buttons: the four F's
  - Secondary Hot Buttons
- Physical and emotional state

**THE LEARNING ORGANISATION MEME: EMERGENCE OF A MANAGEMENT REPLICATOR, By If Price and Ray Shaw**

- Organisations and organisms are self-maintaining systems which spontaneously seek to preserve an evolved order. Both are enabled by replicators: memes or genes respectively. Whereas genes are the units of transmission of our biological inheritance memes are the units of transmission of our cultural inheritance. They cause organisations to settle into patterns, routines and habits of behaviour: manifestations of a particular memetic inheritance. These patterns enable the organisation but simultaneously limit its performance. Both systems share the evolutionary dynamic of adaptive radiation followed by stabilisation. Memetic examples include new markets, new technologies and new business ideas. Business theories and their derivative, managerial fads, are a class of memes.

## **THE SELFISH TEXT: THE BIBLE AND MEMETICS, By Hugh S. Pyper**

- ... A hen as 'an egg's way of making another egg'.... An organism is a gene's way of making another gene ... *Western culture is the bible's way of making more bibles.*
- ... Dawkins becomes concerned to distinguish between a meme as a unit of information lodged in a brain and the phenotypic effects of that meme, such manifestations as the tune or the idea.
- Definitions are difficult, to say the least, in these areas. What exactly constitutes a meme or a culture defies classification and the recent literature on memes is bedevilled by shifting definitions and unsupportable generalisations and comparisons. Extravagant claims for the explanatory power of this concept have been made, including claims that the key to human self-understanding is in the new 'science' of memetics. ... Dennett elaborates the concept of the meme ... where he defines memes as 'self-replicating complex ideas which form *distinct memorable units*'. Even so, the concept remains notoriously fluid and therefore liable to abuse.
- Dawkins himself uses the 'mistranslation' of the Hebrew for 'young woman' as 'virgin' in the Septuagint version of Isaac 53 as an example of the potentially enormous phenotypic effects of a small change in DNA.
- The replicator must maintain its identity over time. Equally important, however, is its capacity to throw up variants which, when conditions change, may confer an advantage on the organisms which bear them. It is this balance between the ability to reproduce faithfully a particular variant but also to be able to produce variation if the circumstances favour it that confers reproductive success on any replicating system.
- Mere reproduction of a text as a physical artefact is not enough to ensure its continued survival ... Copies of books will only endure so long and the relative youth of even the earliest complete manuscripts of the bible bears this out.

## **DARWINIAN SOUP, By W.G. Runciman**

- The word 'meme', popularised by Richard Dawkins in *The Selfish Gene*, has recently gained entry into the *OED* as 'an element of a culture that may be considered to be passed on by non-genetic means, esp. imitation'.
- The controversies are of two kinds. For some opponents of 'memetics' in any guise, the mere idea of it is totally misconceived, whether because (as the Reductionists argue) cultural evolution is really all about the workings of our genes, or because (as the Creationists argue) it's really all about the souls implanted in us by God, or because (as the Post-Modernists argue) no so-called theory of cultural evolution will ever explain it any better than any other so-called theory. There are also not a few anthropologists and

sociologists for whom the mere association of 'memes' with the name of Dawkins is enough to rule them out of mention in correct academic society.

- The disanalogies with natural selection are obvious, and Susan Blackmore is well aware of them. There aren't, literally speaking, lineages of memes as there are of genes, and memes, unlike genes, can be discarded by the people who carry them and replaced by others.
- But what is - or isn't - a 'meme'? Blackmore sensibly eschews any attempt to find precise 'memetic' analogies for genetic concepts like alleles, meiosis and so forth. But having reviewed some of the many conflicting opinions in the current academic literature, she concludes that the best thing to do is to 'use the term "meme" indiscriminately to refer to memetic information in any of its many forms; including ideas, the brain structures that instantiate those ideas, the behaviours these brain structures produce, and their versions in books, recipes, maps and written music'. This is not only confusing, but defeatist.
- But if 'memetics' is, broadly speaking, about the way instructions affecting phenotype are transmitted from mind to mind by imitation and learning, it is presumably worth drawing an explicit distinction between the instructions transmitted, whether visually, orally or in written form, by parents, teachers, peer-group members or role-models, and the behaviour which the instructions dictate. Nobody can deny that such transmission is happening every day. Why, then, should even those whom Daniel Dennett calls 'Darwindreaders' object to the adoption of a single word for whatever specific bundles of instructions affecting phenotype - whether encoded in gestures, speech, writing or any other medium - do in fact guide the behaviour of the members of different cultures or sub-cultures?
- 'Memes' can include discrete units of memorable information (whatever they are) within the totality of values and beliefs which guide our cultural behaviour, as well as mental representations (however they function) of recipes, drawings, musical scores, prayer-books, folk maxims, manuals of etiquette and so forth. The different individual 'cultures' of the human populations in the ethnographic and historical record then consist of sets of related 'memes' (Blackmore calls them 'memeplexes') which have been handed down sufficiently widely and accurately over a sufficient number of successive generations to amount to a definable cultural tradition.
- If 'memetics' is to contribute anything significant to what is already well known to anthropologists, sociologists and historians of religion, it will have to be by identifying specific components of different religious 'memeplexes' whose probability of replication can be shown to have been enhanced by specific features of their carriers' environment - a rewarding task, but a very difficult one.
- Only detailed and wide-ranging re-examination of the relevant ethnographic, historical and archaeological evidence for the heritable variation and competitive selection of clearly identifiable 'memes' whose spread can accurately be traced and functions convincingly specified will turn 'memetics' from a project into a science.

## **DARWINIAN PROCESSES AND MEMES IN ARCHITECTURE: A MEMETIC THEORY OF MODERNISM, By Nikos Salingaros and Terry Mikiten**

- This paper applies the theory of memes to the field of architecture. Two main points are argued: (1) Darwinian processes (combining variation and selection) are important to architecture; and (2) the specific case of modernist architecture corresponds to a 'parasitic' meme, which has spread in spite of its being non-adaptive for the people that make use of modernist buildings. These two theses are logically independent, though both are necessary to present a picture of how architectural styles propagate.

## **MEMECOSYSTEMS: ARE ANIMAL MINDS SUITABLE HABITATS FOR MEMES? By Brent Silby**

- Memes are best thought of as sets of instructions that can be followed to produce behavior ... Instructions can be encoded in a number of formats, including:
  - Musical notation,
  - Written text,
  - Visible (or vocal) action,
  - Connectionist networks such as the neural structure of the brain.

## **EVOLUTION OF TECHNOLOGY: EXPOSING THE MYTH OF CREATIVE DESIGN, By Brent Silby**

- Many thinkers have resisted the idea of memes on the grounds that it is not possible to determine exactly *what* the meme is. We know that it is an idea that replicates, but how can we *point* to a meme? How can we isolate a memetic unit? This is a valid question. After all, there seems to be big differences between memes such as the first four notes of Beethoven's fifth symphony, the phrase "You're damned if you do, and you're damned if you don't", and the behavior of shaking hands. How can each of these be considered to be single memes? A similar problem arises when scientists talk about genes. There is no real gene 'unit'. Genes are packets of information that are encoded in DNA. Distinct genes can vary markedly in length. The same goes for memes. Memes are best thought of as packets of information that can be encoded in a number of different mediums. They can be encoded in the complex neural architecture of the brain, and they can be encoded in magnetic patterns on a hard-drive. A memetic 'unit' can be described as a *self-contained information packet that reproduces*. So, the first *three* notes of Beethoven's fifth symphony do not constitute a meme because it is not a self-contained information packet that reproduces. The *fourth* note is required to complete the reproductive unit. Of course, the question of the rest of the symphony now arises. Is it made up of a multitude of four note units? The answer is no. The symphony *is* made up of a vast collection of memetic units, but they are all different lengths. The entire symphony constitutes a memeplex that has good replicative power, but many of the individual memes would not make it alone. Some do, however -- you often find small pieces of music reproduced in other work --

but for the most part, the memes are dependent on their counterparts in the memeplex for survival. It is possible that the memes that comprise the fifth symphony memeplex were common in Beethoven's day. He was undoubtedly influenced by the tunes and musical themes of his time, and these would have found their way into his work. Even though they would no longer survive on their own, they manage to survive by being a part of the large symphonic memeplex that continues to reproduce.

### **A MEMETIC ANALYSIS OF POLICY MAKING, By Hans-Cees Speel**

- We argue that, in memetics, concepts are often still unnecessarily ambiguous and that thus the establishment of clear definitions is important.
- **Memes:**
  - Pieces of data that are a) copied from individual to individual without too much alteration, or B) that are interactors.
- **Interactor:**
  - An entity that interacts in an arena, where this action results in differential perpetuation of memes into a retention system.

### **THE REPLICATOR: A MISNOMER: CONCEPTUAL IMPLICATIONS FOR GENETICS AND MEMETICS, By Mario Vaneechoutte**

- ... The use of replicator for gene or meme is twice erroneous: they are not self-replicators (remark that 'replicator' is generally used to have the implicit meaning of self-replicator) and they even are not processor-replicators. Genes and memes are replicates or bits of replicable information. For memetics this means that we have to ask the question why some information is replicated more successfully by replicators like human minds and not why they replicate more successfully than other memes. This [is the] more scientifically correct approach ...

### **MERGERS AND TAKEOVERS; A MEMETIC APPROACH, By Ed Vos and Ben Kelleher**

- This paper constructively diagnoses problems within the current merger & acquisition (M&A) theories and provides an alternative theory of corporate behaviour. We contend that humans are the hosts for a replicating entity known as 'memes'. Since finance based motivational studies on M&A activities have not established that this activity 'adds value' to the acquiring firm, it is our thesis that certain managers gain power through mergers and acquisitions. Thus, M&A from the point of view of the acquiring firm can be seen as driving the evolution of ideas, shaping the flow of technology, information, and tastes rather than as 'value adding'. In simple terms, managers (the meme holders) use mergers and acquisitions to enhance their power, and in gaining this power managers

unconsciously provide an improved medium through which their memetic 'stories' may be replicated.

# **CRITIQUE AND DEFENSE OF MEMESIS**

By  
**Onar Aam**  
May 1996

In his article "Memesis Critique" Richard Barbrook [please see below] presents a massive critique of memetics and memesis. While great portions of the article are based on a rather uninformed and projective understanding of memetics Barbrook does make some valid points. One of them is:

"The other major error in the Memesis statement is its use of dodgy biological analogies."

## **Memetics as an idea**

Barbrook is correct that the concepts of memetics are not rigidly scientific. Currently memetics is little more than a pop-theory. Based on this many people draw the erroneous conclusion that memetics is a bad idea. This is not the case. The problem lies in the methodology and support fields rather than in the idea itself. Indeed, Darwin's own idea of biological evolution was at the time he conceived it nothing more than a promising idea. The reason it later became a rigid scientific fact is that Darwin was able to base his idea on existing scientific disciplines: botany, paleontology, geology and taxonomy. (Later also biochemistry and genetics) Without these disciplines as foundation Darwin's nice idea would be just that: a nice idea. Therefore Barbrook is only partially correct when he states that:

"Crucially, evolutionary theory was conceived in the first truly modern society: Great Britain. The rapid economic and social changes taking place in this first industrial society enabled Darwin to understand that nature itself was also in flux."

This alone was not enough. The crucial element which distinguished Darwin's idea from other neat ideas was that he had very respectable disciplines to back it up. As of now memetics lacks such supporting disciplines. This is partially due to the lack of coherence in the social sciences and partially due to memetics' reluctance/inability to coordinate its project with existing fields. I can't imagine any serious scientist who would object to the above critical characterization of memetics.

As a scientist who have been working a lot with these ideas I feel competent to outline a sparse list of fields which memetics should be developed upon: behaviorist sciences, phenomenology, communication theory, systems theory, Artificial Sociality (ASoc) and cognitive sciences. This list is far from complete, but it illustrates the diversity that is required to turn memetics into a science.

## **Ontology**

Genetics has a primitive, but well-defined ontology. The ontology of memetics on the other hand is a disaster. It is a disaster because most people in the field do not even realize that there is a problem, much less what the problem is. Even so I will try to briefly summarize it.

If you ask a geneticist what a gene is he can give you a quite elaborate technical definition. If you however ask a memeticist what a meme is she can give you a lot of examples, but no coherent definition. Typical examples of memes are catch phrases, pop songs, ideas, fashions etc. The first problem is that of substance, what memes are "made" of. Note that the above list of examples appear in a wide range of media and substances: sound waves, CD's, books, memories, computers, clothes etc. As you can see memes appear in virtually any medium or substance. The only thing these seem to have in common is "information". Information -- or rather culturally transmitted information -- is therefore the substance which memes are made of. This is how far traditional memetic ontology extends.

Now, if the meme-concept is to have any scientific meaning whatsoever there must exist a consistent method of measurement. Herein lay a major problem. How do we measure memes? Let me give you an example to illustrate the problem. Suppose we have a "red scarf with blue dots" meme. Then after a while people come up with variations of this fashion. They start wearing blue scarves with red dots, or a shirt instead of a scarf, stars instead of dots etc. Are these mutations new memes or are do we have a cluster of memes? We wouldn't know because a taxonomy for memes does not exist.

## **Consciousness**

Another serious problem is what role consciousness plays in memetics. In this matter the memetic community is approximately divided into two. 1) Those who ban consciousness (the behaviorists) and 2) those who haven't reflected over the role of consciousness. Behaviorist memetics is definitely the branch which has the best developed taxonomy. To a behaviorist memes are primarily stored as habits, and habit formation is something behaviorists have scrutinized for a long time. But unfortunately this method has great problems dealing with mental objects such as ideas. Surely ideas are stored in the form of habits, but without a consciousness observer included in the model it is impossible to determine the structure of the idea. Hence measuring ideas is virtually impossible with current behaviorist methodology.

In any case, no matter which branch they belong to few people in the memetic community address the problem of consciousness. In my opinion understanding the machinery of consciousness is as important to memetics as understanding the machinery of the cell is to genetics. If this is true then the realization of a memetic science is far away.

## **Emergence: the problem of nurture vs. nature**

One of the places where the role of consciousness comes most into play is in the grey zone between emergence and replication. Consider the following scenario: suppose a woman is called Sherrie Berry (This is a real person!). Many people who meet her comment that her name

reminds them of cherries. Even so this information is not something that has spread culturally. People seem to come up with this idea independent of each other. So then, is the cherry-association a meme or not? It is not a meme. What we have here is an example of "emergence". That is, the arising of a pattern due to the particular structure of the systems with which Sherrie interacts. This is equivalent to the emergence of e.g. waves. If you splash your hand in the water you will create a wave. If you splash it again another wave will emerge. But the waves you create are not replications. They are rather the internal dynamics of the system at hand. This is how it is with the cherry-association. Many people get the same association because they have common biological associative cognitive machinery that behaves similarly, not because the association replicates.

The complexity of this problem is perhaps made clearer if I say that consciousness is precisely a field of emergence. The above example is, of course, simple and obvious; but let us consider more complex phenomena. In the last century racism has popped up in suspiciously many places in one form or another (social Darwinism, scientific racism in the US and Soviet, Nazism in Germany, apartheid in South Africa etc.) A natural question to ask then is if this perhaps is a culturally replicated pattern, i.e. a meme. But as in the above example we will find that these ideas have often popped up rather independently of each other. Besides, why have racism become so popular and not e.g. shoe making? It is as if racism "strikes a chord" in many people, as if it emerges rather than replicates. Or that it also more easily replicates because it lies people's hearts more closer than shoe making? What we have encountered is the problem of human nature. Humans are by nature culturally discriminating beings and ideas of racism will therefore tend to pop up more often than others. Just like waves tend to arise in water, racism tends to arise in humans that encounter foreign cultures.

Another example to illustrate the point is the history of philosophy. If you compare the ideas of the many great thinkers you will often find extremely similar ideas across centuries, yes, even across millennia. Again these ideas often arise independently of each other and so again we are dealing with emergence rather than replication.

Thus, the problem that memetics has to deal with greatly resembles that of nurture vs. nature. I.e. is the behavior of people environmentally encoded (i.e. cultural transmitted) or is it biological (i.e. emergent)? The answer is yes! (i.e., both). There is an intimate interaction between emergence and replication. They are inseparable concepts.

This makes the task very difficult for memeticists. Ultimately memeticists have to map the structure of human nature in order to fully understand memetic evolution.

## System and Form

The problem of emergence is perhaps the best illustration of the ontological problems that memetics are up against. But this far from concludes the list of problems. In my opinion the most pressing practical and theoretical problem is that of system and form. The biological world is not just a soup of living creatures, but a highly structured dynamical system with many emergent forms and archetypes. If culture is an evolutionary arena then we should expect to see the same complexity and diversity of forms and archetypes as we see in the biological world. Naturally a

theory that tries to capture the evolutionary structure of culture should reflect this complexity. Does memetics reflect this complexity? No. The key structure of memetics is based on the concept of the selfish gene coined by Richard Dawkins. The problem with this idea is that it innately reduces the complexity of the biosphere by placing too much importance on the gene. The cell, the multi-cellular organism, the species, the ecosystem or in general the “organism” suffers badly under the selfish gene idea. Even so the idea is viable biology since there exists a multitude of biological fields that prevents the idea from going astray. But in the cultural sphere the idea stands alone without fields to support it or correct it. Therefore the idea has gone astray in the case of memetics. Whereas the virus is generally considered an amusing but peripheral phenomenon in biology, memeticists use it as their fundament. This in my view is the single most important problem with memetics. If a biological analogy should be used as a fundament for an evolutionary discipline in social science then that analogy should be built on well-established biological concepts rather than on a peripheral concept like the selfish gene. In the short term the selfish gene idea may have more popular appeal, but in the long run this is going to insure the death of memetics as a science.

## The Defense

Above I have presented a massive critique of memetics as a science. But memetics as a science is something totally different from Memesis as an idea. I will therefore present a defense of this vision. Barbrook neatly summarizes the essence of the Memesis statement:

"The central error of the Memesis statement has become obvious. It regards machines and information as autonomous things outside our control."

However, instead of accepting and embracing the idea he denies it:

"Yet, in reality, both technology and culture are expressions of the social relationships between individual humans."

What Barbrook fails to realize is that technology outlives us. Thousands of years ago individual humans started writing their thoughts and ideas down on paper. Those people are now long gone, but people today still write down their ideas and thoughts on paper. In other words, even though technology is an expression of the social relationships between people it tends to outlive the individual. Therefore inevitably technology and culture are to some degree autonomous things outside our control. Barbrook is wrong when he states that Memesis reiterates “the old McLuhanite lie that technology shapes our minds” because that “lie” has a ring of truth to it. The reason is that technology has a historical tendency to be used to promote one culture. A culture with a superior technology also has the power to take over other cultures. The other cultures have two options: 1) surrender and perish, or 2) adapt to the technology of the enemy and survive. Since history is written by the survivors history is full of arms races, and in the long run the only winner is technology. As such we are undeniably slaves of technology whether we like it or not.

Arms races are just one example of technological autonomy. A more peaceful example is Internet. Internet will probably outlive us all, and it certainly supersedes the power of

governments. No single government can control Internet. And as the net becomes increasingly more important for inter-cultural communication and business the only opportunity is to embrace it.

The following statement by Barbrook was meant as a rhetorical question but I will take the liberty to answer it:

"But, if there is active human involvement in the construction of cyberspace, how can technology and [memes] be autonomous forces outside our control?"

In my view this question is best answered by comparing culture to the multi-cellular organism. Like cells constitute the multicellular organism, people constitute culture. Just like the cells in an organism have a common genetic material, the people in a culture have a common language. Let me now rephrase Barbrook's question in terms of the organism-analogy: If there is cellular activity involved in the construction of the multi-cellular organism, how can it be an autonomous entity outside the control of individual cells? In biology we have no problems viewing the multi-cellular organism as an autonomous unity built up of other autonomous unities. Why then should the idea of culture as autonomous be any more problematic?

## **Conclusion**

As of today memetics is very far from fulfilling the requirements of a science. But I think it is just a matter of time before most of these problems have been resolved. The end result may not be as the current memeticists envision it, but it will certainly be an evolutionary founded social science. The lacks of current memetics does however not discredit the idea of Memesis, or what Francis Heylighen would call a Meta-System Transition.

## **MEMETICS AND MONTAG**

Five philosophical questions put on the web at 8 October 1996

You can refer to this interview as  
Speel HC, (1997) Le idee come entità materiali. In: La Transmissione delle idee. Montag, statica  
e dinamica delle idee.  
Antomarini, B; Biscuso, M; D'ancunto, G; Gozzano, S; Ferretti, F; Traversa, G (ed.). 35-39.  
Edizione Fahrenheit 451 Roma

This interview is published in MONTAG. A philosophical series of 8 books in Italian. There will appear two other interviews here when they become available. The Editors are young doctorates: Brunella Antomarini, Francesco Ianneo, Francesco Ferretti, Simone Gozzano, Guido Traversa, Giuseppe D'Acunto, Massimiliano Biscuso. Publishers: Federico Scanni and Catia Gabrielli. Publishing house: Fahrenheit 451, Campo de' Fiori 44 - Roma.

Montag is the hero of the novel Fahrenheit 451 (R.Bradbury), whose plot deals with a fictional society where books are all burnt. Montag, together with other clandestines, keeps the contents of books in mind in order not to lose them.

Brunella Antomarini at Brunella@uni.net, invited me to answer the five questions below. Questions about the volumes of MON- TAG can be mailed to her. The first volume is scheduled this year, but alas it will be in Italian only. Therefore I have the original English version right below.

### **Questions:**

1. Is the memes' theory a scientific theory? That is, could we say it is an analogical theory - meant to arrive to a biological law similar to the genetic one - or a metaphoric theory - using the gene-theory as just an explicatory aid?
2. But, according to this perspective, isn't the act of knowing totally replaced with the memes' representation of knowledge? If even knowing is just memes resulting stronger than others, isn't this theory itself such a representation? What is left of the relationship with concrete things?
3. Isn't the model of natural selection - and that of mutation - a weak means to explain the way fast-changing ideas are spread? Doesn't intention have a more plausible role in transmission of ideas?
4. What kinds of entities are memes? Are they material entities entering minds? If there's anything outside the mind, which is not biologically verifiable, how can they ever enter the mind?
5. Do memes imprint themselves on the brains? If so, isn't their diffusion intentional?

**Question:**

1. Is the memes' theory a scientific theory? That is, could we say it is an analogical theory - meant to arrive to a biological law similar to the genetic one - or a metaphoric theory - using the gene-theory as just an explicatory aid?

**Answer:**

First of all, let us look at the question. In general metaphor is used as being broader than analogy. In this way analogy seems to refer to a law in biology, and metaphor to an explanatory aid. I think it can be used as you please. In my opinion biology does not have many laws, and certainly no laws in the way physics has them. But aside from this, my efforts are not to arrive at any laws of biology. Memes are in the realm from behavior to thought so it depends how broad you take biology to be.

I think that in the common sense of the word, they are not in the realm of biology, since genetics has little to do with memes, apart for the fact that the things memes are in, are biological, and that parts of these things (things are organisms) can be determined by genes to a large or small extent to take up memes without too much interference by intentions.

I think that the most positive way to look at it is that the memetic perspective, both the virus-metaphor, and the selective metaphor, can be an aid to understand otherwise difficult to understand phenomena. That would be more metaphoric I think. This does not mean that the theory about the structure of thought is very analogical to biological theory, especially the Darwinian Theory. By the virus metaphor I mean the memetic perspective that Dawkins takes, where memes spread without much resistance.

By the selective perspective I mean the perspective where there is a focus on selection of memes; people can and do resist ideas, especially when those ideas prescribe them to behave in ways they don't like, or if those ideas tell them to agree on morally wrong ideas, in their judgment.

**Question:**

2. But, according to this perspective, isn't the act of knowing totally replaced with the memes' representation of knowledge? If even knowing is just memes resulting stronger than others, isn't this theory itself such a representation? What is left of the relationship with concrete things?

**Answer:**

No, the act of knowing is not totally replaced with the memes perspective of knowledge. The human is the one who does the knowing, the meme is replicated. It is possible that humans accept and use memes without knowing they do, but still the memes do not do anything. My phrase to state this is that 'humans do, memes don't'.

The theory of memes is such a representation, but memes are not only representations. In my opinion memes can be 'ways of thought', which can connote to a way to represent. But memes

can also be behavioral patterns. Mothers teach children, sometimes quite intentionally, not to steal money or candy [and so do fathers]. The pattern of behavior, or rather the forbiddance of a pattern, is replicated to the child by punishment, or otherwise. That is not really a representation.

In the same way, songs in the hit parade are replicated, and they are not a way of representation. In the same way, the attachment of emotions to certain stimuli is replicated by movies. We all cry if we see some scene from a touching movie. First of all because it reminds us of emotional events, but later this can be strengthened by the fact that everybody knows the movie, and the reference to that scene can become a 'sign' to feel the emotion. The most fascinating examples for me are the way memes shape government processes. Ways of financing are replicated, for instance because the actor that gives money requires you to order your finances in a specific way, so he can compare different requests for the money he supplies.

If you don't follow his instructions, he will not give you money. I am still thinking about more examples, and what the memetic perspective can add to what I already knew.

**Question:**

3. Isn't the model of natural selection - and that of mutation - a weak means to explain the way fast-changing ideas are spread? Doesn't intention have a more plausible role in transmission of ideas?

**Answer:**

The model of mutation is not weak. Let us look at biology. We have all kinds of mutations: point mutations are the first kind that people probably will associate with mutations. But we have chromosomal mutations, we have inversions, we have crossing over [technically not really a mutation], there are processes in which chromosomes are doubled, where double chromosomes are cut up again, etc, etc.

There are a lot of mutations, some very strong in changing genes, and some not. So I doubt if the model of mutation is weak. Take a textbook like Susuki, an introduction to genetics, or some title like that, and read it! You will be amazed how much of mutation is possible.

And now that I am talking about books, no philosopher that thinks about memetics can go through life without reading David Hull. Because David Hull is one who does not write too much, he writes his things with a great feeling for clear definitions, and knows about biology, and memetics.(Footnote about Hull). Now to the 'way fast changing ideas are spread'. First of all memes are not only ideas. But secondly the question is twofold; mentioned are spreading memes and fast-changing memes. These are two very different matters.

In populations of bacteria genes can spread very fast from individual bacteria to other bacteria, for instance by plasmids [little circle-shaped pieces of DNA, that have no direct function in mitosis etc, ].In the same way viruses can spread very fast through a host population. So this analogy is no problem.

Fast-changing memes are another matter. They are the thing you focus on when speaking about mutation-types. Ideas can stick together in complexes that are copied from individual to individual in one complex. That is the idea of a meme. But memes can be shuffled with other memes, in the brain, for some purpose, or without clear purpose. It is clear that the mechanisms working in this are different from the mechanisms of genetic re-shuffling. In many ways that are fundamental. That should be clear. But is it this process more extensive, are memes more radically altered than genes? I do not know how I should measure two processes that are so different in their mechanisms. Especially since we know so little about how ideas are processed in brains [the mind-body problem].

Intention, I would say, does play a role, but the question is strange. Tell me what intentions are; tell me what free will is. There are libraries full about these questions.

Second, as mentioned, memes are not only ideas.

Ideas can be re-shuffled by intentional thought. But they can also be re-shuffled by rather unconscious processes. That is, we normally don't know how we instantly see that ideas don't fit together. We don't know how creative processes work exactly. Should we call that intentional?

Another example. Do we choose what we see as good and bad? Do we really choose what we think of as 'proper', or 'morally good'? I think we have little choice in that. We seem to copy what we see as good from our parents, or from the papers. Furthermore, we do not think about the dichotomy of 'good and bad', we just use it. Why don't we use very good, a bit good, a bit worse, a bit bad, very bad, and dead-evil? Why do we use a dichotomy, and not a threechotomy, or more numbers?

We do not choose; we copy. And we keep doing that mainly because we do not see that we could choose. So do intentions play a part in memetic evolution? Yes, they do, but often they don't. Is it impossible for intentions to play a part, where they don't at a certain moment? Probably not, but still it doesn't happen as far as I can see. By the way, memes are of course transmitted by brains, along with some other things, but we cannot think that we automatically have intentional influence on what our brains transmit. The things brains do are not always intentional.

#### **Footnote:**

Hull DL (1980) Individuality and selection. Ann. Rev. Ecol. Syst., 11, 311-332

Hull DL (1982) The naked meme. In Plotkin HC (ed.) Learning Development and culture, essays in evolutionary epistemology. John Wiley & Sons.

Hull DL (1988a) A mechanism and its metaphysics: an evolutionary account of the social and conceptual development of science. Biology and Philosophy 3, 123-155

Hull DL (1988b) Interactors versus vehicles. In: Plotkin HC (ed) The role of behavior in evolution, MIT Press.

#### **Question:**

4. What kind of entities are memes? Are they material entities entering minds? If there's anything outside the mind, which is not biologically verifiable, how can they ever enter the mind?

**Answer:**

I think they are material entities. I say think, because I don't know how thoughts in general are located or translated in the brain, in patterns, material structure or anything like that. I think nobody knows. I believe that memes are in some way material, meaning that there must be a way in which memes, like any other thought, memory, etc are somehow stored in the brain. They are given a meaning, in the sense that they are connected to an environment of other thoughts, or at least concepts in which the make sense in the meaning of memes being ideas.

If memes are behavioral patterns that are learned, without a necessary translation in words, the storage could be different, or not, I do not know that. One thing is for sure, behavioral patterns, are copied by means of learning processes for instance from parents to children. If you teach your child not to be rude, you are teaching him a meme. If you do that by explaining it in words, you use words and thus ideas to help the copying process, if you just slap him if he does something wrong, as in the way you train dogs [no intention meant to say that you should] ideas are not needed, and straight conditioning takes over from words as memes. In this sense memes are not only a way to comprehend, but are material [if you include pattern as material] entities, that are real in a biological sense.

The question arises here if that is all they are. I think not. As in a biological theory, the way we describe what memes are is a method of comprehension, like any theory, and thus the theory of memes, or theories of memes are a way to structure phenomena in the minds of scientists or any other humans that think by means of theories. I do not think that one excludes the other.

**Question:**

5. Do memes imprint themselves on the brains? If so, isn't their diffusion intentional?

**Answer:**

How can memes enter minds? Well in the same way that sound enters minds. In the same way in which light enters minds. How? I think nobody knows. In all these cases the question is relevant, but I am not sure someone has the answer. In my book, the question should not be how it enters minds, that is, by the senses, but how the hell it is translated into understanding. How do we know what sound means? How do we interpret light patterns, when a big lion attacks us? How do we know that there is a lion in a case when somebody tells us? How is it that we understand those words, and are careful to stay away from the bars even if we cannot see that lion right away? I don't know. I do know that we can, and thus think there must be a way we do that.

Are there intentions involved? In general we over emphasize that. If we see the lion without knowing that he was there, and we were dangerously close, we would be instantly scared, and that memory would stay with us for a while. So the brain can be seen as being more ready to memorize some events, rather than other ones. But do memes actively imprint themselves?

I think not. They are imprinted because we interact with other humans, and are susceptible to their influence. But memes themselves are just sitting, being more, or less copied, dependent on the copying processes that humans are involved in. Does that mean that we are always intentionally copying? No, I don't think so.

# A REPORT ON THE CONFERENCE: “DO MEMES ACCOUNT FOR CULTURE?”

Held At King's College, Cambridge

By

**Robert Aunger**

Department of Biological Anthropology

University of Cambridge

CAMBS CB2 3DZ, UK.

rva20@cam.ac.uk

The following report summarizes the issues which arose in the workshops associated with the King's College conference on memes held in June 1999. It is organized along the general lines of academic disciplines, since that is the way in which the workshops were arranged. Details of a forthcoming book deriving from the conference (including a chapter by Boyd and Richerson) can be found at: <http://www.cus.cam.ac.uk/~rva20/Darwin.html>.

## 1. Evolutionary Biology

Given its origin in the work of the zoologist and evolutionary theorist Richard Dawkins, the memetics literature has continued to exhibit a strong flavor of evolutionary biology. Many of the problems with pursuing this line of research therefore tended to arise from consideration of memes as analogous to genes -- that is, as cultural replicators.

Where are memes? It was generally agreed that the concept of memes should not be restricted to being only thoughts or behaviors or artifacts, but should be generally conceived (i.e., substrate neutral) for now. The possibility that it is behaviors rather than ideas which replicate was not specifically considered, although that is a position adopted by others within the memetics community (e.g., Derek Gatherer, William Benson). David Hull forcefully argued, however, that definitional uncertainties should not stand in the way of progress; it certainly did not in the case of genetics during the first part of this century.

The relationship of memes and genes: It was argued that memes can certainly influence genetic evolution. One need only think of lactose intolerance (an example discussed by Kevin Laland), where cultural practices (such as drinking milk) influence gene frequencies in a group. Likewise, genes may determine meme frequencies, if indirectly, through psychological biases against the adoption of some kinds of beliefs or values. The general consensus was that both replicators should be seen as coevolving in potentially complex ways. It has been a failure of some work in memetics to believe that memes operate without constraint.

Mention was also made of the immune system as another replicator system, acting at a different temporal and spatial scale (within the lifetime of a single organism, and confined to that body). The implication is that we need not rely only on the gene analogy to understand cultural replication; a variety of replicator systems exist, the investigation of each of which can provide

insights and heuristic principles for memetics, as well as a general awareness that there is likely to be more than one way to accomplish any evolutionary goal.

A major problem -- which many adherents admit to troubling over for a long time -- is establishing how the genotype/phenotype distinction might work for memes. This functional difference in the genetic system has been generalized by Dawkins and Hull as the replicator/interactor distinction. Although it is possible for a replicator to serve both as replicator and interactor (as in the case of a ribosome, for instance), it is generally considered unlikely to persist, if only because replicators and interactors have fundamentally different roles to play in the evolutionary drama (as store of information and as survivor/transmitter, respectively). It is usually inefficient for the same entity to play both roles, so a competitor system with greater specialization would almost certainly win out in an evolutionary race, if only because a more specialized replicator would likely be more robust in its ability to duplicate itself. So it seems memeticists must develop a notion of a memetic phenotype, or "phemotype." While there are a number of contenders for this role, none has achieved a widespread tip of the hat.

Part of the problem with developing a rigorous notion of a memetic interactor is coming up with a criterion which surely identifies it as distinct from its progenitor, the memetic replicator. A related example is the prion: is this protein a replicator or interactor (or both)? David Hull put forward one criterion for making this distinction, which is generalizable regardless of substrate and replicator system (and thus a candidate for Universal Darwinism): the relative difficulty of reconstituting the replicator from an interactor. This is a generalization of the Weismannian notion that you can't go "backwards" from protein to gene. This inability arises because there tends to be some slop in the production of phenotypes: genes don't code for one phenotype, they code for a reaction norm, thanks to the impact of environmental conditions on development. So the relationship between replicators and their material instantiations is not simple (i.e., not one-to-one). This implies that information will be lost in the translation from meme to phenotype. It is this loss of information which makes the project of "reverse engineering" (or inferring the instructions from the product, as Susan Blackmore puts it) so difficult.

If one allows that memes can be present in artifacts, then memes such as ink on paper can be replicated with very high fidelity: using photocopiers, we have direct replicator-replicator reproduction, and consequently no loss of information. However, many memetic replication cycles seem to involve stages of translation from one form to another, and hence some information leakage. In particular, more "traditional" memetic life cycles involve brains. If a meme must pass through someone's head, the general inability of bits of brain to duplicate themselves directly means that they must travel between hosts to replicate. But to do this, they must be translated into another form for social transmission -- for example, as bits of speech -- since bits of brain don't themselves make the journey from one head to another.

The first aspect of the problem is that, per the above argument, such phenotypes are compromised as message carriers. This information loss means there must be a reconstitution of the message by its receiver: this is the famous Chomskian "poverty of the stimulus" argument concerning linguistic message-passing. But if there is significant reconstruction of the informational content of a meme by each host brain, then the likelihood of message replication is low, thanks to the vagaries of how each brain processes in-coming information (due to the

different background information individuals have acquired, the inferencing algorithms they use, etc.).

One way out of this problem, suggested by Dan Sperber (who also brought this problem to our attention), would be for the brain to have a general decoder -- a utility enabling it to reliably infer the intention of the sender, and hence the substance of the message. In fact, this is what the much-lauded "theory of mind" module is likely to be for (although this, in itself, is contentious; most cognitive scientists believe agents reason upon others' minds as a special application of "instrumental reasoning" with special application to social world, rather than granting such a capacity a special modular status). In this view, brains have evolved filters to assess the utility of information coming in from the social environment to keep us from rapidly being swamped with bad information (i.e., duped into stupid behaviors by people with ulterior motives). This normalizing inferential machine might also ensure the replication of memetic material during social transmission. However, its operation is unlikely to be perfect, so a high mutation rate remains a potential problem.

The need to communicate memes between brains also introduces another problem. If such psychological normalization of memetic inputs is important for communication to be successful, then memetic information is not, strictly speaking, inherited because it is not passed from person A to B. Instead, the similarity of socially-acquired information between individuals has another cause: inherently structured inferential processing by the brain. These reconstructive processes depend on a long history of genetic selection on the human cortex, not the passing of information from person to person in cultural lines of descent. In effect, the cause of the similarity between the information in A's and B's brains is thanks to evolutionary psychology, not memetics. Since the causes are different, one can expect the population-level dynamics (e.g., rates of mutation, types of selection) to also be different. This creates a fundamental problem for memetics as an inheritance process (the most general view on memetics).

Dan Dennett pointed out, however, that the memetic process -- even if dependent on error-correction routines in the brain to produce cultural similarity of beliefs and values -- still confers an evolutionary advantage. This is because the same information is acquired through transmission-plus-correction more efficiently and cost effectively than individual learning through trial-and-error could have done. Further, error correction is an important aspect of genetic inheritance as well, so replicator systems can operate with such assistance without having to be called something else.

It was also noted by Susan Blackmore that Dan Sperber's reasoning leads to the expectation that, if there is a cultural replicator, there should also be selection for improved mechanisms for its transmission over time. In this way, the reliance on reconstituting information from local resources each iteration would be reduced and the proportion of information actually being transmitted increased. Her presumption is that this is indeed what has happened during the major transitions in cultural evolution, such as language, writing, and computer-based communication. But whether these have increased the transmissibility of memes, or merely their copying fidelity, remains to be determined.

A somewhat different view of memetics was presented by Liane Gabora. She takes her cue from neuroscience and complexity theory. More particularly, her critique of the replicator perspective is that memes do not function in isolation, but as parts of a complex conceptual network or worldview; this should be the basic level at which cultural evolution is analyzed. In Gabora's vision, the tendency to see memes as discrete, identifiable units is problematic. As parts of the brain, memes are intimately involved in highly distributed, highly structured and interconnected networks of neurons. Therefore, the means by which they are stored and evoked are highly contextualized. So, memes spend most of their time as parts of streams of thought which are continuously renewing, revising, and reassessing the constituent memes. It is the analysis of these more holistic streams of information which should be our primary focus. This can be accomplished, Gabora argues, along lines she has pursued in her own work.

Initially, Richard Dawkins argued it was important to consider the evolution of memeplexes such as religious doctrines. What happens if some component of the complex is missing? Does its ability to function degrade significantly? Memeplexes are another instance of a transition in evolutionary complexity, which should be approachable using "major transition theory" as it is being developed by Eoers Szathmary, John Maynard Smith, Richard Michod and others.

## 2. Psychology

Another major set of issues concerns the psychology of memes. The question which dominated discussion in this workshop is whether memetics can proceed without a clear idea of what kinds of transformations memes undergo during storage and retrieval by brains. Can memetics leave the brain as a black box, and deal only with social transmission aspects? The virtue of ignoring psychology is that we can simply talk about inheritance processes at the population level and not worry about something we don't know too much about anyway: how the brain processes information. On the other hand, if memetics disregards psychology, and there are major transformative processes at work in the brain, then memetics is only explaining part of the cultural evolutionary process, and may therefore "get it wrong." It was felt by some (particularly Rosaria Conte) that no social theory, including memetics, can succeed without a proper psychological underpinning.

Related to this question is the relationship of memetics to imitation. Two interconnected questions pop up here: First, Does imitation require a complicated brain to do? This issue is important because it determines who gets to have memes: only complex intentional agents like people, or more lowly creatures without cortices, such as birds? Many (including Henry Plotkin) argue that there is no consensus concerning the psychological mechanisms of imitation. This is significant because, as Rosaria Conte says, you cannot define imitation without reference to the mental abilities involved, because using behavior as the sole criterion leads to confounds. For example, automatic contagion (such as yawning) is direct phenotypic copying without the inferencing of mental contents. Counting contagion as a kind of imitation suggests that agents don't need to correctly infer another's intention (plus her beliefs and needs, etc.) in order to adopt or imitate her behaviors. What psychological resources imitation demands remains an open question.

Second, Should memetic transmission be restricted to imitation? Susan Blackmore restricts memetics to cases of imitative behavior because, she asserts, only imitation serves as a direct copying process, and if memetics is to be founded on replication events, then only imitation can be counted as a memetic mechanism. But as we have just seen, the jury is still out on whether imitation is behavior copying or mental state inferencing (as assumed in the "theory of mind" literature). This leaves Blackmore's contention somewhat up in the air.

Because of the general discontent with imitation (due to its problematic psychological status), a satisfactory resolution of these interlocking issues was not achieved. However, it was generally felt that all social learning, rather than imitation alone, is a better psychological foundation for the cultural evolutionary process. The famous example of milk bottle-top opening by birds was presented as evidence. The pecking of bottle-tops has now gone on for many bird-generations, and spread through several European countries. Since it is generally felt that birds learn this bit of cleverness not by observing others, but by seeing opened bottle-tops, which inspires their own creativity (a process psychologists call "stimulus enhancement"), it seemed a pity to exclude such an example from the purview of memetics by limiting it to imitation-based diffusion.

However, if this liberal position on social learning is adopted, many repercussions ensue. For example, the phylogenetic history of memes suddenly becomes considerably longer, with birds and perhaps even more "primitive" creatures being allowed to have meme-based "protocultures." If so, then what distinguishes human culture from non-human culture? In addition, it means that direct contact between hosts is no longer required for memetic transmission, since the source of a meme (such as the tit which pecked a bottle-top) can be absent when a new, naive tit arrives on the doorstep. It is the artifact left behind -- that is, the pecked bottle-top itself -- which serves as the proximate stimulus for transmission of the pecking meme to the new arrival. This implies, in turn, that memetics must account for artifact production, since memes can be associated with these constructions, and not just brains. Such implications were not discussed at the meeting.

A second major point of contention was whether the memetic dynamic can be extended into the brain. Can we call individual learning a selection process just like the social transmission process? This proposal met with some disdain, and Henry Plotkin noted that at least among academic psychologists, this is definitely a minority position. Susan Blackmore argued that whatever is happening inside the head should not be considered part of the memetic process; if it is in fact selectionist, it still should be recognized as an independent replicator system. Others argued that including selection among alternative mental representations was crucial to a successful memetics. Two benefits were seen to result from this move by proponents. First, only through an analysis of mental properties and processes can good models of transmission mechanisms, such as imitation, be understood. Second, by extending the Darwinian process into the brain, the confusion of calling thinking --the manipulation of memes -- "directed," "intentionalist" or "Lamarckian" could be avoided. It was felt that the unpopularity of a selectionist psychology may be largely due to the appearance it gives of there being no room for human agency or decision-making, that all human psychology is merely a random selection process among alternative behavioral choices. Of course, the abandonment of intentionality and free will was hailed as a victory for memetics by the mental selectionists. But no consensus was reached on this issue either.

### 3. Social Science

From the perspective of this group, memetics is largely a promise at present, with no real results to show for itself. As such, the question is whether it will contribute anything new. It was felt that a quasi-epidemiological approach similar to memetics is already in widespread use in the social sciences. The idea that some cultures are more stable, or produce a higher quality of life because certain ideas spread better than others, has long been around.

What memeticists don't recognize, these critics argue, is that one can have a theory of cultural change which is not memetic because it does not involve social learning or the spread of particulate bits of information. For example, evolutionary psychology explains cultural change simply by invoking variation in what the environment stimulates people to recall. All the information for cultural processes is considered by evolutionary psychologists to be already in place in people's heads; the inheritance of these bits of information is genetic. What remains to be explained is not social transmission dynamics, but recall dynamics: what kinds of responses do different environments cause to arise? Thus, existing explanations of diffusion and spread exist which do not invoke memes.

Memeticists miss this "Big Picture" because they are largely unaware of the comprehensive literature which has accumulated in anthropology on cultural change, or the actual history of earlier views (e.g., the cultural diffusionists of the early twentieth century). This leads them to reinvent old wheels and postulate new terms for rejected ideas. Memeticists do not even recognize that the concept of culture -- the thing which memetics intends to explain -- is itself sufficiently problematic that some social scientists advocate its abandonment. The notion simply covers too complex and varied a set of processes to be useful in their view. (What exactly would replace the concept of culture, or what sub-concepts it should be divided into, is not obvious, however.)

What remains unclear to this group is the central claim of memetics: whether there is a novel replicator-based process underlying the population-level, epidemiological dynamic that is culture change. The primary problem of memetics, therefore, is whether there is a new entity on the horizon in whose interests things can be said to happen (the "meme's eye view"). This would be a new kind of function which a social institution might serve: that of the memes. As such, it would represent a real and novel alternative to group-level functionalism, or the various flavors of structuralist thought current in the social sciences. Unfortunately, this central claim has not yet been proven.

In sum, memetics is seen as simply another case of those from outside the discipline, in this case largely biologists, "having a go at explaining culture," but without taking into account many of the complexities this project is widely recognized by contemporary social scientists to entail. The meme critics are happy with the general notion that cultural change involves the diffusion of some vaguely characterized entity, but not with an explanation couched solely in terms of the selection, variation and inheritance of a particulate replicator.

It was also remarked that there is a problem of circularity in the way memetics is generally conducted. Memeticists only study things which seem likely to follow a memetic process, like

fashions and fads (e.g., the infamous backward baseball cap). The perceived success of such empirical adventures leads memeticists to self-congratulation. But many aspects of culture aren't small, isolatable bits of information or practices that readily diffuse in observable time. Take the example of language, which permeates every aspect of culture. How does memetics expect to explain these more fundamental components of culture?

The word "meme" itself has problems. Its close parallel to "gene" may lead memetics astray, if in fact memes are not the same kind of thing. It also produces a "revulsion factor" among those who would otherwise be friendly to the Darwinian cause. Memetics is perceived from outside as an arrogant usurper, making extreme, unwarranted claims. This only serves to put memetics in the same basket with a related attempt at explaining human social life, sociobiology, which was widely seen as what Dan Dennett calls "greedily reductionist." Sociobiology left no ground for social scientists to stand on, and all the interesting questions were subsumed under a single algorithm: the maximization of biological fitness. This is unpalatable to social scientists not just because of territoriality disputes, but because such a greedy reduction is bound to failure. Can all social processes really be reduced to selection and transmission? The box of concepts available from Darwinism doesn't impress this group. It seems a very small toolkit when so many theoretical alternatives are already available and there is so much complexity to explain. In fact, theory abounds in the social sciences. What is lacking is insight into real social processes. Explaining these seems a goal quite far removed from the concerns of most memeticists, who are laboring much further down the organizational hierarchy, worrying about replicators. An uphill battle against a wide variety of other approaches therefore lies ahead for memetics in the social realm.

#### **4. Final Remarks**

Many participants observed that despite the shared belief that an evolutionary approach to culture was necessary, significant barriers to communication remained between those from different disciplines. This perhaps derived from the varying histories these disciplines have with evolutionary approaches. In particular, social anthropology has a long history of such thought, which has generally not proven successful. Indeed, a common refrain among those social anthropologists participating in the meeting was "been there, done that." It was difficult for "believers" in memes to convince these historically mindful and hence reticent social scientists that this time around things might be different. Similarly, it was difficult for the anthropologists to explain exactly what went wrong previously, or specifically how the memetic perspective was likely to go wrong itself, even if given a clear run at explaining culture.

This incommensurability of ethos led to an undercurrent of dissatisfaction on both sides. One side seemed to feel that having to address the concerns of "non-believers" kept progress back, while the opposite side felt that the believers "just weren't getting it." Nevertheless, most agreed that bringing both sides together decreased the likelihood that proponents would engage in unchecked, hubristic claims about having explained culture (along with other conundrums such as consciousness), or that social anthropologists would continue to ignore the memetic alternative. Nevertheless, while I don't think anyone was persuaded to jump from one camp to the other, both sides did go away with a lot to think about, and increased respect for those who disagree with them.

A general disappointment was the lack of discussion about what might be called "applied memetics." More time certainly needs to be devoted in future to thinking of ways to do memetics. This should include discussion of existing empirical studies that don't go under the banner of memetics but which could be interpreted as falling within the general purview of this incipient discipline, as well as the development of methodologies for conducting specifically memetic studies in the future. This is because the ultimate test -- which would preempt theoretical objections -- is whether memetics can produce novel empirical work or insightful interpretations of previous results. Everyone agreed it has not yet done so, but must do so in the near future, given the extensive theoretical work already accomplished and the high level of current interest in the subject. Otherwise, it is likely that memetics will soon be perceived to be a failure. This might be considered unlikely if only because, as one participant remarked, just being able to assemble such an eminent, multidisciplinary group to discuss the topic underlines how these ideas are coming to have real force in contemporary intellectual discourse.

## **Workshop Participants**

Dr. Robert Aunger King's College Cambridge CB2 1ST

Dr. Michael Best, Media Laboratory E15-325 MIT 20 Ames Street Cambridge, MA 02139 USA

Professor Susan Blackmore Department of Psychology University of the West of England St Matthias College Bristol BS16 2JP

Professor Maurice Bloch Department of Anthropology London School of Economics Houghton Street London WC2A 2AE

Professor Robert Boyd Department of Anthropology University of California 405 Hilgard Avenue Los Angeles, CA 90024 USA

Professor Rosaria Conte Division "AI, Cognitive and Interaction Modelling" PSS (Project on Social Simulation) IP/Cnr, V.LE Marx 15 - 00137 Roma, Italia

Professor Richard Dawkins Department of Animal Behaviour University of Oxford

Professor Daniel C. Dennett Center for Cognitive Studies Tufts University Medford, MA 02155-7059 USA

Professor James E. Doran University of Essex Department of Computer Science Wivenhoe Park Colchester, Essex CO4 3SQ

Liane M. Gabora Center Leo Apostel Brussels Free University Krijgskundestraat 33, B-1160 Brussels, Belgium

Professor David Hull Department of Philosophy Northwestern University Evanston, Illinois 60208 USA

Professor Nicholas Humphrey Centre for Philosophy of Natural and Social Science (CPNSS)  
Tymes Court Building London School of Economics Houghton Street London WC2A 2AE

Professor Adam Kuper Department of Human Sciences Brunel University Uxbridge, Middlesex UB8 3PH

Dr. Mark Lake Institute of Archaeology University College London 31-34 Gordon Square London, WC1H 0PY

Dr. Kevin Laland Royal Society University Research Fellow Sub-Department of Animal Behaviour University of Cambridge Madingley, Cambridge CB3 8AA United Kingdom

Dr. Neil Manson Fellow in Philosophy King's College Cambridge CB2 1ST

Professor John Odling-Smee Institute of Biological Anthropology University of Oxford 58 Banbury Road Oxford OX2 6QS

Professor Henry Plotkin Department of Psychology University College London London WC1E 6BT, U.K.

Lord W.G. Runciman Trinity College University of Cambridge Cambridge CB2 1TQ

Professor Stephen Shennan Insitute of Archaeology University College London 31-34 Gordon Sq. London WC1H 0PY

Hans-Cees Speel School of Systems Engineering, Policy Analysis and Management Technical University Delft Jaffalaan 5 2600 GA Delft PO Box 5015 The Netherlands

Professor Dan Sperber Centre de Recherche en Epistémologie Appliquée Ecole Polytechnique 1, rue Descartes 75005 Paris, France

Dr. Richard Webb Centre for Philosophy of Natural and Social Science (CPNSS) Tymes Court Building London School of Economics Houghton Street London WC2A 2AE

## Acknowledgements

Thanks to Rosaria Conte, Liane Gabora, David Hull and Henry Plotkin for feedback on this report.

# CULTURAL EVOLUTION OF “GUIDING CRITERIA” AND BEHAVIOR IN A POPULATION OF NEURAL NETWORK AGENTS

By

Gianluca Baldassarre

Department of Computer Science, University of Essex

Colchester, CO4 3SQ, United Kingdom

gbalda@essex.ac.uk

## Abstract

An important form of cultural evolution involves individual learning of behavior by the members of a population of agents and cultural transmission of learned behavior to the following generations. The selection of behaviors generated in the process of individual learning requires some "guiding criteria". As with behavior, guiding criteria can be innate or originate from individual or social learning. Guiding criteria play a fundamental role in cultural evolution because they strongly contribute to determine the behaviors that will enter the pool of cultural traits of the population. This work presents a computational model that investigates the nature and function of some forms of "guiding criteria" in the cultural evolution of a population of agents that learn and adapt to the environment using neural networks. The model focuses on the interplay of individual learning and cultural transmission of behavior and those forms of guiding criteria. The model contributes to clarify the nature and role in culture evolution of the guiding criteria studied. Also, within the assumptions of the model, it shows that the cultural transmission of behavior is more effective than the transmission of the guiding criteria.

**Keywords:** Multi-agent simulation, neural networks, cultural evolution, cultural transmission, reinforcement learning, imitation, behavior, guiding criteria, values, evaluations.

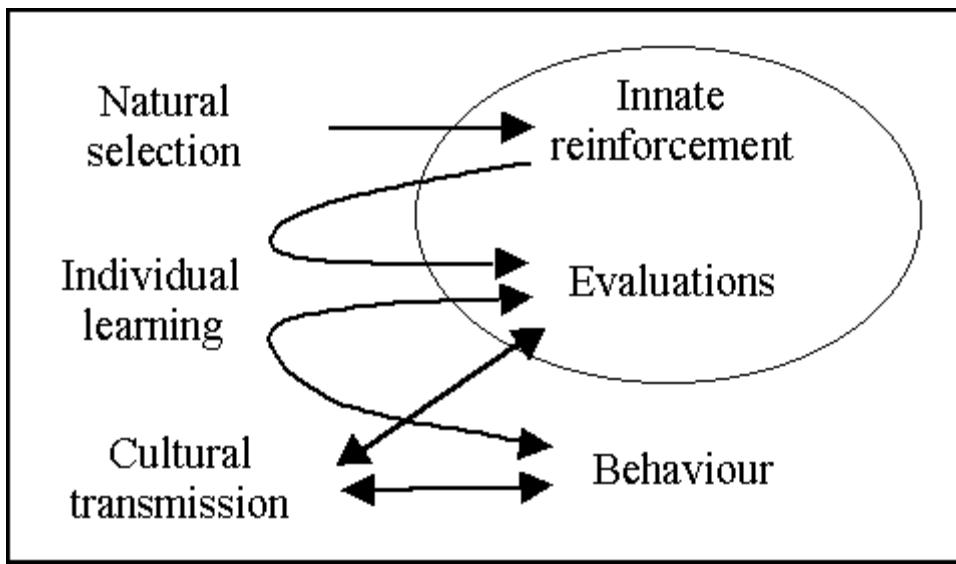
## 1. Introduction

This work investigates some phenomena related to cultural evolution. Culture can be defined as the total pattern of behavior (and its products) of a population of agents, embodied in thought, action and artifacts, and dependent upon the capacity for learning and transmitting knowledge to succeeding generations (cf. Cavalli-Sforza and Feldman, 1981, page 3). Boyd and Richerson (1985) have presented some mathematical models that specify some of the forms with which cultural evolution can take place. "Direct bias" implies that the descendants directly test (some of) the cultural traits present in the population, and adopt the best ones. "Indirect bias" implies that the descendants adopt the cultural traits of the most successful individuals, while "frequency bias" implies that the descendants adopt the cultural traits with the highest frequency within the population. Finally "guided variation" implies that the behavior is transmitted by one generation to the other ("cultural transmission") and that the agents in the new generation go through a process of individual learning that modifies the acquired behavior before transmitting it to the next generation. Guided variation leads to the evolutionary emergence of new behaviors in the population ("cultural evolution") in a Lamarckian-like fashion. Guided variation is the only form of cultural evolution investigated in this work. Denaro and Parisi (1996) have studied another

type of process that yields cultural evolution in a Darwinian-like fashion, and is similar to the "indirect bias" mechanism. The individuals of the population are ranked according to their success, and the most successful ones are used by the new generation as "cultural models" to imitate. Baldassarre and Parisi (1999) have compared this process of cultural evolution with guided-variation.

Boyd and Richerson (1985, page 132 and 136) underline the importance of "guiding criteria" in the process of individual learning and cultural evolution. They argue that guiding criteria are things like the "sense of pleasure and pain that allows individuals to select among variants", where these (behavioral) variants are generated in the process of individual learning or observed in other individuals. Guiding criteria play a central role in cultural evolution because by determining which behaviors are kept and which are discarded during individual learning, they strongly affect the cultural traits that enter the population's pool of traits through guided variation (this is the focus of this work). The same guiding criteria are also used for the selection of cultural traits in other forms of cultural evolution, such as direct bias (this is not investigated here).

Despite the importance of guiding criteria, Boyd and Richerson do not give a precise description of their nature and origin. They only say that they "could be inherited genetically or culturally or learned individually". Guiding criteria, as their name hints, are a complex compound concept that will require much investigation to be fully understood. The goal of this paper is (a) to contribute to this investigation by presenting a computational model that, by drawing from the biological models of animal learning, defines the nature of some guiding criteria and (b) to study the role of such guiding criteria in the cultural evolution of behavior.



**Figure 1:** Relationship between the processes and entities investigated in this work. A circle highlights the guiding criteria. Bent arcs indicate relations mediated by individual learning.

The processes studied in this work are summarized in figure 1. The model used in this investigation mimics a population of artificial agents that adapt to the environment by learning to search for "good" food and to avoid "bad" food. This process of learning is based on an innate

capacity to judge the food as good or bad tasting. The capacity to search and avoid food can also be acquired by imitating other agents. Two kinds of guiding criteria are studied in the model. The first is "reinforcement", widely studied in animal learning literature (Lieberman, 1993, for a review). Reinforcement roughly corresponds to an internal neural activation of the agent's brain associated with pleasure or pain. The neural mechanisms underlying this activation are mainly innate (Rolls, 1999). The second kind of guiding criteria are "evaluations" of the perceived state of the world. An evaluation is an internal neural activation of the agent's brain that quantifies the potential of that state to deliver reward in the future. As with the searching and avoiding behavior, the capacity to express correct evaluations can be learned individually or socially from other agents. As we shall see, the capacities to perform correct evaluations and to exhibit adaptive behavior are closely related. The study focuses on the effects that the cultural transmission of evaluations and behavior produces on the level of adaptation of the whole population.

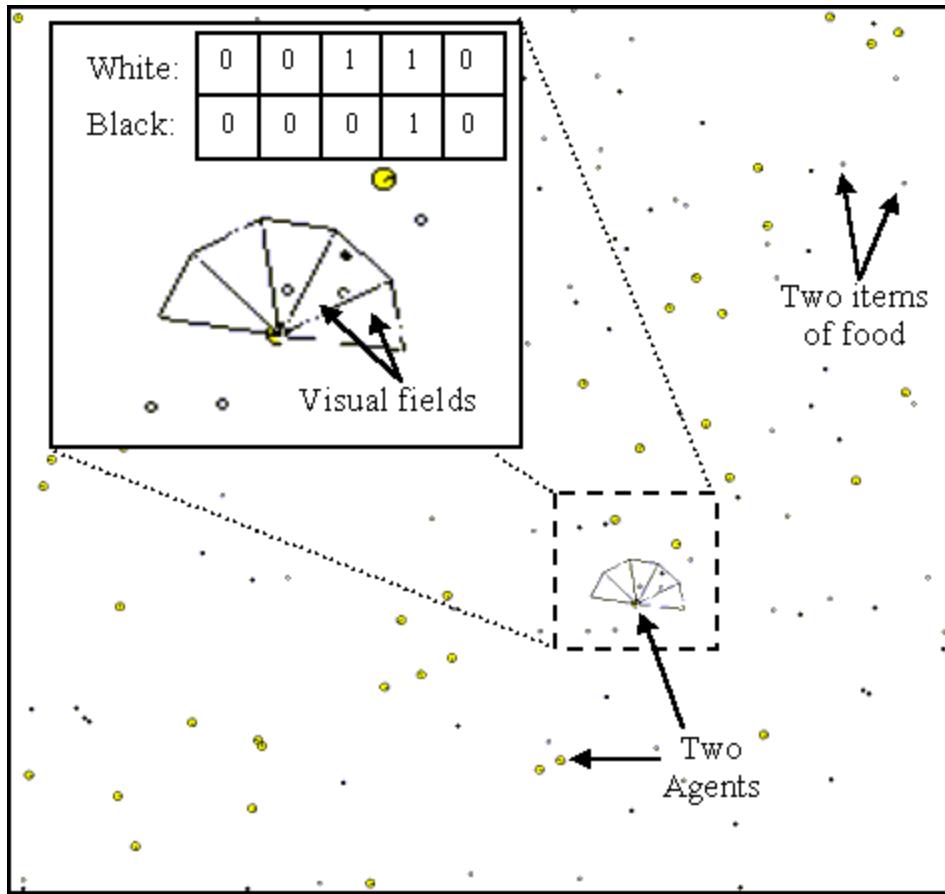
Section two describes the computational models used in the simulations and in particular the neural network controlling the agents and the algorithms mimicking individual learning, imitation, and the transmission of evaluations. Section three presents the results of the simulations and their possible interpretation. Section four describes some related work and, finally, section five draws the conclusions.

## 2. Methods

### 2.1 The environment and the task

The environment of the simulations is a square arena with sides measuring 1 unit. Within the arena there are 50 items of white ("good tasting") food, 50 items of black ("bad tasting") food, and a population of 50 agents randomly placed (figure 2). When an agent steps on an item of food, the food is consumed and a new item of the same kind is reintroduced at a random location of the arena. Each agent has a one-dimension "retina" of 5 non-overlapping sensor pairs that receive information from a 180xfb frontal visual field. Each pair contains one sensor sensitive to the color white and one sensitive to the color black. In each cycle of the simulation an agent perceives the environment with its sensors and selects and executes one of three actions: going forward left, going straight, going forward right.

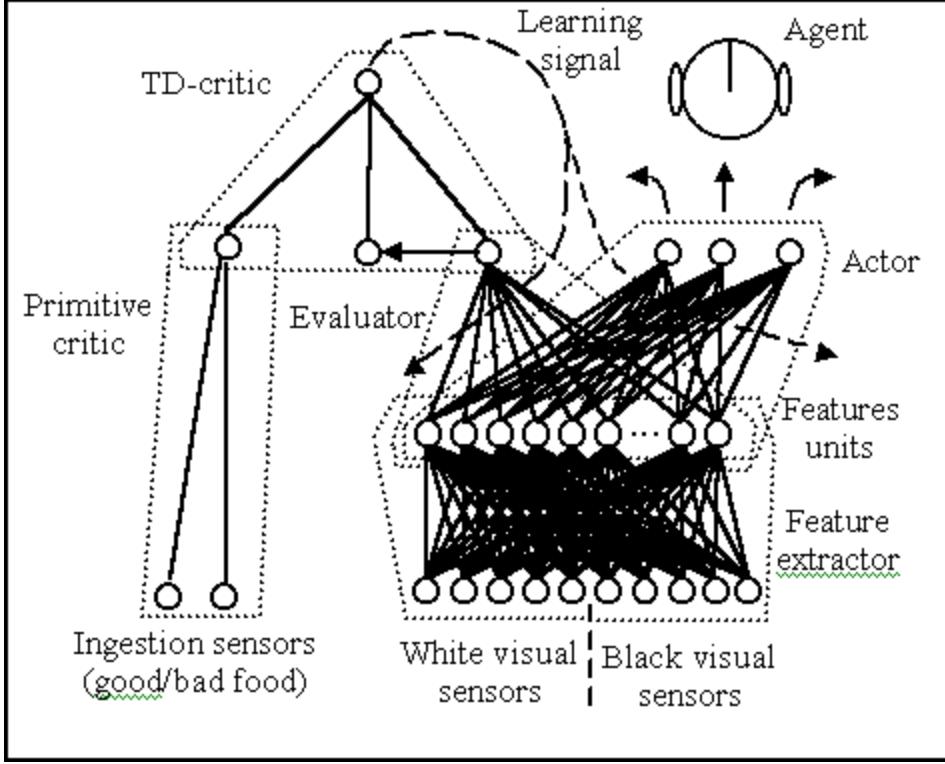
In the simulations, succeeding generations of agents (overlapping in time) live in the environment. Each agent is capable of learning to search for good food and to avoid bad food during its life (individual learning). Also it is capable of learning this capacity from its (only) parent (cultural transmission).



**Figure 2:** The environment (1x1) with the agents and food items (radius 0.005 and 0.0025 respectively). The "zoom window" shows the 5 visual fields of the 5 white and black pairs of sensors of one agent, and their activation.

## 2.2 Individual learning

The individual learning process is now described. Figure 3 shows the main components of the learning controller of one agent. The learning controller allows the agents to search for good food and to avoid bad food. The learning algorithms used to train the controller mimic a trial-and-error process and are based on the actor-critic model of Sutton and Barto (1998). See Houk et al. (1995) for a hypothesis of the primate brain's neural structures that correspond to the neural components of the model.



**Figure 3:** The neural architecture controlling an agent. The circles and arcs represent neurons and connections. The dotted arrows represent the learning signal that allows updating the weights of the evaluator and the actor.

The primitive critic incorporates the guiding criterion of reinforcement considered previously. This guiding criterion is innate. The primitive critic is made up of a simple neural network with two input units (one for the good-tasting food,  $g$ , and one for the bad-tasting food,  $b$ ) and one output unit,  $r$ . The input units assume the value of 1 when an item of the corresponding (good or bad) food is ingested and a value of 0 otherwise. The two (innate) connection weights  $w_g$  and  $w_b$  are set to the value +1 and -1 respectively, so that a reward or punishment ("pleasure" or "pain") is signaled by the linear output unit when an item of good or bad food is ingested. The activation of the output unit is computed as follows:

$$(1) r = w_g g + w_b b$$

The feature extractor takes the 10 signals from the visual sensors as input and implements a "Kanerva re-coding" of them (Sutton and Barto, 1998). The (innate) weights of the feature extractor are randomly drawn from the set {0, 1}. Each of the feature units (150 in the simulation) activates with 1 if the Hamming distance (standardized to 1 by dividing it by the number of feature units. The Hamming distance between two binary arrays is the number of elements with same position but different values) between the input pattern and the "prototype" encoded by its weights is less than a certain threshold (0.4 in the simulations). If the Hamming distance is greater than the threshold then the feature unit activates with 0. The main function of the feature extractor is to map the input space into a space with a higher dimensionality so to avoid eventual problems of non-linear separability and to attenuate interference problems during learning (Sutton and Barto, 1998).

The actor that is equivalent to the agent's action-selection policy (behavior), is a two-layer feed-forward neural network that takes the activation of the feature units as input and has three sigmoidal output units that locally encode the three actions. To select one action, the activation  $p_k$  (interpretable as "action merit") of the three output units is used in a stochastic winner-take-all competition. The probability  $P[.]$  that a given action  $a_g$  among the  $a_k$  actions becomes the winning action  $a_w$  is given by:

$$(2) P[a_g = a_w] = p_g / \sum_k p_k$$

The evaluator incorporates the second kind of guiding criteria considered in this work, the "evaluations". An evaluation is a measure of the potential of a state of the world to deliver future reward or punishment. The evaluator is a two-layer feed-forward neural network that gets the activation of the feature units as input. With its linear output unit it learns to express the estimation  $V^\pi[s_t]$  of the evaluation  $V^\pi[s_t]$  of the current state  $s_t$ .  $V^\pi[s_t]$  is defined as the expected discounted sum of all future reinforcements  $r$ , given the current action-selection policy  $\pi$  expressed by the actor:

$$(3) V^\pi[s_t] = E[\gamma^0 r_{t+1} + \gamma^1 r_{t+2} + \gamma^2 r_{t+3} + \dots]$$

where  $\gamma \in (0, 1)$  is the discount factor, set to 0.95 in the simulations, and  $E[.]$  is the mean operator.

The TD-critic is an implementation in neural terms (weights to be considered as innate) of the computation of the Temporal-Difference error  $e$  defined as (Sutton and Barto, 1998):

$$(4) e_t = (r_{t+1} + \gamma V^\pi[s_{t+1}]) - V^\pi[s_t]$$

The evaluator is trained with a Widrow-Hoff algorithm (Widrow and Hoff, 1960) that uses as error the error signal coming from the TD-critic. The weights  $w_i$  are updated so that the estimation  $V^\pi[s_t]$  expressed at time  $t$  by the evaluator, tends to be closer to the target value  $(r_{t+1} + \gamma V^\pi[s_{t+1}])$ . This target is a more precise evaluation of  $s_t$  because it is expressed at time  $t+1$  on the basis of the observed  $r_{t+1}$  and the new estimation  $V^\pi[s_{t+1}]$ . The formula of the updating rule is:

$$(5) \Delta w_i = \eta e y_i$$

where  $\eta$  is a learning rate (set to 0.01 in the simulation) and  $y_i$  is the activation of the feature unit  $i$ . This algorithm implements the individual learning of the evaluations. The idea behind this algorithm can be explained by assuming that the action-selection policy expressed by the actor is fixed. At the beginning of the simulation the weights of the evaluator are set randomly so it will associate evaluations close to 0 to each perceived state of the world (in fact its transfer function is linear). The first time that the agent bumps into a good (or bad) item of food, it perceives a

reward of +1 (or -1), so the TD-critic expresses an error of about +1 (or -1). Given this error, the evaluator learns to associate a higher (or lower) evaluation to the state of the world that preceded the one where the food was ingested. The next time that this state is perceived, it will cause the state preceding it to be evaluated with a higher (or lower) evaluation. This process will continue in a backward fashion so that the agent will finally assign positive (or negative) decreasing (because of the discount factor) evaluations to the sequence of states preceding the ingestion of good (or bad) food.

The actor is also trained according to the error signal coming from the TD-critic, so that the action-selection policy is improved with experience. At the beginning of the simulation the actors' weights are set randomly, so the output units have an activation close to 0.5 for each input pattern (because its transfer function is sigmoid), and the probability of selecting each of the three actions is close to 0.33. This induces the actor to explore the environment randomly. Given that the evaluator (once it has been trained) expresses an evaluation  $V^\pi[s_t]$  of  $s_t$  according to the average effect of the actions that the actor selects in association with  $s_t$ , an error  $e_t > 0$  means that the selected action  $a_w$  has led to a new state of the world  $s_{t+1}$  (evaluated  $r_{t+1} + \gamma V^\pi[s_{t+1}]$ ) that is better than the one previously experienced after  $s_t$ . In this case the probability of selecting  $a_w$  in association with  $s_t$  is increased. Similarly an  $e_t < 0$  means that the selected action  $a_w$  has led to a state with an evaluation smaller than the average, so its probability is decreased. Formally, the change of the probability is done by updating the weights of the neural unit correspondent to  $a_w$  (and only this) as follows:

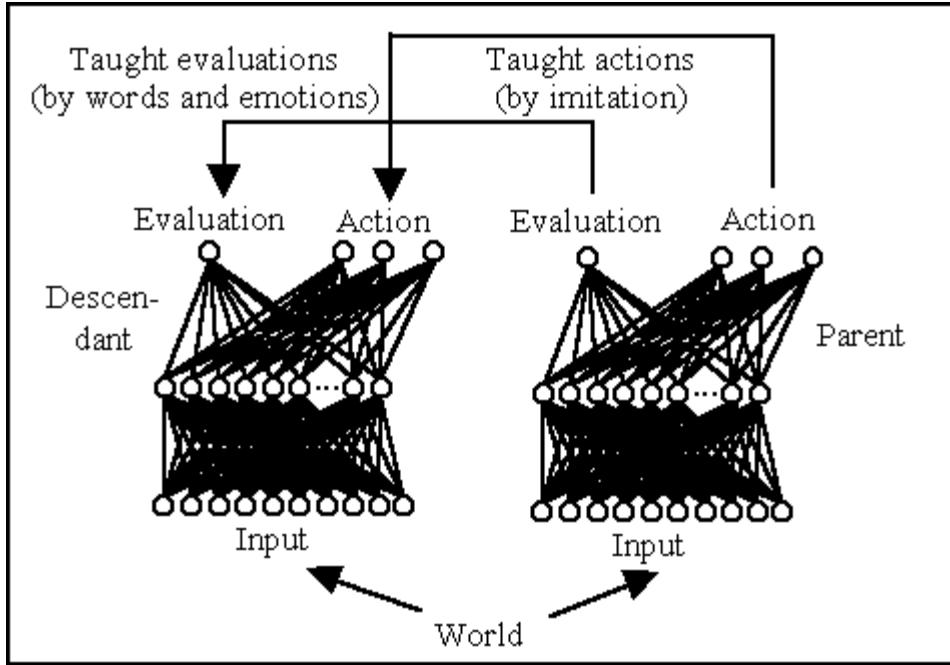
$$(6) \Delta_{w_{wi}} = \zeta e_t y_i$$

where  $\zeta$  is a learning rate, set to 0.01 in the simulation.

In the simulation the evaluator and the actor are trained simultaneously. The evaluator learns to evaluate the states of the world on the basis of the action-selection policy currently adopted by the actor, and the actor learns to improve the action-selection policy by increasing the probabilities of those actions that positively surprise the evaluator, i.e. that produce better results than the actions previously selected in the same conditions.

### 2.3 Cultural transmission

The cultural transmission process is now described. There are two kinds of cultural transmission. Within the first kind, that involves the transmission of behavior, the descendant learns to imitate the behavior of the parent. The descendant should be thought of as following the parent "on its shoulders", perceiving the same visual input, selecting an action (using its actor), observing the action of the parent, and trying to conform its own action to the parent's one. As in Denaro and Parisi (1996), this is implemented by training the descendant's actor network with a Widrow-Hoff algorithm (with learning rate of 0.01). The Widrow-Hoff algorithm uses as output the output signals  $p_k$  that the descendant associates to the current input pattern and as teaching input the output signals given by the parent. Figure 4 represents this process.

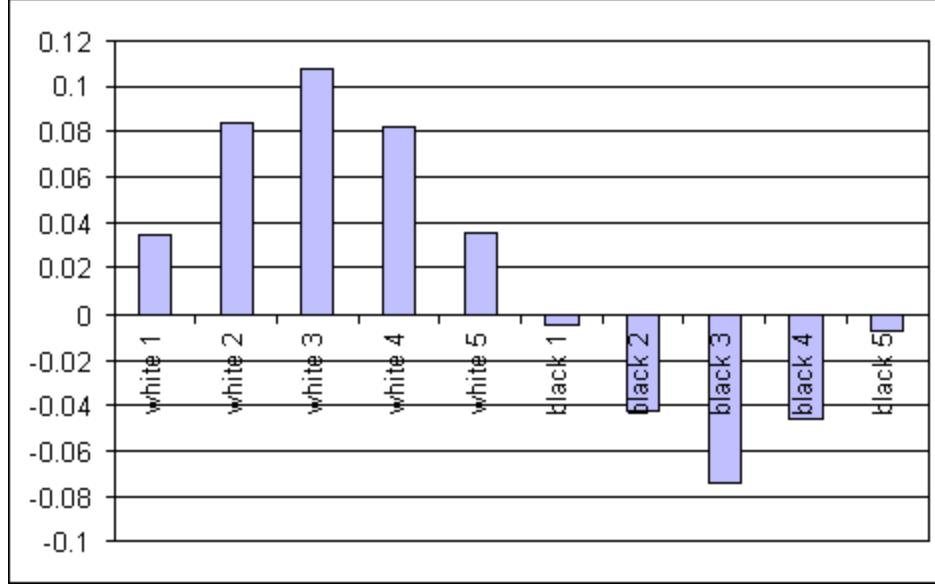


**Figure 4:** Transmission of evaluations through words and emotions, and transmission of behavior through imitation.

The second type of cultural transmission involves the transmission of "guiding criteria", in this case the evaluations associated by the parent to the states of the world (figure 4). A parent should be thought of as verbally (or emotionally) expressing the evaluation that it associates to the perceived situation and the descendant as learning by listening (or observing). This is implemented by training the descendant's evaluator network with a Widrow-Hoff algorithm (with learning rate 0.01) that uses as output the evaluation that the descendant assigns to the current perceived situation ( $V^\pi[s_t]$ ), and as teaching input the parent's evaluation.

### 3. Results and interpretation

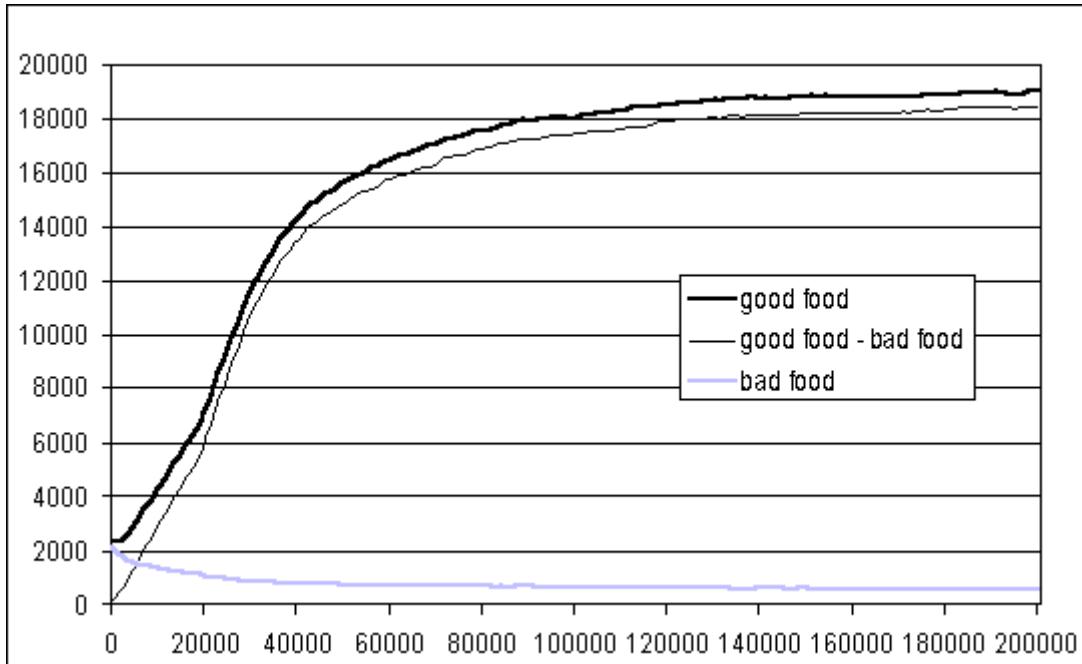
The first simulation is intended to clarify the meaning of "evaluation" used in this work. An agent wanders randomly in the arena (this is done by setting the learning rate of the actor to 0) for 20000 cycles. Its evaluator learns to assign evaluations to each perceived state of the world. A low learning rate of 0.001 has been used to produce stable evaluations. Figure 5 shows the evaluation (averaged across 20 agents) assigned to 10 possible states of the world, each corresponding to the activation of one (and only one) sensor, after 20000 cycles. It can be seen that the evaluation assigned to each state of the world as perceived reflects the probability with which the organism will bump into a good (or bad) item of food. For example, the perception of an item of good food in front (3rd white sensor active) receives the highest evaluation.



**Figure 5:** Average (over 20 agents) evaluation of 10 different world states, after 20000 cycles of experience.

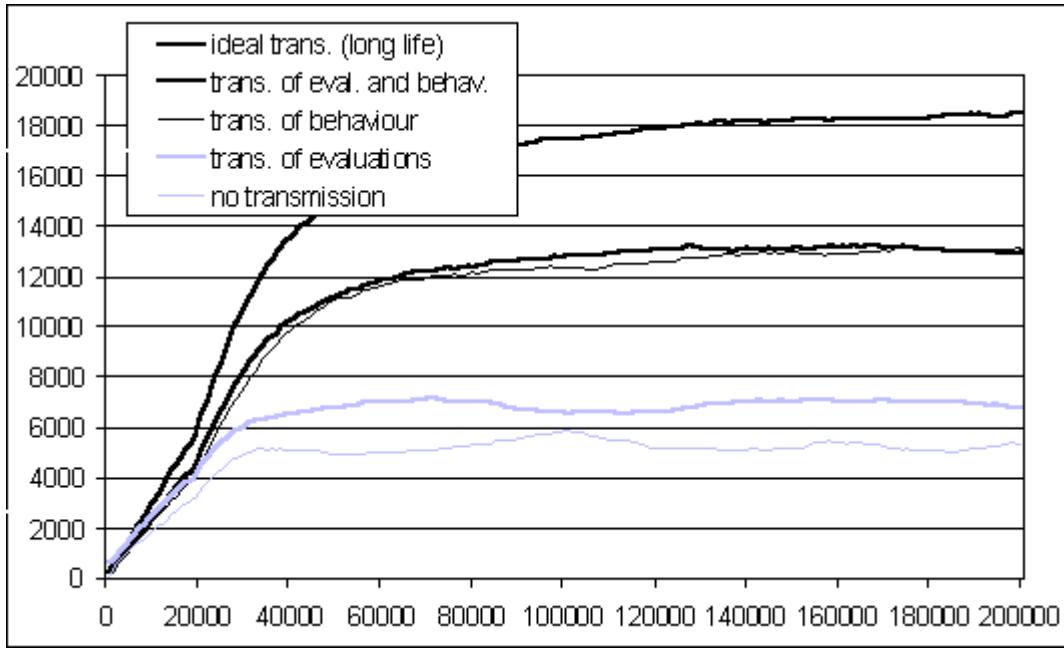
In order to test the effectiveness of the agents' individual learning process, a simulation where each agent had a very long life (200000 cycles) was run. Figure 6 shows the total number of items of good food, bad food and their difference, collected by the whole population of 50 agents against the number of cycles (the graph plots a moving average over 20000 cycles, and shows the average of three simulations run with different random seeds). This graph and direct observation of agent behavior shows a good capacity to search for good food and avoid bad food.

The following simulations have been run in order to evaluate the effectiveness and role of the different kinds of cultural transmission (introduced in section 2) in the population's cultural evolution. Each agent of the first generation has a life length randomly drawn from the interval [0, 20000]. 5000 cycles before its death, each agent generates one descendant to which it transmits some knowledge (vertical transmission, see Cavalli Sforza and Feldman, 1981). Each descendant has a life of 20000 plus a random number of cycles drawn from the interval [-1000, +1000]. Each descendant has the same neural structure shown previously and a feature extractor, evaluator and actor with initial random weights. The performance of the population is measured under four different conditions: transmission of both behavior and evaluations; transmission of behavior; transmission of evaluations; and no transmission.



**Figure 6:** Moving average (20000 cycles) of the items of good food, bad food, and their difference, collected by the whole population.

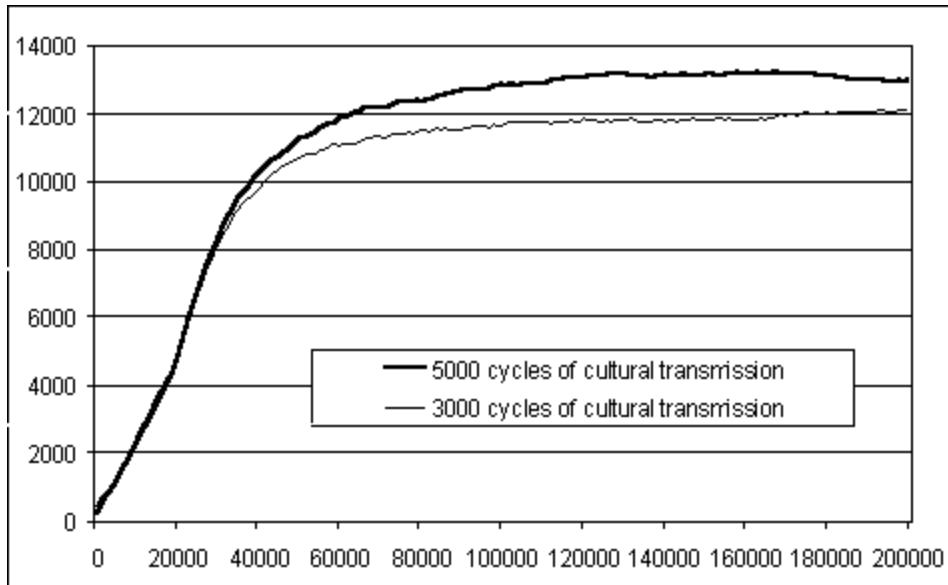
Figure 7 reports the results of the simulations run under these four conditions (for each condition three simulations with different random seeds were run and averaged). It also reports the plot of figure 6 relative to one organism with a long life. This case can be considered equivalent to a hypothetical situation ("ideal transmission of knowledge") where the cultural transmission from parent to descendant does not include any loss of knowledge due to errors of imitation or communication. In the simulations these errors are reproduced by the fact that the Widrow-Hoff algorithm never brings the error to 0.



**Figure 7:** Amount of good food minus bad food collected by the population in the last 20000 cycles.

Three relevant facts are identified through these simulations. The first is that all the three conditions with cultural transmission (transmission of both behavior and evaluations, transmission of behavior, and transmission of evaluations) yield a better performance of the whole population (about 13000, 13000, and 7000 respectively) than the condition where only individual learning is present (about 5000). This means that there is an accumulation of knowledge within the population across the generations: cultural evolution is taking place.

The second interesting fact is that even the most favorable cultural transmission condition (where both behavior and evaluations are transmitted), leads to a performance (13000) inferior to the ideal condition (where there is no loss of cultural knowledge, 18000). This means that the very process of transmission implies a continuous loss of knowledge that prevents the population from reaching the maximum performance. As mentioned, this is due to the fact that the Widrow-Hoff rule never reduces the error to 0. It is also due to the fact that the period of cultural transmission for each agent is limited in time. A further simulation has shown that the shorter the cultural training period for each agent, the lower the population performance level reached in the long term (see figure 8, averaged for three random seeds). This result is attenuated by the fact that in the simulations cultural transmission has (unrealistically) no costs.

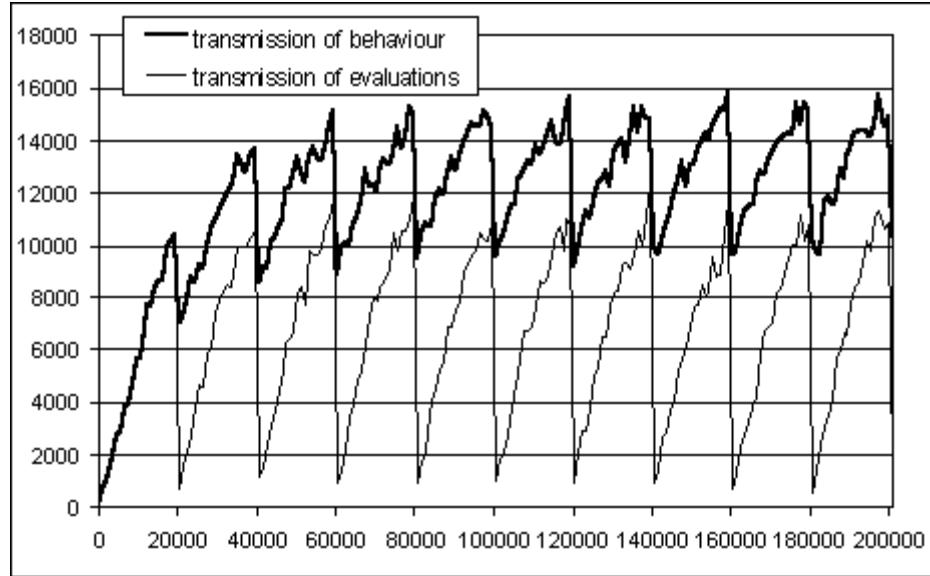


**Figure 8:** Performance of the population when the cultural transmission period for each agent is 5000 versus 3000 cycles.

In order to test the interpretation of the first and second facts, another simulation was run with the three distinct cultural-transmission conditions, but with a "synchronized" population: each agent has a life lasting precisely 20000 cycles, first generation included. To reveal changes at a finer time granularity, the performance of the population (determined by the difference between the number of good and bad food items collected), has been measured with a moving average of 1000 cycles (versus 20000 cycles used in the previous simulations; to keep the scale consistency the performance has been multiplied by 20). The results are shown in figure 9 (averaged for 3 random seeds). For clarity the plot relative to the condition with both transmissions of behavior and evaluations is not reported in the graph because it closely resembles the condition with the transmission of behavior only.

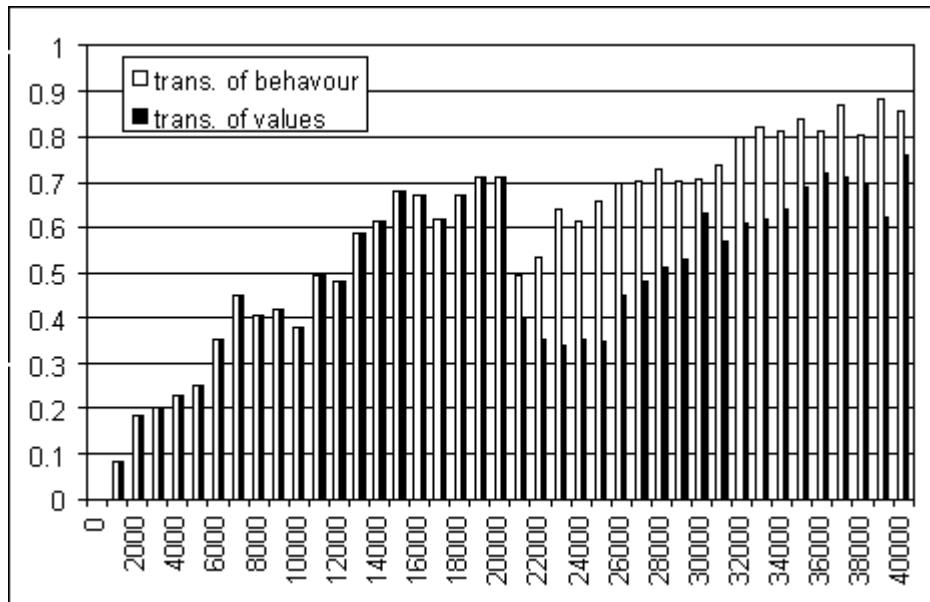
It can be noted that in both conditions (but in particular with the transmission of behavior) each new generation, with the exception of the first one, starts its life with skills above the "ignorance" level (i.e. 0). This explains the population's improving performance shown in figure 7, and shows the nature of the cultural evolution process: each newly born agent enters the adult life already possessing some skills. During its life it further improves these skills, and then passes the refined skills to its descendent. The result is an increase of the population average performance across generations.

It can also be seen from figure 9 that at each passage from one generation to another the performance decays abruptly. This is due to the errors of transmission of behavior and evaluations mentioned before.



**Figure 9:** Performance of the population when the generations are synchronized.

The third relevant fact that emerges from the simulation reported in figure 7, quite unexpected, is that cultural transmission of behavior is greatly superior to cultural transmission of evaluations (13000 against 7000).



**Figure 10:** Evaluation assigned to an item of good food perceived with the central sensor. Average for the population.

Figure 10 explains this finding by showing the dynamics of the evaluations of the first and second generations (20000+20000 cycles) of the preceding simulation. Every 1000 cycles the evaluation that each agent has assigned to the last item of good food seen with the white central sensor (all the other sensors being off) is measured and an average is computed for the population. This evaluation can be considered as an indicator of the overall capacity to evaluate.

In fact good food in a frontal position is highly promising of reward and deserves a high positive evaluation. So an agent that has learned to evaluate correctly the states of the world should express a high evaluation in this circumstance.

The most surprising fact is that the second generation's capacity to give correct evaluations at the beginning of life (cycle 21000), is higher with the transmission of behavior than with the transmission of evaluations (0.5 versus 0.4). The possible explanation of this fact is that once the descendants have culturally inherited the behavior that leads to search for good food and avoid bad food, they rapidly learn the evaluations by individual experience. The graph shows that after 7000 cycles (cycle 27000) the descendants have already learned to express evaluations similar to the parents' ones, and even improve on their parent's evaluations. At the level of population's performance this fact renders cultural transmission of behavior quite effective, as shown in the previous simulations.

Where the descendants inherit the evaluations, they are not capable of directly using them to search for good food and avoid bad food, because they do not possess the behavior necessary to do so. The immediate consequence is that the evaluations are themselves corrupted (from cycle 20000 to cycle 25000) because they prove to be incorrect: after all, having good food in front of you is not a very promising situation if you are not capable of reaching for it! The long-term consequence is that the descendants recover evaluations similar to the parents' ones only after 16000 cycles (cycle 36000), near the end of their life. At the level of population this causes a severe loss of skills passing from one generation to the other, and this makes the process of cultural transmission of evaluations quite ineffective, as shown in the previous simulations.

## 4. Related work

Several computational models of imitation and social transmission of skills and knowledge have been developed. Here those that use reinforcement learning, and that are closely related to the model presented in this paper, are reviewed.

Lin (1992) presents a model (tested within a food-gatherer predator grid world) where a reinforcement learning agent learns from its own recent past experience ("experience replay") by being (re-) exposed to histories of quadruples of the type "state-action-next state-reward" suitably stored in a buffer. This increases the speed of learning. This model is relevant for social learning because the "experience" could potentially derive from another agent. Tan (1993) presents a model (tested within scout-hunter/co-operative hunters grid worlds) where reinforcement learning agents communicate perceptions or use and update the same policy, so obtaining advantages in terms of speed of learning and performance. Clouse (1996) presents a model (tested using the Race Track game) where a reinforcement learning agent can execute an action "suggested" by an expert teacher instead of its own action. It is shown that higher rates of learning are obtained with high probabilities of suggestion per step. Finally Price and Boutilier (1999) present a model (tested with grid-maze tasks) of "implicit imitation". Here a reinforcement learning agent observes the effects produced by actions executed by a mentor, uses them to build a model of the world (state transitions induced by the mentor's actions), and then trains itself to act by using this model (model-based reinforcement learning).

All these models use a form of "greedy policy" to select actions. This implies that the action to be executed is directly selected as the one with the highest expected (cumulated discounted) future reward. No merits or probabilities are explicitly stored for the selection of actions. In contrast, this storage takes place in the actor-critic model adopted in this paper, and appears to happen in real organisms (Houk et al., 1995). In the research presented here it has been the adoption of a model which includes explicitly stored merits that has allowed separation of the modification of the evaluations from the modification of behavior, and the investigation of their relationship.

## 5. Conclusion and future work

An important form of cultural evolution involves the social transmission of behaviors and skills from one generation to the next, and a change and eventually improvement of these by means of individual learning and experience. The change of behavior and skills by individual learning requires some "guiding criteria" to be accomplished, which determine which new behaviors and skills will be selected and retained, and which will be discarded. In the long term, the guiding criteria strongly bias which behaviors and skills enter the pool of cultural traits of the population, and as such they play a central role in determining the cultural evolution of a population. This work is intended to contribute to the clarification of the nature of "guiding criteria". With this purpose it focused on some guiding criteria with a well-established biological origin: the reinforcement, that is mainly innate, and the evaluations, that are learned with experience or themselves culturally transmitted. After presenting a computational model incorporating these concepts, the different effects produced by the transmission of behavior and evaluations on one aspect of the adaptation of the population have been investigated. First, the simulations have helped to clarify the nature of the form of cultural evolution considered here, and have shown that the cultural transmission of evaluations and behavior produces positive effects on the performance of the population. Second, they have shown that cultural transmission is more effective if the loss of knowledge and skills due to errors of imitation and communication is attenuated by longer training of new generations. Third, and most importantly, they have shown that the transmission of behavior is more effective than the transmission of evaluations. It should be noted that this result holds within the assumptions of the model (where evaluations are changed quickly if shown to be inconsistent with direct experience). However, it could be the case that culturally learned evaluations persist in the face of contrary direct experience in real agents (this could eventually be modeled with different reinforcement learning algorithms). The final answer can only come from an empirical validation of the prediction of the model here presented.

This work has considered only a few forms of guiding criteria, namely reinforcement and the evaluations originating from reinforcement. These have a biological origin, and have been generated (or the mechanisms that generates them has been generated) by natural selection because they enhance organisms' fitness. Other forms of guiding criteria probably have an origin more strongly related to cultural processes and may eventually counter-act biologically generated criteria. For example, Miceli and Castelfranchi (1999) have investigated a notion of evaluation different from that considered here, related to the cognitive formulation of goals, and the notion of "values" (a special kind of evaluation where the goal is left unspecified). These other kinds of

guiding criteria, and their relation with the biological ones, should be the object of investigation in future work.

## Acknowledgements

I thank the Department of Computer Science, University of Essex, which funded my research. I also thank Prof. Jim Doran (University of Essex) and Domenico Parisi (Italian National Research Council) for their valuable support and James Adam for his precious help in the preparation of the article.

## References

- Baldassarre G., Parisi D. (1999), Trial-and-Error Learning, Noise and Selection in Cultural Evolution: A Study Through Artificial Life Simulations, *Proceedings of the AISB'99 Symposium on Imitation in Animals and Artifacts*, AISB - The Society for the Study of Artificial Intelligence and Simulation of Behaviour, Sussex - UK.
- Boyd R., Richerson P. J. (1985), *Culture and the Evolutionary Process*, University of Chicago Press, Chicago.
- Cavalli-Sforza L. L., Feldman M. W. (1981), *Cultural Transmission and Evolution: A quantitative Approach*, Princeton University Press, Princeton - NJ.
- Clouse A. J. (1996), Learning from an automated training agent, In Weiband G., Sen S. (eds.), *Adaptation and Learning in Multi-agent Systems*, Springer Verlag, Berlin.
- Denaro D., Parisi D. (1996), Cultural evolution in a population of neural networks, In Marinaro M., Tagliaferri R. (eds.), *Neural nets: Wirn96*, Springer, New York.
- Houk C. J., Adams L. J., Barto G. A. (1995), A model of how the basal ganglia generate and use neural signals that predict reinforcement, In Houk C. J., Davis L. J., Beiser G. D. (eds.), *Models of Information Processing in the Basal Ganglia*, The MIT Press, Cambridge - Mass.
- Lieberman A. D. (1993), *Learning - Behaviour and Cognition*, Brooks / Cole publishing, Pacific Grove - Ca.
- Lin L. J. (1992), Self-improving reactive agents based on reinforcement learning, planning and teaching, *Machine Learning*, Vol. 8, Pp. 293-321.
- Miceli M., Castelfranchi C. (1999), The Role of Evaluation in Cognition and Social Interaction, In Dautenhahn K. (ed.), *Human Cognition and Social Agent Technology*, John Benjamins, Amsterdam.
- Price B., Boutilier C. (1999), Implicit imitation in multi-agent reinforcement learning, *ICML'99 - International Conference on Machine Learning*, Pp. 325-334.

Rolls E. (1999), *Brain and Emotion*, Oxford University Press, Oxford.

Sutton R. S., Barto A. G. (1998), *Reinforcement Learning: An Introduction*, The MIT Press, Cambridge - Mass.

Tan M. (1993), Multi-agent reinforcement learning: Independent vs. cooperative agents, *ICML'93 - International Conference on Machine Learning*, Pp. 330-337.

Widrow B., Hoff M. E. (1960), Adaptive switching circuits, *IRE WESCON Convention Record*, Part IV, Pp 96-104.

## **MEMESIS CRITIQUE**

By  
Richard Barbrook

### **ASSERTION 1**

The human being, characterized by a remarkable ability to process information...

In the first sub-clause of the opening sentence, the human ability to communicate complex thoughts to each other is elevated above all other aspects of our existence.

Although expressed in computer jargon, this separation of thinking from doing is one of the oldest inversions found in philosophy, as in Plato's parable of the cave. For centuries, it was used the conjuring trick of the priestly caste - elevating the speaking of sacred words above physical toil on the earth of the peasants. After over two centuries of modernity, it is surprising to find this ancient idea at the beginning of a declaration supposedly describing the future. But, as we will see, the process of seeing the world upside down is one of the most important failings of the whole Memesis statement.

### **ASSERTION 2**

Let's start again:

The human being, characterized by a remarkable ability to process information, has extended his phenotype further than any other species. Complex tools and technologies are an integral part of our evolutionary "fitness". Human evolution is fundamentally intertwined with technological development; the two cannot be considered apart from one another. Humanity has co-evolved with its artifacts; genes that are not able to cope with this reality will not survive the next millennium.

The other major error in the Memesis statement is its use of dodgy biological analogies. The discovery of evolution was one of the key intellectual moments in the development of modern society. By offering a rational understanding of the origins of humanity in nature, it destroyed the intellectual basis of revealed religion. Crucially, evolutionary theory was conceived in the first truly modern society: Great Britain. The rapid economic and social changes taking place in this first industrial society enabled Darwin to understand that nature itself was also in flux.

However, problems arise when the relationship is drawn in the other direction: when natural evolution is used to explain social development. In this century, millions of people were shoved into gas chambers because it was believed that they possessed 'genes that are not able to cope' as the Memesis statement puts it. Following the defeat of fascism, the biological metaphor is now more often used to revive an earlier illegitimate use of Darwinian Theory for political purposes: Social Darwinism. As championed by Herbert Spencer, this theory claimed that unregulated market competition between private property owners was a natural phenomenon rather than a social one. The moral claims of liberalism were restated in positivist language. For instance,

Spencer even opposed the installation of municipal sewerage systems to prevent cholera and other diseases as an obstruction of the natural laws of the market!

Although discredited in the early part of this century, the globalization of capitalism over the past few decades has been accompanied by a renewed faith in the simplicities of liberal economics. In turn, the idiocies of Spencer have been revived in an updated form - as can be seen in Kevin Kelly's book 'Out of Control'. By adopting the rhetoric of biology, the Memesis statement is not simply trying to reduce the complexities of millennia of social development to the self-replication of DNA proteins. It is also implicitly supporting the failed social and economic policies of the 1980s.

### **ASSERTION 3**

As an analogy to the building blocks of biology, the genes, memes describe cultural units of information, cognitive behavioral patterns that propagate and replicate themselves through communication. From the "bio-adapter" of language as a proto-meme to the "infosphere" of global networks as the ultimate habitat for the human mind.

If you combine two errors, you go further into confusion. Having separated thinking from doing and reduced society to biology, the Memesis statement now goes on to assert that human consciousness is an autonomous gene! Dressed up in newly fashionable biospeak, what we have here is a garbled fusion of:

- The old Stalinist lie that history is a process without a subject;
- The old post-structuralist lie that language speaks us;
- And the old McLuhanite lie that technology shapes our minds.

For if memes 'replicate themselves', what are humans doing in the meantime? We're not the blind objects of genes or memes. Rather, whether using language or global networks, humans transform themselves and nature through our own activity. We are the subject of history - even if it is not always in circumstances of our own choosing.

### **ASSERTION 4**

The discussion is intended to probe specific segments of the techno-cultural revolution against the background of the idea of a "culturally based history of creation".

So now we get to the core of the Memesis statement. The current process of the convergence of telecommunications, the media and computing is to be explained through the dubious biological metaphor. The prosaic task of developing the hardware and software for an integrated communications and information network is to be transformed into a positivist mysticism: the fulfillment of our pre-programmed genetic destiny. Henri Lefebvre remarked that structuralism, semiotics and psychoanalysis were usually used by intellectuals as ways of avoiding examining the social relations of capitalism. Now we can add biology to the list.

## **ASSERTION 5**

...and also science fiction!

This is not to develop new utopias, but rather to critically assess the current scenario, which promises the fulfillment of long prophesied visions of the future. The possibility of the emergence of a post-biological, cyberorganic line of evolution out of universal binary code systems, of which the first protozoans have names like Internet, Cyberspace and I-way.

The Memesis statement is certainly not advocating a 'new utopia'. It is instead repeating one of the oldest plots in science fiction: the Frankenstein monster, Hal in 2001, the Terminator and so on. But, whereas in the past we were supposed to fear the destructive powers of technology, we're now invited to celebrate 'the emergence of a post-biological, cyberorganic line of evolution'. The ending might be more optimistic, but we've read the book and seen the film before.

And where is this new life form supposed to come from? Allegedly it emerges from the 'Internet, Cyberspace and I-way'. Yet the Net is a creation of human labor. Someone has to dig holes in the road to lay the fiber-optic wires. Someone has to work on the production line building the PCs. Someone has to design the hardware which will control the flow of data across the networks. Someone has to write the software to enable people to use the Net. And all of us participate in communicating across cyberspace with one other. Without human activity, the Net is nothing but an inert mass of metal, plastic and sand. We are the only living beings in cyberspace.

## **ASSERTION 6**

As the biological body coincides with its mechanical and now informational clone as well, neurobionic, robotic prosthetics question our relationship to the body and to gender; cyborg theory and cyberbody fetish as response.

If the Net can be explained by Social Darwinism, why not recent advances in medical science also? In an orgy of hyperbole, people with pacemakers - or even glasses - can become characters from science fiction tales: the cyborgs. Deprived of the consolation of religion, we can now rediscover the promise of eternal life in a new hi-tech form.

But by embracing this positivist mysticism, we ignore the creative powers of our own species to transform its own conditions of existence. Above all, it is not technology by itself which has changed our perception of gender relations. Rather the possibilities of modernity have been grasped by women – and some men - to transform the limitations on human existence imposed by patriarchal societies and biological necessity. It might seem deliciously wicked to indulge in 'cyborg fetish', but worshiping the machine and the gene obscures past and present struggles of flesh 'n' blood humans to civilize social relationships between the genders.

## **ASSERTION 7**

Media memory - the collective memory and experience of humanity externalized in world-wide networks. Memes, as a "mass crystal", the identification and integration of virtual communities that gather only in network interfaces. After repeatedly claiming that memes are self-replicating, the Memesis statement suddenly announces that 'media memory [is] the collective memory and experience of humanity externalized in world-wide networks'. So the memes are actually our own creation. But, if there is active human involvement in the construction of cyberspace, how can technology and genes be autonomous forces outside our control? The fetishistic conception of technology and information within the Memesis statement is now openly revealed for the first time. What we're dealing with here is a moment of Feurbachian revelation: the Memesis statement now admits that it wants us to worship the magical power of graven images built with our own labor.

The central error of the Memesis statement has become obvious. It regards machines and information as autonomous things outside our control. Yet, in reality, both technology and culture are expressions of the social relationships between individual humans. It is human activity which is crystallized into machines and information, not memes which create 'mass crystal'. Crucially, by denying the Promethean power of collective creativity, the Memesis statement ignores one of the central questions of modernity: how are the rewards of labor to be divided among the different groups involved in the social production of machines and information? Ah, but the social question is so unfashionable nowadays...

## **ASSERTION 8**

Memes, the cognitive pixels as a blueprint for the cultural practice of sampling, of the universal "copy and paste", which has emerged from the new conditions of media.

Having briefly admitted that humans are the subject and culture is the object of their labor, the Memesis statement quickly reverts to its favorite philosophical inversion dressed up in computer jargon. Once again, it is not clever artists, musicians or designers who use digital technologies to create innovative new forms of cultural expression. On the contrary, the software mystically writes itself!

Moreover, the practice of sampling the work of others is hardly new. For centuries, artists have plagiarized their predecessors and contemporaries. All collective endeavors involves a constant process of sampling and 'copy and paste'. What the new information technologies have done is make the process much easier and more aesthetically pleasing.

Above all, they have enhanced and deepened the possibilities of strange juxtapositions and hallucinatory combinations. The Futurists could only dream of machine music. Jungle musicians can now actually create the digital rhythms of drum 'n' bass.

## **ASSERTION 9**

Memesis: a synonym for the current process of compression, for the convergence of various developmental vectors, which achieve a breakthrough as a whole.

Exhausted, we at last reach the end of this journey through intellectual confusion. Assuring that this was all just a 'synonym' for the process of convergence, the statement concludes with the promise of imminent rapture: 'a breakthrough as a whole'. But, at best, the Memesis statement is a piece of bad poetry. At worst, it is an apology for the defunct neo-liberalism and tired post-modernism of the last decade. The Memesis statement presents a 'Californian Ideology' for Europeans: radical rhetoric hiding conservative ideas.

The real crime of the Memesis statement is the way that it willfully obscures the process of human innovation and creativity under a mass of dodgy biological metaphors. In contrast, we must celebrate the Promethean power of humans to create - and recreate -themselves. It is precisely our refusal to accept our biological destiny which makes us more than insects. Unlike our fellow species, we can transform ourselves through thought and action.

The advent of modernity has radically accelerated this process of human self- transformation. The 'convergence of various developmental vectors' - as the Memesis statement so inelegantly puts it - is only the latest stage in this process of modernization.

What is interesting about our present situation is how imaginative people are using new technologies to push forward the limits of social and cultural creativity. Instead of being mesmerized by memes, what we should be doing instead is celebrating the achievement of these digital artisans - the artist- engineers who're pioneering the ways in which everyone will be able to participate within cyberspace in the future.

# **MEMETIC MEANINGS**

## **A COMMENTARY ON ROSE'S PAPER:**

### **CONTROVERSIES IN MEME THEORY**

**By**

**Francis A. Beer**

Political Science Department,  
Box 333, University of Colorado,  
Boulder CO 80309.  
[beer@spot.colorado.edu](mailto:beer@spot.colorado.edu)  
Tel. +1 303-492-7802

Rose (1998) raises a number of important issues for the developing area of memetics. These include the ambiguous definition of central terms like 'meme', memetics' use of central analogies and metaphors from genetics, and the role of philosophical concepts like 'self' and 'mind' in memetics. All these can be considered as issues of meaning.

## **1. Definitions and Meanings**

"The definition of a meme," Rose begins, "is currently ambiguous.... Without some kind of firm definition the word 'meme' becomes almost meaningless."

The set of problems centering on definition and meaning is serious but hardly limited to memetics. The empirical study of semantics tells us that various modifiers in different contexts subtly transform the definitions and meanings of most words. As our language becomes more distant from direct physical experience of the physical world, it builds upon itself to create and use abstract, general concepts. The core meanings, the constants diminish; contextual parameters gain heavier weight.

This semantic indeterminacy is true for words such as 'reason', 'validity', and democracy for example (Beer, 1997; Beer, 1994 ; Speel, 1997). It is not surprising, and should not be that the word 'meme' is similarly multivalent. The general nature of semantics leads us to expect no less.

## **2. Analogies and Metaphors**

Rose provides a number of alternative definitions of 'meme' together with current sources. Yet memetic genealogy is much longer than even his capable review suggests (Best, 1998). The field of rhetoric, for example, has long included the classical term 'enthymeme' which dates from classical times. In contemporary usage, the biological term 'gene' is a central current orienting analogue or metaphor for 'meme'. Genetics thus provides a primary, foundational orientation for memetics.

The field of cognitive linguistics suggests that there are two ways in which this genetic metaphor may work. We may call one of these, the strong program. In this program, memes are considered as nominal categories. Further, there is a direct mapping of attributes from the genetic source to the memetic target. (Lakoff, 1994). In this situation, we should expect all genetic characteristics

to be faithfully mapped into memetic space. All genetic attributes, however, do not faithfully reproduce in memetic space. Much of Rose's article, as well as other work, focuses on such general mapping anomalies from genetics to memetics, as well as differences in memetic applications in different fields (Beer, 1997). The strong program carries within it the seeds of its own disappointments. Memes may be like genes in certain ways, but they cannot be identical if only because of the differences between physical and cultural phenomena.

We may call another model of metaphorical process the weak program. In this program, memes are considered as interval categories, values for which may be derived through statistical regression or factor-analytic techniques. It implies a more subtle, complex, and non-linear dynamic of metaphorical meaning construction. This model is more forgiving since it allows genes to provide a general orientation for the meaning of memes, but does not place memes into a conceptual straight jacket. Instead of fitting memes to genes, it blends them. Further, the meanings of memes and memetics might take part of their inspiration from genetics, but there would be other sources and associations. The fit would be looser. Memetic meanings would be freer to adapt and evolve independently from their origins, depending on their uses and contexts. (LSA; Rohrer, 1997; Turner and Fauconnier, 1995).

### 3. Selves

This brings us to thorny philosophical problems of the self and others, free will and determinism. These terms have developed specific meanings within the context of genetic theory. For example, while genetics has accepted the identity of the organism, it has also partly relocated selfhood to the genetic level, the level of the 'selfish gene'. This genetic self is similar to the individual organism self in many respects, but it differs in several important ways. First, the genetic self is prior to and determinative of the individual self. Second, the genetic self acts randomly and blindly - not consciously and intentionally with goals and foresight.

Memetics moves forward from this position by posing new questions about the self. It asks us to consider the possibility that individuals or cultures may develop selves with both genetic and memetic cognitive capabilities for foresight and prediction, intentionality and goal specification, planning and structured action. This appears to involve a reorientation of the basic genetic model of the self, yet there need be nothing particularly shocking about this evolutionary memetic adaptation. In the first place, as we suggested earlier, memetics need not directly mimic genetics. Indeed, memetics may suggest new directions for genetic theory. Second, such conscious capabilities may be easily integrated into the basic evolutionary framework. Genetic foundations exist for such attributes. Evolutionary psychology, moreover, suggests that normal mechanisms of cellular growth, memory and learning may explain memetic learning.

Are memetic capabilities helpful to survival? Evidence suggests that they probably are. The Darwinian theory of evolution was itself a memetic adaptation. Though it is not often noted, Darwinian theory suggested that human beings were not blind, that the evolutionary theorist, at least, could perceive with some accuracy the structures and processes of nature. Darwinism contains within itself both memetic and genetic models for the self - memes for scientists, genes for nature. Darwinism implied that scientists, armed with proper knowledge, could predict future

natural events. Scientists could, further, plan and act intentionally to collect and interpret natural data.

## 4. Minds

Memetics advances from these origins along multiple parallel fronts. One of these is in the area of "mind." As we have suggested, it seems reasonable that minds include cognitive tools for representation, foresight, choice, and action and that these are genetically, biologically based, culturally and socially developed and programmed. These capabilities interact with other biological systems through complex feedback mechanisms and influence organisms' survival capabilities in different environments.

At the same time, we must recognize that the meaning of 'mind', like 'meme', is multivalent and metaphorical. 'Mind' has different meanings in biological, psychological, linguistic, and philosophical contexts.

To pose the problem as free will versus determinism may obfuscate the complexity of the situation. Free will or determinism may simply exist as different interpretations of the same phenomena in different theoretical contexts. From the naïve perspective of everyday common sense, free will begins at the body's edge. Philosophically and legally, free will is the essential component of our individuality, of our dignity, of our rights and responsibilities. From a biological or medical perspective, free will may represent the effects of unknown variables within deterministic systems; indeterminable interactive effects of multiple known variables within complex systems; or random error of otherwise deterministic processes.

As memetics develops, it will generate new meanings both for itself and for associated concepts and terms. Old analogies and metaphors, for example genetics and genes, may diminish in importance. New ones, perhaps centering on patterns of information, will gain. In any case, memetics will help reshape how we think about ourselves and our minds. For the time being, free will and determinism uneasily coexist. The emerging discipline of memetics will help us track their meanings as they evolve.

## References

Beer, F.A. (1993). Validates: a political science perspective. *Social Epistemology*, 7:85-105.  
(<http://socsci.colorado.edu/~beerball/valid.html>)

Beer, F.A. (1994). Words of reason. *Political Communication*, 11:185-201.  
(<http://socsci.colorado.edu/~beerball/papers/reason.html>)

Beer, F.A. (1997). A Review of: the International Studies Quarterly special issue on Evolutionary Paradigms in the Social Sciences, 40(3) in September, 1996. *Journal of Memetics-Evolutionary Models of Information Transmission*, 1. ([http://www.fmb.mmu.ac.uk/jom-emit/1997/vol1/beer\\_fa.html](http://www.fmb.mmu.ac.uk/jom-emit/1997/vol1/beer_fa.html)).

Best, M. L. (1998). A letter on: Memes on memes - a critique of memetic models. *Journal of Memetics - Evolutionary Models of Information Transmission*, 2. ([http://jom-emit.cfpm.org/1997/vol2/best\\_ml.html](http://jom-emit.cfpm.org/1997/vol2/best_ml.html)).

Lakoff, G. Conceptual Metaphor Home Page. (<http://cogsci.berkeley.edu/>).  
LSA. Latent Semantic Analysis Group Papers. (<http://lsa.colorado.edu/>).

Rohrer, T. (1997). Conceptual blending on the information highway: how metaphorical inferences work. *International Cognitive Linguistics Conference '95 Proceedings*, v.2. Amsterdam: Johns Benjamin. (<http://metaphor.uoregon.edu/iclacnf4.htm>)

Rose, N. (1998). Controversies in meme theory. *Journal of Memetics - Evolutionary Models of Information Transmission*, 2. ([http://jom-emit.cfpm.org/1998/vol2/rose\\_n.html](http://jom-emit.cfpm.org/1998/vol2/rose_n.html)).

Speel, H.-C. (1997). A memetic analysis of policy making. *Journal of Memetics - Evolutionary Models of Information Transmission*, 1. ([http://jom-emit.cfpm.org/1997/vol1/speel\\_h-c.html](http://jom-emit.cfpm.org/1997/vol1/speel_h-c.html)).

Turner, M. and G. Fauconnier. (forthcoming). Conceptual integration and formal expression. *Journal of Metaphor and Symbolic Activity*, 10, 3. (<http://metaphor.uoregon.edu/turner.htm>).

# MODELS FOR INTERACTING POPULATIONS OF MEMES: COMPETITION AND NICHE BEHAVIOR

By

**Michael L. Best**

Media Laboratory,

Massachusetts Institute of Technology,

Cambridge, Massachusetts 02139, USA.

mikeb@media.mit.edu

The original version of this paper was published as Best, M.L. (1997). Models for Interacting Populations of Memes: Competition and Niche Behavior. In Husbands, P. and Harvey, I. (eds.) (1997). *Fourth European Conference on Artificial Life*, Cambridge, MA: MIT Press. MIT Press has kindly given us permission to re-publish this new version.

## Abstract

We make use of a set of text analysis tools, primarily based on Latent Semantic Indexing, to study the dynamics of memes on the Net. Our analysis discovers replicating memes within posts to the USENET News (or NetNews) system. We cluster the posts to NetNews into clouds within a conceptual sequence space; these clusters describe quasi-species. We then go on to study the pairwise interactions between these quasi-species by computing the cross-correlation between the interacting population's level of post activity. We analyze a particular corpus of posts to the soc.women newsgroup and argue that strong negative cross-correlations are examples of competition between the quasi-species. We find that high levels of competition occur more frequently among quasi-species who exist within a narrow ecological niche. We note that this phenomena also occurs in natural ecologies.

## 1 Introduction

Ideas do not exist in a vacuum. Neither does discourse, the interconnected ideas which make up conversation and texts. In this research we investigate the pairwise interaction between populations of ideas within discourse: Are our text populations in competition with each other? Do they mutually benefit each other? Do they prey on one another?

This work attempts to build models of *population memetics* by bringing together two disciplines: Alife and text analysis. Through techniques of text analysis we determine the salient co-occurring word sets, texts, and text clusters, and track their temporal dynamics. We then study the life-like properties of this human-made system by considering its behavior in terms of replicators, organisms, and species.

Richard Dawkins coined the term *meme* to describe replicating conceptual units (Dawkins [7]). In studying the population dynamics of ideas we consider the meme to be the largest reliably replicating unit within our text corpus (Pocklington & Best [27], Pocklington [26]). Through text analysis we identify memes within a corpus and cluster together those texts which make use of a common set of memes. These clusters describe species-like relationships among the texts.

The particular texts we study are posts to the popular USENET News (or NetNews) system. These posts form the basis of a new Alife environment, the *corporal ecology* (Best [4, 5]). In this ecology texts are the organisms, the digital system defined by NetNews describes an environment, and human authors operating within some culturally defined parameters are the scarce resource.

At the core of our study sits a large text analysis software system based primarily on Latent Semantic Indexing (LSI) (Furnas, et.al [16], Deerwester, et.al. [9]; Dumais [10, 11]). This system reads each post and computes the frequency with which each word appears. These word counts are then used in computing a vector representation for each text. A principal component analysis is performed on this collection of vectors to discover re-occurring word sets; these are our memes. Each post is then re-represented in terms of these memes. By grouping texts which are close to one another within this meme-space we cluster semantically similar texts into species-like categories or *quasi-species* (Eigen et.al. [12]).

We proceed to study the interactions between those populations which coincide temporally. For each cluster we compute a series which represents its volume of post activity over time, for instance how many texts of a given cluster were posted on a given day. Cross-correlations between each pair of time series are then determined. We find that some pairs have strong negative correlations and argue that these are examples of texts in competition. A number of examples of such competition are explored in depth. We argue that high competition is correlated with those text clusters which exist within a narrow ecological niche; this phenomena is also observed in natural ecologies (Pianka [24]).

Note that this is an unusual shift from the typical Alife environment. We are not synthesizing replicators, embodying them into agents, and observing their life-like interactions. Instead we are studying a pre-existing artifact. Through our analysis we *discover* replicators within organisms, and use computational techniques to observe their dynamics.

In this paper we first briefly overview the NetNews environment and describe the LSI based text analysis system. Next we describe the mechanism used to determine the temporal dynamics and cross-correlations given a corpus of posts. We then relate the cross-correlations to models of interacting populations. In the next section we examine in depth a couple pairs of post clusters with strong interactions. We then describe a theory of niches within the corporal ecology and note that narrow ecological niches are correlated with significant competition. We end with our conclusions.

## 2 The NetNews Corpus

Understanding our corpus requires a basic knowledge of the NetNews system. NetNews is an electronic discussion system developed for and supported on the Internet (Kantor & Lapsley [18]). Discussion groups have formed along subjects ranging from science to politics to literature to various hobbies. The collections of messages are organized into particular subject groups called *newsgroups*. The newsgroups themselves are organized in a tree-like hierarchy which has general top-level categories at the root and moves to more specific topics as you progress

towards the leaves. A newsgroup name is defined as the entire path from the top-level category through any subsequent refining categories down to the name of the group itself. Category and group names are delimited by the period symbol. Thus, ``soc.religion'' is the name of a newsgroup concerned with social issues around the world's religions and ``soc.religion.hindu'' is a more specific group devoted to Hinduism.

Texts sent to NetNews, the *posts*, are composed of a number of fields only a few of which are relevant here. The user creating the post is responsible for the post body (that is, the actual text of the message) as well as a subject line. The subject line is composed of a few words which describe what the post is about. NetNews software will attach a number of additional fields to posted messages including a timestamp and the user name of the person who created the post.

Posts can be either an independent message or a follow-up to a previous message. A follow-up, or ``in-reply-to'' message, will have special threading information in its header linking it to the previous posts to which it is a reply. This header information allows news readers to reconstruct the discussion thread.

NetNews today has grown considerably from its beginnings in the late 70's and 80's. With over 80,000 posts arriving each day, it provides an excellent dataset for the study of cultural microevolution.

```
From: mikeb@media.mit.edu
Newsgroups: soc.religion.hindu
Subject: Angkor Wat
Date: 26 Jan 1997 02:17:05 -0700
Is Angkor Wat, the magnificent temple complex in the jungles
of Cambodia, considered a Buddhist or Hindu shrine?
```

**Figure 1.** A fictitious example post sent to the soc.religion.hindu newsgroup along with some of its header information.

### 3 The Text Analysis Method

We analyze a corpus of posts to NetNews to distill their salient replicating unit or memes, and to cluster together posts which make common use of those memes. We do this by employing a large system of text analysis software we have built. The techniques employed are based on the vector space model of text retrieval and Latent Semantic Indexing (LSI).

#### 3.1 Vector Space Representation

We begin with a corpus composed of the full-text of a group of posts. We analyze the corpus and identify a high-dimensioned space which describes the conceptual elements within the texts. For each post we identify a point within this space which captures it semantically. This technique is known as a vector space representation (Salton & Buckley [28]; Frakes & Baeza-Yates [15]). Each dimension in this space will represent a *term* from the corpus where a term is a word that occurs with some frequency (e.g. in at least three posts) but not too frequently (e.g. the word

``not'' is dropped from the term list). The goal is to arrive at a set of terms which semantically capture the texts within the corpus.

Given the conceptual space described by this set of terms each post can be represented as a point within this space. We score each document according to the frequency each term occurs within its text, and assign each term/document pairing this *term weight*. The weighting we use for each term/document pair is a function of the *term frequency* (simply the number of times the term occurs in the post) and the *inverse document frequency* (IDF). Consider a corpus of  $m$  posts and a particular term,  $j$ , within a list of  $n$  terms. Then the IDF is given by,

$$IDF_j = \log(m - m_j) / m_j$$

where  $m_j$  is the number of posts across the entire corpus in which term  $j$  appears. Thus, if a term occurs in 50% or more of the texts the IDF for that term will vanish to zero. But if, for instance, a term occurs in 10% of the documents the IDF will be  $\log(10)$ . In words, rare terms have a large IDF.

The term weight for a document,  $i$ , and term,  $j$ , is then defined by,

$$TermWeight_{ij} = w_{ij} = \log(TermFrequency_{ij}) \cdot IDF_j.$$

Each term weight, then, is a function of the inter- and intra-document term frequencies. Each post,  $i$ , is now represented by a particular term vector,  $r_i = (w_{i1}, w_{i2}, \dots, w_{in})$ . The entire collection of  $m$  term vectors, one for each post, define the *term/document matrix*,  $A$ .

This set of steps, culminating in the term/document matrix, form the basis for much of modern text retrieval or filtering and are at the core of most Web search engines.

### 3.2 Latent Semantic Indexing

LSI is a technique used to distill high-order structures from a term/document matrix, consisting of sets of terms which re-occur together through the corpus with appreciable frequency. The re-occurring term sets are discovered through a principal component method called Singular Value Decomposition (SVD). While LSI was primarily developed to improve text retrieval, we are interested in its ability to find replicating term sets which act as memes. We will first overview the LSI technique and then discuss how it discovers memes.

LSI was originally proposed and has been extensively studied by Susan Dumais of Bell Communications Research and her colleagues (Furnas, et.al. [16]; Deerwester, et.al. [9]; Dumais [10, 11]). Peter Foltz investigated the use of LSI in clustering NetNews articles for information filtering (Foltz [14]). Michael Berry and co-authors researched a variety of numerical approaches to efficiently perform SVD on large sparse matrices such as those found in text retrieval (Berry [1]; Berry, et.al. [2]; Berry & Fierro [3]).

The SVD technique decomposes the term/document matrix into a left and right orthonormal matrix of eigenvectors and a diagonal matrix of eigenvalues. The decomposition is formalized as,  $A_k = U V^T$ .

The term/document matrix,  $A$ , is approximated by a rank- $k$  decomposition,  $A_k$ ; in fact the SVD technique is known to produce the *best* rank- $k$  approximation to a low-rank matrix (Berry [1]).

We are interested in only the right orthonormal matrix of eigenvectors,  $V^T$ . Each row of this matrix defines a set of terms whose co-occurrences have some statistically salient re-occurrences throughout the corpus. That is, each eigenvector describes a subspace of the term vector space for which the terms are frequently found together. These *term-subspaces* describe a set of semantically significant associative patterns in the words of the underlying corpus of documents; we can think of each subspace as a *conceptual index* into the corpus (Furnas et.al. [16]).

For instance, an example term-subspace generated by analyzing a collection of military posts found three words as having significant re-occurrences, and therefore replicating together with success: ``harbor'', ``japan'', and ``pearl''. These term-subspaces make up our replicators and are our putative memes. Memes are not single re-occurring words but are made up of *sets* of re-occurring words.

Our final text analysis step is to ``compress'' the original term/document matrix by multiplying it with this right orthonormal matrix of eigenvectors (in other words we perform a projection). This, in effect, produces a *term-subspace/document* matrix. Each post is represented by a collection of weights where each weight now describes the degree to which a term-subspace is expressed within its post's text.

## 4 Meme and Quasi-species

### 4.1 Term-subspace as Putative Meme

We are looking for replicators within the corpus which are subject to natural selection. Elsewhere we have argued at length as to why the term-subspace captures the requirements of a true meme because its word sets act as a unit of selection within the corpus (Best [4, 5]; Pocklington & Best [27]). The strengths of this term-set as a replicating unit of selection are due to it meeting the following conditions:

- it is subject to replication by copying,
- it has strong copying fidelity,
- but not perfect fidelity, it is subject to mutation,
- it has a strong covariance with replicative success

(Lewontin [19]; Eigen [13]).

We will quickly review each of these points in turn.

SVD techniques exploit structure within the term/document matrix by locating co-occurring sets of terms. Clearly these term sets are replicating through the corpus since that is the precise statistical phenomena the SVD analysis detects. However it is not obvious that this replication is generally due to copying. Instances of precise copying occur when an in-reply-to thread includes elements of a previous post's text via the copying mechanism provided by the software system. Other instances of copying occur within a particular context or discussion thread when authors copy by hand words or phrases from previous posts into their new texts. More abstractly, replication occurs because certain memes are traveling outside of the NetNews environment (and thus outside of our means of analysis) and authors again act as copying agents injecting them into the corporal ecology. But, clearly, some re-occurrences are not due to copying but are a chance process where unrelated texts bring together similar words. The likelihood of such chance re-occurrences will be a function of the size and quality of our replicating unit. In summary, term-subspaces are instances of replication often due to copying.

The copying fidelity of a term-subspace is also a direct outcome of the SVD statistical analysis. But importantly, the copying fidelity of re-occurring term sets is not perfect across the entire corpus; the term sets will co-occur with some variation. These mutations are both changes *designed* by human authors and chance variation due to copying errors. In either case the mutations are *random* from the vantage of selection; in other words, human authors are not able to perfectly predict the adaptive significance of their inputted variations. These mutations work ``backwards'' into the actual term-subspace representation for a post organism. That is, a random mutation at the post level will actually result in a random mutation in the vector subspace representation (the memotype) for the post organism. In this way, the memes as represented in the memotype are subject to mutation.

Finally, we have elsewhere shown there can be a strong covariance between the replicative success of a cluster or thread of posts and the degree to which they express certain term-subspaces (Pocklington & Best [27]). In other words, a group of posts can increase its volume of activity over time by increasing the degree to which it expresses certain term sets within its post's text. This, then, is a covariance between the *fitness* of a population of posts and the expression of a particular *trait* as defined by a term-subspace. The demonstration of this covariance is critical to establishing that a replicator is subject to natural selection.

## 4.2 Quasi-species

If the term-subspace is a reasonable model for the meme then the term-subspace vector representation of a post is a good model of the post's *memotype*. Much as a genotype describes a point within genetic sequence-space for each organism, the memotype describes a point within conceptual sequence-space. By *sequence-space* we mean any of the search spaces defined by a replicator undergoing selection. Examples of sequence-spaces include the gene space, protein spaces under molecular evolution, and the meme space defined within a corporal ecology.

The notion of a *quasi-species* is due primarily to Manfred Eigen (Eigen, et.al. [12]; Eigen [13]). He states that the ``quasi-species represents a weighted distribution of mutants centered around one or several master sequences. It is the target of selection in a system of replicating individuals that replicate without co-operating with one another (RNA molecules, viruses, bacteria)," (Eigen

[13]). One organism is a *mutant* of another if it is particularly close to the other in sequence-space.

We wish to group our posts into quasi-species. This requires finding groups of memotypes which are centered together within the conceptual sequence space. To do so we employ a simple clustering algorithm, the Nearest Neighbor Algorithm (Jain & Dubes [17]). We first normalize each post memotype to unit length; this amounts to discarding text length information and representing only the *relative* strength of each meme within a text. The clustering algorithm then considers each post memotype in turn. The current memotype is compared to each memotype which has already been assigned to a cluster. If the closest of such vectors is not farther than a threshold distance, then the current vector is assigned to that cluster. Otherwise the current vector is assigned to a new cluster. This continues until each and every vector is assigned to a cluster.

This process assigns each post to a quasi-species defined as those posts which are close to one another in conceptual sequence-space.

The overall aim in grouping organisms is to bring to light certain evolutionarily significant relationships. Clearly, our quasi-species clustering method is ahistorical; that is, it does not directly account for decent when grouping together text organisms. The extent to which such groupings are effective when studying the relatedness of natural organisms is a matter of continued controversy as can be seen in the debates of the cladists versus evolutionary systematists versus pheneticists. While we are currently agnostic to this controversy we do agree with an original claim of the pheneticists: the more traits used when assessing the relatedness of individuals the more accurate are the groupings (Mettler, Gregg & Schaffer [23]).

We are in the happy situation of clustering based on the complete memotype for each of our organisms. The result is that under empirical verification our clusters exhibit extremely strong historical relatedness. We have found that the vast majority of texts clustered together come from the same in-reply-to thread and thus are related by decent (Best [5]). But our clustering method has the added benefit of grouping related texts even when the in-reply-to mechanism is not used and, alternatively, breaking up texts that are within the same thread but are not semantically related. This is of value since many posters to NetNews use the in-reply-to mechanism to post unrelated texts or, alternatively, post follow-up texts without bothering to use the in-reply-to facility. Thus, we claim that our clustering mechanism, due to its access to hundreds of traits, is actually superior at grouping together both related and descendent texts then would be a simple reliance on the threading mechanism. The clustering method meets our goal of illuminating evolutionarily significant relationships.

### 4.3 Comparison to Natural Ecologies

We are describing phenomena within a corpus of texts in terms of population ecology and population genetics. This is not simply a metaphorical device; we believe that interacting populations of texts and their constituent memes are evolving ecologies quite exactly. However there are clearly a number of interesting differences between genes and memes (as here operationally defined), natural organisms and texts, ecologies and corpora. Important differences include the driving forces behind mutation within the texts and the role of self-replication and

lineage within the corpora. We leave to future work a more complete analysis of these differences.

## 5 Models for Interacting Populations

We now turn to studying the interaction between quasi-species of posts. We have so far only studied the pairwise interactions between post quasi-species. Similar pairwise interactions have been widely studied within theoretical ecology. Consider two interacting populations: one population can either have a positive effect (+) on another by increasing the other's chance for survival and reproduction, a negative effect (-) by decreasing the other population's survival chances, or a neutral (0) effect. The ecological community has assigned terms to the most prevalent forms of pairwise interaction, in particular:

- Mutualism (+, +)
- Competition (-, -)
- Neutralism (0, 0)
- Predator/prey (+, -)

(Pielou [25]; May [22]).

Our goal is to study the pairwise interactions of quasi-species within the corporal ecology with the hope of discovering some of these interaction types.

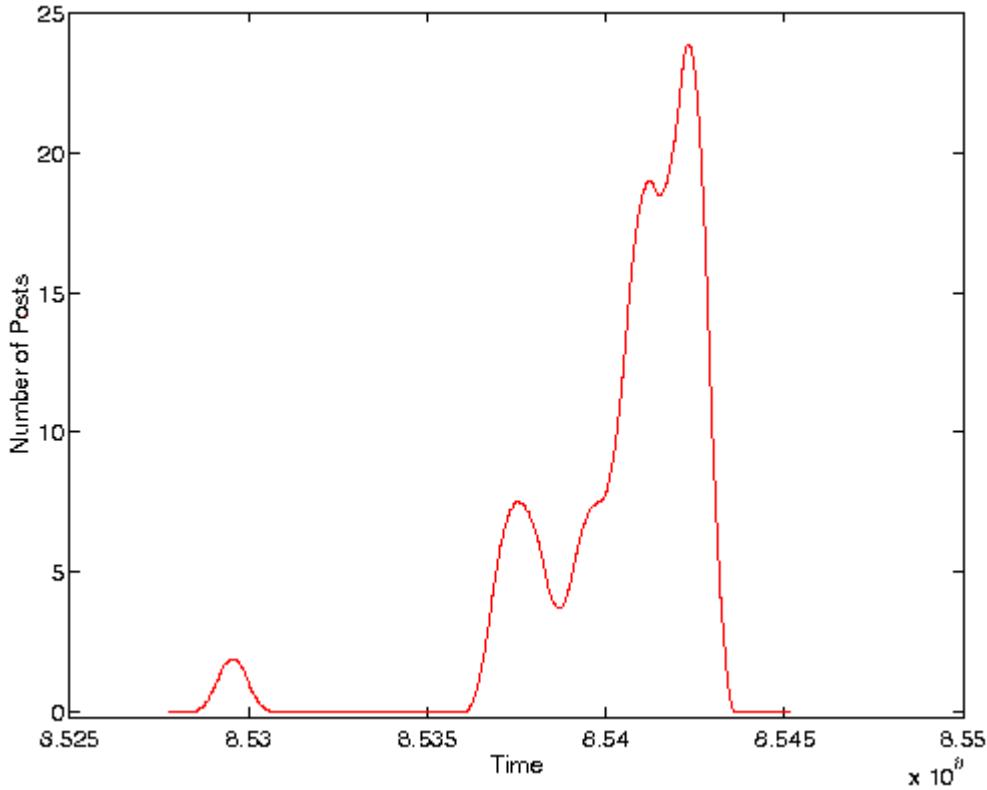
### 5.1 Time Series

To study how the interactions of populations affects growth rates we must define a method to measure a quasi-species' growth over time. Recall that a quasi-species describes a collection of posts which are close to one another in sequence-space. Each of these posts has associated with it a *timestamp* identifying when that text was posted to the system; in effect, its birth time and date. (Note that a post organism has something of a zero-length life-span; it comes into existence when posted but has no clear time of death.)

A histogram of the timestamp data is created with a 24 hour bucket size. That is, for each quasi-species we count how many member texts were posted on one day, how many on the next, and so forth through the entire population of texts. The datasets currently used span on the order of two weeks and consist of thousands of posts. So for each day a quasi-species has a volume of activity which can range from zero to 10's of posts. This rather coarse unit, the day, has been chosen to neutralize the strong daily patterns of post activities (e.g. activity may concentrate in the afternoons and drop off late at night; different time zones will shift this behavior and thus encode geographic biases). Thus the patterns of rise and fall in the volume of posts within a quasi-species when measured at the day level will, hopefully, reflect true changes in interest level and authorship activity rather than other external or systemic factors

## 5.2 The Test Corpus

Figure 2 is a typical graph for the volume of posts within a particular quasi-species over a period of ten days. This cluster was found within a corpus of all posts sent to the soc.women newsgroup between January 8, 1997 (the far left of the graph) and January 28, 1997 (the far right). In the figure the number of posts in a day is represented by the height of the graph. This particular cluster of texts exhibited an initial set of posts, a few days worth of silence, then a rapid building up of activity which then declined precipitously at the end of the dataset. The entire corpus used consisted of 1,793 posts over the same ten day period. The clustering mechanism arrived at 292 quasi-species the largest of which contained 103 posts.



**Figure 2.** Typical time series of posts to quasi-species. Time axis is measured in seconds since Jan. 1 1970.

## 5.3 Time Series Cross-correlation

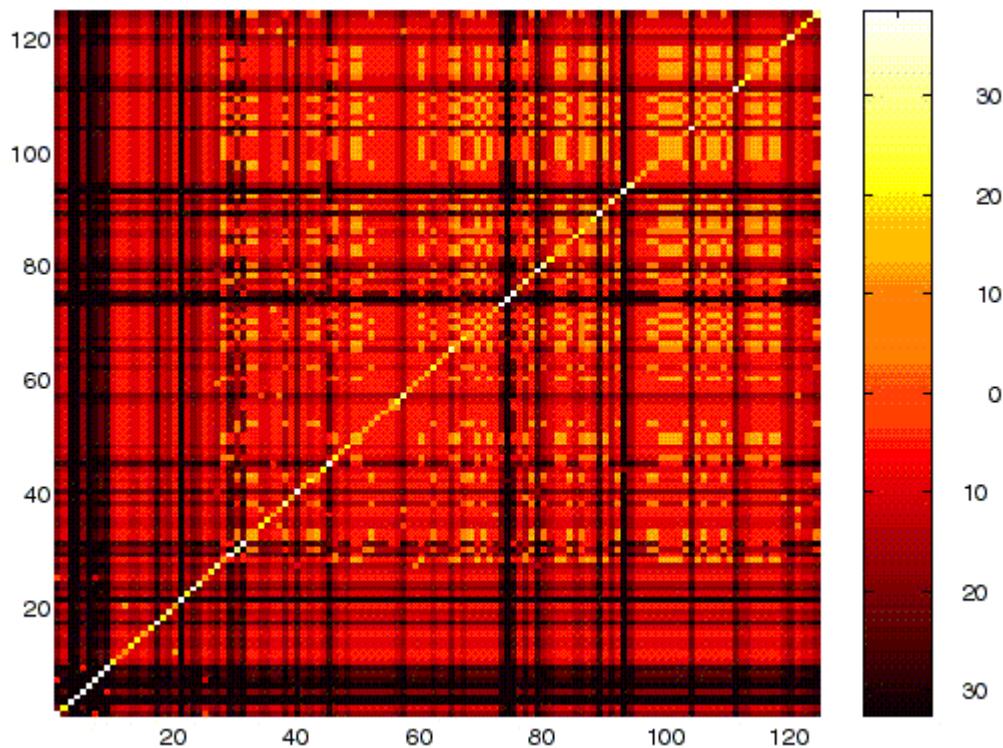
To study the relationship between the time series of two populations of posts we use the cross-correlation function. The use of the cross-correlation to study bivariate processes, and time series in particular, is well known (Chatfield [6]). Each time series is normalized to be of zero mean and unit standard deviation; that is, we subtract off the mean and divide by the standard deviation. In this way, the cross-correlations will not be dominated by the absolute volume of post activity within some cluster and instead will be sensitive to both large and small sized clusters.

We assume the readers are familiar with the regular covariance and correlation functions. Then the cross-correlation for two time series,  $X$  and  $Y$ , is given by,

$$xy = \frac{xy}{\sqrt{xx yy}}$$

Here  $xy = \text{Cov}(X, Y)$  and  $xx$  and  $yy$  are the variance of  $X$  and  $Y$  respectively. Note this formulation only considers the cross-correlation for a zero time lag. That is, it considers how the two time series are correlated for identically matching points in time. With a nonzero lag the cross-correlation would study cases when the two series might have correlations offset by some fixed amount of time. Since we group our time data into day-long chunks the zero-lag cross-correlation will be sensitive to covariances which have a time offset as large as 24 hours; this builds into the time series an adequate time lag.

When the cross-correlation between two sets of data is significantly different than zero it suggests the two sets of data have some relationship between them. A positive value means an increase in one series is likely to co-occur with an increase in the other series. A negative value means an increase in one series is likely to co-occur with a decrease in the other series.



**Figure 3.** The pairwise time series cross-correlation for 125 largest quasi-species clusters.

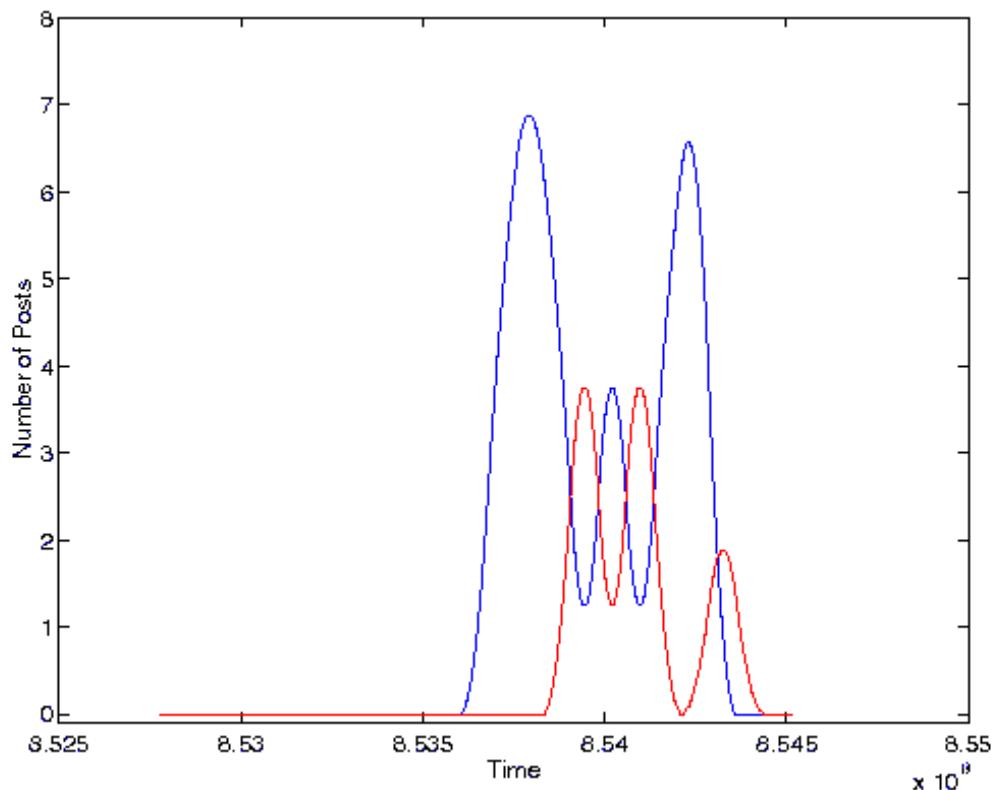
Figure 4 shows the pairwise cross-correlations for the 125 largest quasi-species clusters within our corpus. The diagonal represents the cross-correlation between a time series and itself which, as expected, is identically one. Note that the matrix is symmetric about the diagonal. The off-diagonal values range from near one to -0.26. The mean cross-correlation is 0.3. This value is quite high, indicating that most of these post clusters are somehow positively related. We suspect this high average cross-correlation is at least partially due to external or systemic affects which were not removed by the day-long bucket size. For instance, our analysis would be sensitive to

patterns due to the Monday-Friday work week common in the West. Further, some of this correlation may be due to a high level of mutualistic interactions amongst the posts. Clearly, the ideas conveyed within the soc.women newsgroup often share similar contexts.

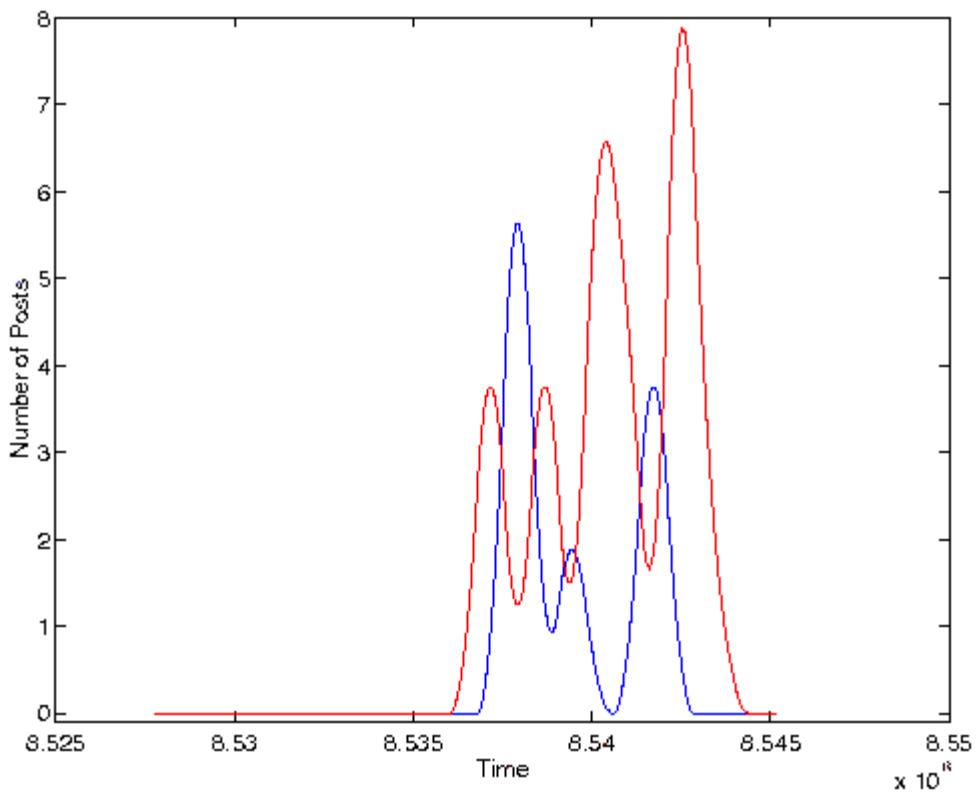
In our analysis this overall high correlation does not particularly matter since we are concerned with the *relative* cross-correlation -- that is, those that are the largest and those that are the smallest.

## 6 Negative Cross-correlations: Competition versus Predator/Prey

We have primarily studied those pairs of quasi-species with relatively strong negative cross-correlations; to wit, those where  $x_{xy} \leq -0.2$ . Note that in all such cases (there are 42)  $P < .001$  suggesting that with extremely high probability the correlations are not due to chance. Figure 4 and Figure 5 plot two such interactions, both fairly characteristic of this population.



**Figure 4.** Volume of activity for two quasi-species. The cross-correlation between these two series is -0.26.



**Figure 5.** Volume of activity for a different set of two quasi-species. The cross-correlation between these two series is -0.23

Both of these figures demonstrate a clear negative covariance between the volume of activity of the two post clusters. This negative covariance is both statistically significant and visually compelling. But what do these graphs signify and can it be interpreted within the rubric of ecological interactions?

At first glance the interactions appear to be of a predator/prey variety; they have a (+, -) relationship to them. However, competition might also produce similar interaction phenomena if the competitors are operating close to some limitation or environmental carrying capacity. In such instances the relationship between population sizes will be a zero-sum game, when one goes up the other must come down. To be able to classify the interactions of Figures 4 and 5 we need to consider the qualitative details of these two interactions through direct study of the texts.

Recall that in the case of a predator/prey relationship, one population enjoys an increased growth rate at the *expense* of another population (e.g. one population feeds on the other). The presence of a relatively large population of predators will result in a diminished level of success for the prey (they get eaten up). Conversely, the relative absence of prey will result in diminished success for the predator (they have nothing to eat).

Now consider the case of competition. In competition two interacting populations inhibit each other in some way, reducing each other's level of success. This often occurs when the two

populations rely on the same limited resource. Unlike the predator/prey relationship where the predator requires the prey for success, with competition the two populations would just as soon avoid each other all together.

This pressure towards avoidance is the source of much ecological diversity since it propels populations to explore new and therefore competition-free niches (Pianka [24]). An ecological *niche*, for some particular species, is simply that collection of resources the species relies on. Interspecific niche overlap occurs when two or more species share one, some, or perhaps all of their resources. When those resources are scarce, interspecific competition will result. The *width* of a niche is simply a qualitative sense of the variety and number of resources a population makes use of.

## 6.1 Competition and Niche Behavior

We have studied the set of posts which make up the four quasi-species shown in Figures 4 and 5 in an attempt to qualitatively classify their interactions. The quasi-species of Figure 4 are made up of posts within a single thread. The subject line for these posts reads, ``Men's Reproductive Rights''. In general, these posts are concerned with the responsibilities and rights of men towards their unborn children. The quasi-species displayed with a dashed line in the figure is centered around the use of contraceptives. It consists of a collection of posts wherein the authors debate who is most responsible, the women or the man, when using contraception. The quasi-species with a solid line deals instead with the use of abortion and whether the father has any intrinsic rights in deciding whether or not to abort an unborn child.

In Figure 5 these two quasi-species are also from a single thread. The subject line here reads, ``Unequal distribution of wealth?''. This particular thread of discussion was rather large. In fact there was a total of 365 posts to this thread which our text analysis tools broke up into a number of quasi-species due to significant bifurcations of the topic. In other words, many parallel discussions occurred all within a single in-reply-to thread. The cluster of discussion shown with the solid line in Figure 5 centered around a debate as to whether the US military was a ``socialist collective''. The quasi-species with the dashed line was a debate on the value of releasing the mentally ill from hospitals. Clearly, these two debates are quite dissimilar even though they span the same set of days and are posts to the same discussion thread.

The quasi-species of Figure 4 are different but related discussions. Those of Figure 5 are different and not clearly related. Still, we believe that both of these sets of interactions demonstrate elements of competition. Within the texts there is no evidence of predator memes; in Figure 5, in fact, the memes seem entirely orthogonal to one another.

However, in both examples the memes are competing for the same collection of human authors who must act as their agents if they are to propagate and succeed. This seems even more likely when we consider that all these posts are to the same newsgroup which due to its narrow subject area supports only a limited supply of human posters. Moreover, each pair of interactions are confined to a single thread of discussion, which again has an even more limited set of potential human authors since users of the NetNews system often zero-in on particular threads they find interesting and ignore others. After inspecting most of the interactions which demonstrated

strong negative correlations we observed no examples of predator/prey interactions but many instances which appeared to be examples of competition.

### 6.1.1 Statistical Artifacts

We computed the cross-correlation between 125 different clusters, arriving at 15,625 different correlations. It is possible, therefore, that the cross-correlations with large negative values exist simply by chance; they represent the tail of the distribution of correlations.

However, we believe that our qualitative analysis provides strong evidence that these negative correlations are *not* artifacts but are indeed due to an interaction phenomena between the two quasi-species. The two pairs of quasi-species described in detail above demonstrate this point. The likelihood that two quasi-species be brought together by mere chance *and* both be from the same thread (out of 324 threads within the corpus) seems vanishingly small.

### 6.1.2 Competition

We now will test our theory that these interactions are of a competitive nature. Again recall that competition is often caused by populations existing within the same (narrow) ecological niche. What makes up an ecological niche for a meme within NetNews? We argue that the newsgroups themselves make up spatially distributed ecological niches. Since there is relatively little interaction between newsgroups (save the phenomena of cross-posting) we would expect these niches to behave something like island ecologies -- they remain relatively isolated from each other. Within a single newsgroup (which is all we have studied so far) niches might be described by threads of discussions. As previously stated, we have found that individual posters to the system tend to become involved in particular in-reply-to threads which interest them. Thus the memes within a particular thread make use of a set of human resources which is smaller than the entire set of potential human resources available to the newsgroup. These resources define the niche.

We theorize that cross-correlations which approach -1 in our corpus are examples of competition, and competition will be more likely between populations which are posted to the same threads and thus have overlapping niches. The most direct way to test this theory is to see if negative cross-correlations between two quasi-species correlates with the degree to which they post to the same threads. For each of the 125x125 pairwise interactions we computed the number of threads each of the quasi-species pairs had in common and divided that by the total number of threads posted to by each quasi-species. For example, one quasi-species may contain posts which went to two different in-reply-to threads. Another quasi-species may have posts which span three different threads one of which is identical to a thread within the first group. So this pair of quasi-species would have posted to a total of four different groups one of which was shared. Their relative niche overlap would therefore be 0.25.

We calculated the correlation coefficient between the negative cross-correlations of Figure 3 and the percentage of thread overlap between these quasi-species pairs. We found this correlation to be -0.04. While this correlation is statistically significant ( $P < .001$ ), it is not very pronounced. The negative sign, though, does indicate that as the level of competition increases (a negative

cross-correlation) the percent of overlap of their niche also increases (a larger positive shared thread percentage).

This small correlation coefficient may be due to a small signal/noise ratio. Since most pairwise interactions result in small correlations, the relative number of large negative correlations is quite small. The number of interactions grows with the square of the number of quasi-species. We suspect that a simpler experiment which grows linearly with the number of quasi-species will have a better signal/noise ratio.

We have studied the correlations between the absolute number of in-reply-to threads a quasi-species is posted to and the average degree to which the quasi-species finds itself correlated with other clusters. Our hypothesis is that the absolute number of threads a quasi-species is posted to will be related to the average degree of competition the quasi-species experiences in its interactions. Since the variety of resources used by an entity defines it niche, if a quasi-species is posted to a relatively small number of threads then it exists in a narrow ecological niche. Should there subsequently be any interspecific overlap of these narrow niches, scarcity will result in competitive encounters. We computed the correlation coefficient between the total number of threads within a quasi-species and its *average* cross-correlation value. The correlation coefficient here is 0.25. Thus, as the number of threads within a quasi-species increases (the set of available resources is widened) the average level of competition diminishes (the mean pairwise cross-correlation also increases). This correlation is statistically significant ( $P < .001$ ) and rather pronounced.

We further computed the correlation coefficient when the absolute number of threads was normalized by the size of the quasi-species. We might expect that the number of threads employed by a quasi-species would grow with the number of posts within that quasi-species. In other words, as a quasi-species gets larger the number of threads increases too. This might affect the analysis above such that instead of measuring niche width we were simply measuring quasi-species size. Dividing out the size amounts to computing the *average* number of threads employed by a post for a given quasi-species. When this set of values was correlated with the mean cross-correlation we arrived at a nearly identical coefficient as above and again clear statistical significance. Thus quasi-species size is not a major factor in level of competition.

## 7 Conclusions

We have described a set of text analysis tools, based primarily on Latent Semantic Indexing, which distill replicating memes from a corpus of text. We have trained this analysis system on a corpus of posts to NetNews. This makes up a corporal ecology where the posts are organisms, NetNews is the environment, and human authors are a scarce resource. We argue that this represents an important bridging of text analysis and the Alife research program. Further, it amounts to a novel shift for Alife research -- rather than synthesizing life-like agents we are analyzing a pre-existing environment and discovering life-like behaviors.

In results reported here we group together posts which make use of similar sets of memes. These groups, clouds within a conceptual sequence-space, describe quasi-species. For each quasi-species we compute its time-wise volume of activity by histogramming its daily post levels. We

then study the pairwise interaction between quasi-species by computing the cross-correlations between their time series. In our corpus, strong negative cross-correlations signify conditions of competition between the interacting populations where the quasi-species are competing for a limited set of human authors. Furthermore, quasi-species with relatively narrow ecological niches, those which make use of a small number of in-reply-to threads, are more likely to be in competition with other quasi-species. This behavior is analogous to what is found in natural ecologies (Pianka [24]).

Why do these quasi-species compete? Qualitative analysis of the posts, such as those described in the previous section, shows that many competing quasi-species are posts sent to the same or similar threads. Competition is over the scarce authorship resources within these specific thread niches. Over time a particular thread of discussion may bifurcate into two or more internal themes which then proceed to compete for ``air-time'' within the thread.

## Acknowledgments

We thank the two anonymous reviewers for their helpful critique of this paper. We gratefully acknowledge the significant input of Richard Pocklington to this work and his careful reading of a draft of this paper. Laurie Hiyakumoto also gave very helpful comments on a draft of this paper. We thank Bernd Schoner and Michael Patrick Johnson. Finally, the author is pleased to thank the sponsors of the MIT Media Laboratory for their financial support.

## References

- [1] Berry, M.W. (1992) Large-scale Sparse Singular Value Computations. *The International Journal of Supercomputer Applications*. Vol. 6, No. 1.
- [2] Berry, M., T. Do, G. O'Brien, V. Krishna, & S. Varadhan (1993). SVDPACKC (Version 1.0) User's Guide. University of Tennessee Computer Science Department Technical Report, CS-93-194.
- [3] Berry, M.W. & R.D. Fierro (1995). Low-Rank Orthogonal Decompositions for Information Retrieval Applications. University of Tennessee Computer Science Department Technical Report, CS-95-284.
- [4] Best, M.L (1996). An Ecology of the Net: Message Morphology and Evolution in NetNews. Massachustes Institute of Technology, Media Laboratory, Machine Understanding Technical Report, 96-001.
- [5] Best, M.L. (1997). An Ecology of Text: Using Text Retrieval to Study Alife on the Net. To appear *Journal of Artificial Life*.
- [6] Chatfield, C. (1989). *The Analysis of Time Series An Introduction*. London: Chapman and Hall.
- [7] Dawkins, R. (1976). *The Selfish Gene*. New York, Oxford University Press.

- [8] Dawkins, R. (1982). *The Extended Phenotype*. San Francisco, WH Freeman.
- [9] Deerwester, S. S.T. Dumais, G.W. Furnas, T.K Landauer, and R. Harshman (1990). Indexing by Latent Semantic Analysis. *Journal of the American Society for Information Science*. 41(6): 391-407.
- [10] Dumais, S.T. (1992). LSI meets TREC: A status report. In *The First Text REtrieval Conference (TREC-1)*, ed. D. Harman. NIST Special Publication 500-207.
- [11] Dumais, S.T. (1993). Latent semantic indexing (LSI) and TREC-2. In *The Second Text REtrieval Conference (TREC-2)*, ed. D. Harman. NIST Special Publication 500-215.
- [12] Eigen, M.J, J. McCaskill, & P. Schuster (1988). Molecular Quasi-Species. *Journal of Physical Chemistry*, Vol. 92, No. 24.
- [13] Eigen, M. (1992). *Steps Towards Life: A Perspective on Evolution*. Oxford: Oxford University Press.
- [14] Foltz, P.W. (1990). Using Latent Semantic Indexing for Information Filtering. *Proceedings of the 5th Conference on Office Information Systems*. ACM SIGOIS Bulletin vol. 11, issues 2,3.
- [15] Frakes, W.B. & R. Baeza-Yates, eds. (1992). *Information Retrieval: Data Structures and Algorithms*. Englewood Cliffs, New Jersey: Prentice Hall.
- [16] Furnas, G.W., S. Deerwester, S.T. Dumais, T.K. Landauer, R.A. Harshman, L.A. Streeter, & K.E. Lochbaum (1988). Information Retrieval using a Singular Value Decomposition Model of Latent Semantic Structure. *Proceedings of the 11th International Conference on Research and Development in Information Retrieval (SIGIR)*. New York: Association for Computing Machinery.
- [17] Jain, A.K. & R.C. Dubes (1988). *Algorithms for Clustering Data*. Englewood Cliffs, New Jersey: Prentice Hall.
- [18] Kantor, B. & P. Lapsley (1986). Network News Transfer Protocol: A Proposed Standard for the Stream-Based Transmission of News. Internet RFC-977.
- [19] Lewontin, R.C. (1970). The Units of Selection. *Annual Review of Ecology and Systematics*, Vol. 1.
- [20] Lewontin, R.C. (1974). *The Genetic Basis of Evolutionary Change*. New York, Columbia University Press.
- [21] May, R.M. ed. (1981A). *Theoretical Ecology Principles and Applications*. Oxford: Blackwell Scientific Publications.
- [22] May, R.M. ed (1981B). Models for Two Interacting Populations. In (May 1981A).

- [23] Mettler, L. E., T.G. Gregg, & H.E. Schaffer (1988). *Population Genetics and Evolution*, 2nd ed. Englewood Cliffs, NJ, Prentice Hall.
- [24] Pianka, E.,R. (1981). Competition and Niche Theory. In (May 1981A).
- [25] Pielou, E.C. (1969). *An Introduction to Mathematical Ecology*. New York: Wiley-Interscience.
- [26] Pocklington, R. (1996). *Population Genetics and Cultural History*, Msc Thesis, Simon Fraser University, Burnaby.
- [27] Pocklington, R. & M.L. Best (1997). Cultural Evolution and Units of Selection in Replicating Text. To appear *Journal of Theoretical Biology*.
- [28] Salton, G. & C. Buckley (1988). Term-weighting Approaches in Automatic Text Retrieval. *Information Processing & Management*. 24: 5, 513-523.

# A LETTER ON: MEMES ON MEMES A CRITIQUE OF MEMETIC MODELS

By  
**Michael L. Best**  
Media Laboratory,  
Massachusetts Institute of Technology,  
Cambridge, Massachusetts 02139, USA.  
[mikeb@media.mit.edu](mailto:mikeb@media.mit.edu)

In these pages and elsewhere a lot has been made of the "new science of memetics". I argue that this phrase misleads. First, work on evolutionary models and indeed replicator based models of cultural phenomena is not terribly new, though too much current work in the field ignores this history. Second, the scientific progress made in the area, if by science one means formalized models, is often quite immature. An exemplar of these weaknesses is the recent contribution to this journal by Aaron Lynch (1998).

Evolutionary theories of cultural change pre-date Darwin himself. By the turn of the century evolutionary theories had established themselves as dominant models of epistemology (e.g. Spencer 1864; see Campbell 1974) and language change (e.g. Lyell 1863; Darmesteter 1886; see Nerlich 1990). In other words, evolutionary theories of "ideas" dominated late nineteenth century thought and are certainly no newer than evolutionary theories for organic biology.

Examining the history of organic evolution one sees that theories of the mechanisms for change (namely natural selection (Darwin 1859)) predate modern theories for transmission (namely population genetics (Haldane 1932; Fisher 1930)). The same seems to be true for cultural evolution with theories for change established well before transmission models. Nonetheless, in the last twenty years a collection of formal replicator based transmission models for cultural phenomenon have been proposed (Cavalli-Sforza & Feldman 1973; Lumsden & Wilson 1981; Cavalli-Sforza & Feldman 1981; Boyd & Richerson 1985; Barkow 1989; Durham 1991; reviewed in Durham 1990).

Dawkins (1976), in a nice turn of phrase, proposed the word "meme" to describe replicating cultural units. However, other terms have also been put forth for this purpose including "culturgen", "symbol", "culture-type", and "theme". Lynch (1998) claims to introduce a new neologism, the "mnemon", though in point of fact this term was first proposed by Campbell (1974). Debate continues on what is meant exactly by these various terms.

In his recent contribution, Lynch (1998) proposes a formal transmission model for population memetics. Thus, he is working within the mode of, in particular, Cavalli-Sforza, Feldman, Boyd, and Richerson; though he does so without reference to any of their work. He devotes some time to the description of differential equations which aim to model the dynamics of two distinct memes which vie for hosts within a population.

A formal model should aim to describe empirical data (present or sought after) and should shed explanatory light on some phenomena. Lynch's do neither. This is due to a number of problems - the wrong tool, focused at the wrong areas of the problem, creating a model which is too complex to make progress with, which encapsulates a set of assumptions that are neither stated nor justified.

$$\begin{aligned}
\frac{dN_1(a,t)}{dt} = & \int_a^{\infty} R_1(p-a) K_{11}(p,a) N_1(p,t) dp \\
& + \int_a^{\infty} R_2(p-a) K_{21}(p,a) N_2(p,t) dp \\
& + \int_0^{\infty} \gamma_{12}(p,a) N_1(p,t) N_2(a,t) dp \\
& + \int_0^{\infty} \frac{\beta_{12}(p,a) N_1(p,t) N_2(a,t)}{N(t)} dp \\
& - \alpha_1 N_1(a,t) + \alpha_2 N_2(a,t) \\
& + \frac{\partial}{\partial a} N_1(a,t) - M_1(a) N_1(a,t)
\end{aligned}$$

Equation (1)

Consider just the first term to Equation 1, taken from Section 16 of Lynch's paper. Here Lynch describes the change with time,  $t$ , of the acquisition of a particular meme in a population of a given age due to transmission from parents. This is expressed in terms of the number of parents at time  $t$  of some particular age  $p$ ,  $N_1(p,t)$ , the number of offspring they have of the appropriate age,  $R_1(p-a)$ , and the ratio of parents who possess the meme that transmit it to their offspring given these ages,  $K_{11}(p,a)$ . Note that this is given as a ratio and not a probability resulting in a deterministic model. Lynch chooses to model everything as continuous functions with time or age and to integrate over all age groups whenever he wants a fixed number. The misapplication of the continuous domain is evident by the fact that even within his formalisms Lynch is compelled to discretise time. Here,  $K_{11}(p,a)$  is the ratio per year of a transmission event and not a continuous ratio over time. I am not sure what it means to integrate over continuous ages a value measured in discrete time steps and to, subsequently, differentiate that over continuous time.

This model, naturally, focuses our attention on its parameterized spaces. Namely, it asks us to consider how transmission dynamics change with respect to the age of parent versus offspring. I question the value of this particular focus of attention. While age specific models might indeed be worthy of exploration we are offered no arguments as to why this particular age parameterization is relevant. In particular, why would the age of the parent be given such attention? While Lynch correctly states that a model must hide certain elements while exposing others I question those he has chosen to expose. Other factors to consider include the relationships between model and learner (parental and non-parental models), the numerical relationships between model and learner, the complexity of societies, spatial distributions, modes of communication, and so forth. Further, the model makes a number of assumptions which go unstated and unsupported. For example uniparental transmission; there is no consideration of two parents. Such a simplification demands defense and should not be included within the model without comment.

Consider again,  $K_{11}(p,a)$ , the ratio per year of a transmission event between a parent of age  $p$  and an offspring of age  $a$ . For this model to be of explanatory value we need to understand what this ratio  $K_{11}$  might look like in two-dimensional space. What would it mean for it to be concave, convex, increasing in  $a$ , increasing in  $p$ , and so forth. These are the sort of questions whose answers might provide value to these models. But Lynch provides no such explanation. He simply states the existence of this ratio,  $K_{11}$ . But it is this very ratio which goes to the phenomena Lynch presumes to be describing – namely, what are the dynamics of transmission between parent and offspring. From the vantage of population memetics, the action is in the likelihood of and factors around successful transmission. Lynch's formulation couches that phenomenon in a quite complex set of continuous functions who's born when and how does that vary with transmission. But it does not make any explanatory progress towards understanding these transmission dynamics. Moreover, it presents a deterministic ratio for the transmission event without offering any hope for actually arriving at this fixed (given a set of ages) rate of transmission. This side-steps what should be the central question of study what is the probability distribution of these transmission events given a sound parameterization.

The overall complexity of Lynch's differential equation also is a drawback. It would be useful to understand, even within his parameterized spaces, what factors would lead to fixation of one meme versus another, what factors would lead to a frequency dependent mix, and so forth. However the sheer complexity of the model makes analytic (and indeed numerical) progress unlikely.

In Section 5 Lynch also presents a more simplified method to symbolize memetic transmission events. While the simplicity I applaud I find the overall mechanisms implausible. Lynch proposes to model the transmission of a "negatively defined mnemonic". This is, apparently, some sort of "forgetting" event where a meme is removed from an agent's memory. Lynch supports this mechanism through an argument by definition he claims that this is consistent with his usage of "memory abstraction". While no doubt one could build a consistent model of the transmission of forgetting that does not make it a mechanism important (indeed even present) within human cultural evolution. Lynch goes on to claim that this is why birth control is a successful meme; it is easier to learn than to forget such concepts. Here we see how an implausible formal model leads to an unjustified conclusion. It seems far more likely that the concept of birth control

maintains itself due to its epistatic co-adaptive relationships to other meme complexes which render it stable in the presence of small numbers of competing memes. But, again, this is the very sort of question that Lynch's models should aim to investigate; our arm-chair theorizing amounts to nothing against well supported models and, of course, empirical data.

My critique of this model is towards a broader purpose. I hold building a science of memetics as an admirable project. Slow progress towards such a science has been made as evidenced by much of the cited literature. But the misapplication of science, that is formal models, does damage to the program. Mathematics, even if it is correct in some formal sense (and even this I question in the case of Lynch), is not of use unless it helps to explain the phenomena in question.

There exist a few practical methods to insure progress with formal modeling. Perhaps the most critical is an open and lively dialogue on these very matters; I hope this essay can help to stimulate such a discussion. Second, it is best to test formal models on real empirical data. In this way predictions can be verified, fit can be measured, hypotheses can be tested. Sadly, empirical data is not always available. Alternatively, computational simulation and analysis can act as a surrogate for empirical data. In this case, hypotheses are stated, and analysis (numerical or symbolic) can serve to test these hypotheses. For instance, if Lynch is keen to employ an age-specific model, he might study what relationships between ages and transmission probabilities would result in points of fixation. Lynch, it must be said, makes some of these exact points himself; he argues for the value of falsifiable models. But, sadly, he does not take his own advise and try to support (or falsify) his own arguments.

Finally, the scientific method relies heavily on an appreciation of past art. Lynch would do well to study the transmission models (which include continuous-time models) within the literature (if he has studied them he would do well to cite and make use of their results). Indeed, cultural transmission is not the only important related work. I find myself, for instance, looking to research on social learning in animals, behavioral ecology, artificial life, and so forth.

## Acknowledgements

I would like to thank Richard Pocklington and Hans-Cees Speel for their comments on a draft of this letter.

## References

- Barkow, J.G. (1989). *Darwin, sex, and status: Biological approaches to mind and culture*. Toronto: University of Toronto Press.
- Boyd, R. & Richerson, P.H. (1985). *Culture and the evolutionary process*. Chicago: University of Chicago Press.
- Campbell, D.T. (1974). Evolutionary epistemology. In P.A. Schilpp (Ed.). *The philosophy of Karl Popper*. La Salle, IL: Open Court Publishing.

Cavalli-Sforza, L. & Feldman, M. (1973). Models for cultural inheritance, I: Group mean and within group variation. *Theoretical Population Biology* 4: 42-55.

Cavalli-Sforza, L. & Feldman, M. (1981). Cultural transmission and evolution: A quantitative approach. Princeton, NJ: Princeton University Press.

Darmesteter, Arsène (1886). The life of words as the symbols of ideas. London: Kegan Paul, French & Co.

Darwin, C. (1859). On the origin of species by means of natural selection, or the preservation of favored races in the struggle for life. London: John Murray.

Dawkins, R. (1976). The selfish gene. Oxford: Oxford University Press.

Durham, W.H. (1990). Advances in evolutionary culture theory. *Annual Review of Anthropology*, 19, 187-210.

Durham, W.H. (1991). Coevolution: Genes culture and human diversity. Stanford, CA: Stanford University Press.

Fisher, R.A. (1930). The genetical theory of natural selection, 2nd rev. ed., 1958. New York: Dover.

Haldane, J.B.S. (1932). The causes of evolution. London: Longmans, Green.

Lumsden, C.J. & Wilson, E.O. (1981). Genes, mind and culture: The coevolutionary process. Cambridge, MA: Harvard University Press.

Lyell, C. (1863). The geological evidences of the antiquity of man, with remarks on theories of the origin of species by variation. London: John Murray.

Lynch, A. (1998). Units, events and dynamics in memetic evolution. *Journal of Memetics - Evolutionary Models of Information Transmission*, 2.  
[http://jomemit.cfpm.org/1998/vol2/lynch\\_a.html](http://jomemit.cfpm.org/1998/vol2/lynch_a.html).

Nerlich, B. (1990). Change in language: Whitney, Bréal, and Wegener. London: Routledge.

Spencer, H. (1864). Illustrations of universal progress; a series of discussions. New York: D. Appleton and Company.

# THE LIFECYCLE OF MEMES

By

Henrik Bjarneskans, Bjarne Grønnevik and Anders Sandberg

## Abstract

Memes, self reproducing mental information structures analogous to genes in biology, can be seen as the basis for an explanatory model of cultural and psychological behaviour. Their properties and effects are evolutionary conditioned and ultimately seeks to promote their replication. To survive in a context the memes must meet certain conditions. We abstract a model of these conditions and use it to analyse three well-known memes: the "Kilroy was here" graffiti, urban legends and Christianity.

## 1. What is a Meme?

Memes were originally described by Richard Dawkins in his book *The Selfish Gene* (1976) [1]

*A unit of cultural transmission or a unit of imitation*

Examples of memes are tunes, ideas, catch-phrases, clothes fashions, and ways of making pots or building arches. Just as genes propagate in the gene pool via a process which, in the broad sense can be called imitation. If a scientist hears, or reads about, a good idea, he passes it on to his colleagues and students. He mentions it in his articles and his lectures. If the idea catches on, it can be said to propagate itself, spreading from brain to brain.

What makes the meme concept so powerful is its close analogies to the theory of natural selection. Natural selection occurs whenever the following conditions exist (Dennet 1990):

- Variation: a continuing abundance of different elements.
- Heredity or replication: the elements have the capacity of creating copies or replicas of themselves.
- Differential "fitness": the number of copies of an element that are created in a given time varies, depending on the interaction between the features of that element (whatever it is that makes it different from other elements) and features of the environment in which it persists.

This is a quite general definition which is not limited to biology, and suggests that memes are subject to natural selection: they vary (due to "mutations" in transmission or mental storage, plus deliberate changes), they replicate (by definition) and have differing fitness. This leads to phenomena of competition, co-evolution, population dynamics and adaptation surprisingly similar to their biological counterparts. The set of shared memes form the memepool (in analogy with the genepool).

It should be noted that human decisions are part of the memetic selection process; from the perspective of the memetic ecology humans and human behaviour corresponds to the climatic and geological environment of biological life. In the meme perspective, it is more accurate to say that the message has evolved into its form in order to encourage people to spread it than to say that people have selected or 'bred' the message into its form.

Although people often make the decision to spread a meme or not consciously, this process is influenced by the meme. Some memes are viewed as important, and hence spread to others after a conscious and sometimes rational evaluation; some memes exploit aspects of cognition or emotion to bias their hosts to spread them. Natural selection favors memes that are good at reproducing, which suggests that in time there will exist many memes that are very efficient replicators. Their accuracy is irrelevant for their survival, only their ability to replicate and find new hosts; memes that interest people and encourage them to spread the meme will thrive at the expense of less attractive versions.

What makes the meme perspective so interesting is that it suggests that some of what we have learned from biology can be applied to human psychology. Dawkins points out that "a cultural trait may have evolved in the way it has simply because it is *advantageous to itself*". Gross (1996) says

*The main shift in thinking that needs to take place is to look at the spread of the legend not so much from the point of view of the people who propagate the warning, but from the point of view of the warning itself.*

In memetics, ideas are viewed as almost independent creatures in a symbiotic relationship with human minds and cultures.

## 1.1 Defining Memes

The meme concept is somewhat slippery to define, and there is an multitude of definitions ranging from the very wide to the very narrow. The definition of meme we will use in this essay is

*A meme is a (cognitive) information-structure able to replicate using human hosts and to influence their behavior to promote replication.*

This is a somewhat strict definition, since it excludes many structures able to replicate without influencing host behavior or using non-human hosts such as chimpanzees, dolphins and computers. It can be seen as a subset of the more general memes described by Dawkins.

Memes do not only influence behavior to promote replication, but many of the most successful memes have other side-effects (for example, being able to invoke various emotions) or promote their replication by being useful or through other features (like parasiting on other memes, e.g. parodies and imitations); using a biological analogy one could say symbiotic memes spread mainly using their usefulness, while parasitic memes compel the host to spread them. This compulsion can be more or less subtle, ranging from explicit orders like in chain letters ("Send

ten copies of this letter to your friends") to implicit influences that link with our attitudes like the "Save the whales" meme described in (Hofstadter 1985, p. 55).

It is quite common that memes are confused with ideas/thoughts. Both are cognitive structures, but an idea is not self-replicating and is spread passively (i.e. for extrinsic reasons) if it is spread beyond its initial host at all. The difference is sometimes hazy; the idea "Isn't it time for us to eat?" can easily spread in a small group, but will not spread well outside the group and will disappear once the question is settled, while a meme usually can spread generally and does not have any limited lifespan.

It should also be noted that memes often form meme complexes, groups of memes mutually supporting each other and replicating together. The dividing line between a meme and a meme complex is yet again diffuse. In this text we will not try to distinguish between the two.

## 1.2. Issues in Memetics

### 1.2.1. The Meme - Gene Analogy

Much fuss has been made over this analogy since it was introduced by Dawkins in his "The Selfish Gene". When Dawkins introduced this analogy it was to give us a meaningful comparison in the light of which we would better understand the concept of memes. This was done to help our *initial phase* of understanding; unfortunately, many memeticists has not left this area. Many writers have scrutinized the comparison with the gene to see if it really is analogous or not, i.e. Hans-Cees Speel (1996). Although this has provided interesting reading we feel that it is a bit beside the point. The importance of memes lies not in whether they are mental copies of the genes and obey the same laws as the genes do or don't, but rather in how they work and what they are capable of (and not capable of). We do not feel that you can reach a complete understanding of this only by comparing them to other things. You have to study the idea of memes *in itself*.

This is why we have chosen to make a detailed study of what we can call "the memetic life-cycle", to try to discover its inner memetic workings. Our aim is to find a model that, whether analogous to the genetic life-cycle or not, is sound and supported by studies of existing memes.

### 1.2.2. Definitions

In the field of memetics there are a couple of different definitions of "host", "vector" and "meme" around, and there is a tendency to make these wide to the point of being meaningless. We want operational definitions that are usable and still distinct. Therefore, in this paper, we are going to use the words "host" and "vector" as such (meme has already been defined above):

Host: A host must be able to possess at least *the potential capacity* to elaborate on the meme and to perform those cognitive tasks connected to the meme that we normally refer to as "understanding". This means that only humans can be hosts (animals can perhaps become hosts for simpler memes, but we will not discuss this here), at least until the development of artificial intelligences reaches further.

Vector: A vector is anything that transports the meme between hosts without the capacity to reflect on the meme. Examples are a wall, a voice, an email-program, or a picture. Can a human be a vector? Yes she can, if she lacks the cognitive capacity (or interest) to elaborate on a specific meme. Then she is just a non-reflective carrier of the meme, much the same as a book. Note though that the human vector is still a potential host - or *inactive* host (Grant, 1990) - for the meme, should she suddenly choose to analyze the meme (in its widest sense) or achieve the contextual understanding which would make this possible.

### 1.2.3. The Conscious Meme

It is worth noting that although the terminology used in genetics and memetics sometimes seems to indicate that genes and memes act upon their own conscious will, this is of course not the case. Genes and memes are not conscious, and they do not have a will as such to act upon. But it is practical and economical to speak as if they do, since their behavior follows such patterns.

This way of speaking can be seen as lazy shorthand; "a meme wants X" means "the fitness of a meme is enhanced by X".

## 2. The Lifecycle of Memes

Memes have a life-cycle similar to parasites (Fig 1). During the transmission phase of the meme it is encoded in a vector, such as a spoken message, text, image, email, observed behavior or slab of stone. When a potential host decodes the meme (reads the text, hears the message) the meme may become active and infects the person, who becomes a new host (the infection phase). At some point the meme is encoded in a suitable vector (not necessarily the same medium it was originally decoded from) and can be spread to infect new hosts.

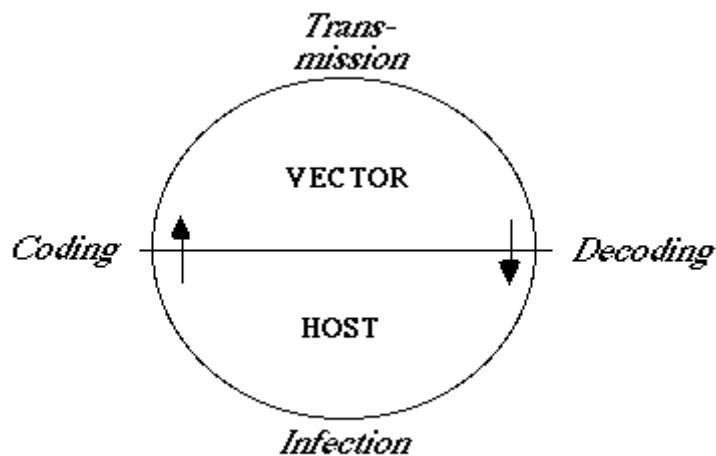


Fig 1. A simple model of the lifecycle of a meme.

This division of the lifecycle makes it easier to discuss memetic selection criteria, such as the list proposed by Heylighen (1994):

- Contribution to individual fitness (how the meme enhances the fitness of its host)
- Reliability of predictions
- Learnability
- Ease of communication
- Tendency to be transmitted
- Conformity pressure ("meme selfishness"; how the meme interacts with other memes in the host)
- Collective fitness (how the meme enhances the fitness of the group or social system of its hosts).

or other, more elaborate divisions such as (Hale-Evans 1995).

In the following we will discuss the intrinsic factors of the meme that contributes to its fitness, and those external factors that mesh with them. We will look at the factors that help or hinder a meme in each of the phases of its lifecycle as depicted in Figure 1. These factors will be summarized below in an extended version of the model (Figure 2).

A successful meme will be good at exploiting these factors in its environment, while memes that cannot exploit them well will be out-competed and eventually go extinct. If one phase presents an insurmountable obstacle to the meme, it will be unable to reproduce and survive. This allows us to make estimates of memetic viability.

## 2.1. Transmission Phase

In the transmission phase the meme is encoded in a vector, some kind of information- carrying medium. Which medium is used strongly depends on the meme, both how it *can* be expressed (influenced by its complexity, the need for copying-fidelity and the requirements of its semantic form) and how it "*wants*" to be expressed. Often the medium is strongly linked with the meme or an actual part of it, as in the case of the "Kilroy was here" meme where part of the meme is the graffiti itself, suggesting the possibility of scrawling it on some suitable surface; "The medium is the message" as Marshall McLuhan put it. Memes are able to shift between media, sometimes with a mutating effect on the meme (such as the major differences between the *Hunchback of Notre Dame* by Victor Hugo and the same story told by Disney).

It is interesting to see how memes and media have co-evolved: many media have been developed as memetic vectors, able to encode memes indefinitely and with a high degree of exactness, while memes have evolved to use them and exploit their peculiarities.

Dawkins (1976) suggests three qualities of a meme that gives it a high survival value: longevity, fecundity and copying fidelity. Longevity and copying-fidelity are most significant in the transmission phase (although longevity can also be seen as the meme's ability to remain in memory for a long time), and are strongly linked with the properties of the medium. Heylighen

(1994) also points out that ease of communication increases memetic fitness, either through creating salient behaviour that is easy to imitate or by being able to be clearly expressed.

### **2.1.1. Reproduction Ability**

The more copies of itself a meme can encode in vectors, the higher fitness does it have (by definition). This implies that memes would thrive in media where it is easy to make numerous copies and distribute them widely. From a memetic standpoint, the ideal medium is a broadcast medium where many copies can be made cheaply. If it is one-to-many such as radio, then the memes of the broadcaster will spread with little feedback; this supports a less diverse memetic ecology than a many-to-many network, but also promotes the memes of the sender more strongly. In a many-to-many medium such as the Internet there are multiple meme sources, and the memetic diversity is high.

Two typical examples of the influence of reproduction ability are "xeroxlore" ("You don't have to be insane to work here, but it helps") that can be found in almost every modern office thanks to photocopiers, and Internet spams ("Make Money Fast!") that thrive in the broadcast environment of the Internet; neither meme would be possible without easy copying.

### **2.1.2. Copying Fidelity**

While mutation leads to evolution, it also risks to destroy or degrade the meme. Just as in biological organisms a balance has to exist between evolvability and copying fidelity (Kelly 1994, p. 537). Many memes have properties that enhance their copying fidelity, often by making errors in encoding noticeable (note the analogy to error-detecting and -correcting codes cf. (Biggs 1985)).

#### **2.1.2.a. Non-Alteration Policy**

Some memes explicitly forbid any alteration of themselves, such as the instructions given in chain letters, the copyright notices of public domain information (only unaltered copying is allowed) and many sacred formulae that warn of awful consequences if they should be altered. They increase their copying fidelity by making the host careful in encoding them well.

#### **2.1.2.b. Structure**

Cohesiveness may also improve copying fidelity by introducing patterns that support the copying and transmission process. Poetry becomes stable by using rhyme and meter, since mutations create obvious errors that can be corrected. Many memes contain noteworthy features such as humor, commonly known symbols or repetition (cf. the common use of groups of three in myths). Incidentally, these properties also make memory encoding and retrieval easier.

#### **2.1.2.c. Simplicity**

Another way of achieving high copying fidelity and a low mutation rate is simplicity. A short meme is less likely to be changed, and may even be "brittle" - any change makes it obviously

unusable to the encoder, and thus prevents mutation. Longer or larger memes can survive only in media with high copying fidelity and low copying costs.

#### **2.1.2.d. Repetition**

A fourth way of achieving a low mutation rate in the vector form is to concentrate on making potential hosts understand the meme, not to ensure perfect copying fidelity (which may not be possible in some media, such as the spoken word). By repeated exposure, a potential host will not only recreate the meme from possibly distorted encodings, but also be more likely to become infected. This is an especially useful strategy for meme complexes, large memes and memes with high abstractability since they can (and often must) be transmitted piecemeal.

#### **2.1.3. Survival**

The meme has to be able to survive in the medium, and the medium itself has to survive. Memes that can be encoded in durable vectors such as books, great art or major myths can spread almost unchanged for millennia, while memes encoded in ephemeral vectors such as the spoken word have to spread from host to host quickly and are also more likely to mutate into new variants.

#### **2.1.4. Abstractability**

Some memes can be encoded in a large variety of vectors, while others are fixed in a single medium. Canonical examples are funny stories, which can be spoken, acted, written or drawn, and graffiti, which is its own vector.

A meme that can pass from one medium to another can spread more easily and change into forms more able to infect new hosts, but is of course more sensitive to mutations. One strategy that works well together with high abstractability is the understanding- repeated exposure strategy, since it decreases the mutation rate and partially relies on the abstractability of the meme.

#### **2.1.5. Decodability**

Finally, a meme must ensure that its vector is possible to decode for new potential hosts. Most artificial media used for information and meme storage are intended to be easy to decode (Dawkins would say there is a memetic selection effect here: the memes of creating media that have low decodability would not be replicated as much as the memes of creating high-decodability media, which would tend to dominate), but many natural media such as behavior are not necessarily easy to decode. Observed behaviour can be viewed as a medium of meme transmission; this is essentially social learning theory (Bandura 1977) restated in a memetic framework. Many behaviors (e.g. gestures such as applauding) are learned this way with no verbal explanation even at an early age and reinforced through positive feedback.

A classic example of behavior transmission (although somewhat outside our strict definition of memes) is the spread of food-washing observed among Japanese monkeys: in 1952 monkeys on the island Koshima were given sweet potatoes by researchers. The potatoes were left on the

beach, and while the monkeys enjoyed the food they disliked the sand. One monkey, 18-month Ima, found that she could get rid of the sand by washing the potatoes in a stream or the sea. Other monkeys observed her behavior and repeated it as they found that it had positive results, until after a few years practically all monkeys, except the oldest, washed their food [2].

It should be noted that some memes increase their fitness by making decoding *harder*. Belonging to an exclusive group increases self-esteem; memes that can only be decoded by some hosts will thus provide them with a feeling of superiority (whose strength depends on the exclusiveness of the meme and other factors) and thus gain a certain infective advantage. The popularity of secret languages among children and secret orders among the upper classes demonstrate this.

A hard-to-decode meme requires more mental activity to decode, which makes infection more likely if the salient aspects of the meme can be reasonably sure to be decoded in the process, and suggests (through cognitive dissonance) a higher value of the meme. If the meme can also be interpreted in several or arbitrary ways, it is also more likely that potential hosts settle for interpretations that fit their attitudes the best. An example of this class of meme is alchemical writings, which are heavily cloaked in symbolism and riddles: the bait is the promise of powerful knowledge (which doesn't have to be delivered if the riddles are hard enough), the meme itself is the symbolical language of alchemy and the hook that leads to transmission is the self-enhancing feeling of spreading esoteric wisdom to the select few who can understand it.

## 2.2. Decoding

Decoding is when a potential host interprets and restructures the information pattern of the meme in the vector by perceiving and understanding, and thus creates a mental copy of the meme. In other words, if a meme is to infect a host, it must first be perceived by the host, then decoded to fit the host's schemata. The possibility of infection arises after decoding. Many of these decoding processes are automatic and/or subconscious.

### 2.2.1. Understanding and Interest

If the host is incapable of understanding the meme or even incapable of perceiving it in the form in which it is being transmitted, the meme will not proceed to the infection phase. Memes that in some way promote their perception, by being encoded in a noticeable vector, by containing strong emotional content or otherwise arouse interest will have an increased likelihood of infecting the host.

Understanding is in this case not limited to conscious understanding, as in the case of copying a popular fashion, neither is it always necessary that the host acquires a complete and correct understanding of the meme (see the "Kilroy was here" example) if it can later be fully decoded using the repetition strategy.

### 2.2.2. Pro- and Contra-memes

There is a possibility that the potential host is infected by memes with the function of counteracting or helping the current meme. Such memes are often symbiotic parts of larger

meme-complexes, with the function of indirectly promoting itself through the promotion of its meme-complex. A typical example is religions, which promotes trust in the teachings of the religious authorities (pro-memes) and often contain memes denouncing competing religions as heresy (contra-memes).

Contra-memes act by making their hosts automatically reject memes that do not fit the dominant cognitive structures. Often they act by creating a strong emotional response or attributing negative traits to the meme (knee-jerk reactions). In the same way pro-memes create positive attributions of certain memes and ease their assimilation.

### **2.3. Infection Phase**

After successful decoding the meme becomes part of the host's mental structures, and this is called infection. A person who does not remember a meme at all is not infected. A person that does remember a meme but whose behavior is not affected has thus become a human vector. A person whose behavior is affected by a meme has been actively infected and can potentially transmit it to other hosts.

The following factors will influence whether an *active* infection will occur or not.

#### **2.3.1. Ability to Fit Cognitive Structures**

The meme must fit into the present schemata of the host to be seriously considered. A meme that assumes that God exists will probably not be successful in the mind of an atheist. To convince this person, an extensive set of pro-memes would have to be successful in infecting the atheist, thus reconfiguring the cognitive structures of its host to that of a religious person.

#### **2.3.2. Threat/Bait**

Most memes use threats or temptations to make the host accept the meme. If the meme appears to provide an advantage over previous memes, it can replace them (but it should be noted that older memes may have infiltrated the motivation structure of the host, making her unwilling to switch views anyway). Threats are even more potent, since potential risks are evaluated more strongly than potential gains (Kahneman and Tversky 1982), which means that a threatening meme often can out-compete a tempting meme.

Often threats and baits are combined to further enhance the meme. Many religions use this in propagating themselves, promising the faithless a hot afterlife (threatening) and the faithful a fluffy and light one (tempting).

#### **2.3.3. Storage**

If a meme is to be spread by a host for a long time, the host must remember the meme. If a host is infected and later forgets the meme and/or stops acting out the new behaviour before the host has spread the meme on, the host has not done the meme any more good than if the host had not been infected in the first place. Thus successful memes encourage permanent or long-lasting

changes in the host. Note that it is not necessary for the hosts to remember the meme itself, just change their behaviours in a way that will promote the spread of the (reconstructed) meme. This can be achieved in several ways:

### **2.3.3.a. Assimilation/Accommodation**

If the meme can be assimilated into existing schemata, it will be supported by them. If the meme can somehow force an accommodation, it will have even better chances of becoming a vital part of a long-lasting mental structure (and this also gives it an excellent position from which to act as a contra- or pro-meme).

### **2.3.3.b. Elaboration**

Something that a host actively thinks about is less likely to be forgotten, and also more likely to influence behaviour. Thus memes that encourage thinking or fantasising about themselves or related concepts have increased chances of survival. Rituals and ceremony are often powerful reminders of the meme.

### **2.3.3.c. External storage**

Since human memory tends to be rather uncertain, external memory aids can also aid memes greatly not just as vectors, but as memory feedback.

If a host is infected by a scientific meme-complex he will be encouraged to read books relating to the meme complex. The host becomes likely to learn more and more about the theories rather than forgetting parts of them, and should he forget something relevant he can look it up again; the books can serve as memory feedback loops and also act as vectors for other parts of the meme-complex causing further infection.

## **2.3.4. Storage Time**

The longer a meme infects its host, the more chances it has to be spread on to other potential hosts, but there is also an increased risk of mutation. Some memes have limited lifetimes determined by outside factors, such as millennial memes ("The world will be destroyed on Tuesday!") whose fitness decrease significantly after a set date (but before that date gain in fitness by being actual and urgent). There are examples of computer viruses that destroy themselves after a predefined goal, like making X copies of itself or residing for X days in the computer (Dawkins 1993), which are surprisingly similar to the instructions of chain letters – after making the copies and sending them, the host is no longer instructed to spread the meme.

## **2.3.5. Survival**

Survival of the meme depends upon whether or not the meme can handle or use the following problems in the host mind:

### **2.3.5.a. External Contra-Memes**

An infection may be subject to contra-memes, trying to "cure" the infection after it has become active, as opposed to a contra-memes struggle to prevent an infection in the decoding phase. A contra-meme does not have to be specifically directed towards its counterpart, it only has to counteract the intention of its counterpart. Survival requires a meme to withstand their attacks. Many obviously false but firmly believed memes can be surprisingly tenacious in the face of opposing evidence (Gross 1996).

### **2.3.5.b. Immunity**

If a meme can make the host assimilate or accommodate contra-schemata to protect itself, it is less likely to have to deal with external contra-memes or competitors. The meme complex of atheism is likely to give its host a defense against any religious meme complex. This defense is practically an immunity since a convinced atheist won't even consider listening to a religious host with an open attitude. The opposite seems to apply too.

### **2.3.5.c. Universality**

If the meme is widely enough defined, or is a generally accepted concept, it can assimilate all other memes because it is more general and can be used as a context which can explain all necessary functions. Many religions claim to explain the world's creation and other phenomena better than modern science.

## **2.4. Encoding and Spread**

The meme must replicate itself if it wants to be successful, that is, to spread to and infect more hosts. Quantity is important for some memes in order to survive, especially weak memes which do not occupy the ideosphere (the universe of ideas, an analogy of the biosphere (Monod 1970)) of their host for long since they are easily forgotten. As a reaction they try to infect as many hosts in as little time as possible. A typical example of this is chain-letters.

Other memes have a more targeted area of hosts. The target hosts are the only ones necessary for the survival of the meme, but they are on the other hand crucial. This kind of meme is often evolved to survive better in a single host over time than the quality-oriented one, since the spreading of this second meme type often involves a more complex and time-consuming infection process (as in religion).

### **2.4.1. Hooks and motivation**

A host doesn't start to spread a particular meme all by herself. She has to be motivated. To this extent, memes have a particular trait, or a co-meme, called a *hook*. The hook is what encourages the host to spread the meme. A common hook relies on humans' altruism. It works like this (after Hofstadter 1985, p. 55):

ALTRUISTIC PREMISE: *I (the host) don't want any harm to befall my friends.*

MEME HOOK: "Anyone who doesn't believe in this doctrine will be tormented in the afterlife."

CONCLUSION: *Since the doctrine is true, and since the premise is true, I will make sure that my friends starts to believe in the doctrine.*

Another type of hook is of course the opposite, the threat directed towards the host. This is an efficient tool not only for spreading, but works also to minimize mutation ("You will be tormented if you misread the doctrine") and to guarantee a firm place in the host's ideosphere ("You will be tormented if you lose faith").

Naturally, all hooks aren't as elaborate as the ones above. These types are normally found in the larger meme complexes. But singular memes are equipped with the hook co-meme too. Take, for instance, the joke. Why do you tell a joke? Maybe because you want to make people happy, and/or you want to be appreciated or popular. The hook for telling - and spreading - a joke would then be something like this:

MEME HOOK: "If you tell people something funny they will be happy and they will appreciate you."

Hopefully, this is enough to motivate you to start spreading the meme. A hook should obviously not contradict common sense and in a meme complex it should ultimately feel like a natural extension of the meme complex. This is to make sure the hook is easily, and preferably automatically, activated. A complicated hook that demands a lot of conscious elaboration in order to be activated has doomed the meme by inhibiting the possibilities of swift reproduction.

Note that the above-mentioned meme hook is not married to a particular joke. Rather, it is implicit in the host's disposition towards jokes in general. Whereas the hook for, say, a chain letter is explicit and tells you what to with this particular letter and this letter only.

#### 2.4.2. Feedback

Another important determinant in how successful the spreading of a particular meme will be is the feedback. This includes direct as well as indirect feedback.

Direct feedback is the immediate response you get when trying to spread the meme. Did the new host get infected (Did she laugh at your joke)? Was the vector you chose a satisfying medium? Indirect feedback concerns matters such as how often you recognize the meme in different media and whether it seems like a lot of people inhabit the meme (and thereby strengthen your own belief in it). In other words, indirect feedback is your recognition of the spreading meme apart from your own participation in the process. You can naturally recognize some indirect feedback as being partly a result of your own reproduction of the meme, but only in the sense of your meme spreading being part of a larger meme-spreading complex. If your actions are the exclusive producer of the feedback, then the feedback is always direct.

Positive feedback strengthens the belief in the meme and encourages the spread, whether negative feedback works in the opposite direction. Strong enough negative feedback can actually kill the meme in the original host, if the meme is not equipped to handle a situation like that. Successful memplexes are well prepared, e.g. the religious defense "If they don't believe you they work for the devil, and you should avoid further contact." But for smaller memes, negative feedback often proves to be fatal. This is what happens when we say that a joke "wears out", we start to receive negative feedback because everybody is tired of the joke and eventually we stop telling it.

#### 2.4.3. Survival through Vectors

As always, it is important for the meme to mutate as little as possible. But reproduction must also be as easy as possible. An optimal vector should satisfy both these needs. The Torah is an example of minimal mutation but extremely low reproduction rate. The rabbis were forced to copy the original by hand, page by page, letter by letter. If one letter got wrong they had to do the page all over again (Eco 1988, p. 569). Gutenberg changed all that, and today the printed word is one of our most common vectors.

Another element which enhances the meme's possibilities is whether it survives a transition from one media to another. Of course, this flexibility also enhances the risk for mutation. One might think that memes lie dormant while in vectors, but this is not necessarily so. With the advent of television and more interactive types of media, not to mention the Internet, the meme have got the possibility to shape its vectors to the extent that the media becomes the message, to paraphrase McLuhan. Examples of this are the various cyberchurches which exist only on the Internet and preaches technosophy, which aims to foster a spiritual appreciation for technology (Wright, 1996).

### 2.5. Summary of Model

In order to stay successful and reproductive, the meme has to complete the described cycle over and over again - preferably with as little mutation as possible (unless mutation is one of its particular characteristics, as in the example of urban legends, below). The meme will die if it is unable to complete the cycle.

Our ambition in writing this paper has been to analyze the obstacles and ordeals which the meme has to face in its life-cycle, and to get an understanding of its workings. The result is summarized in fig. 2.

To show how this cycle applies in real life, we will now end this study by analyzing three different well-known memes or meme complexes. While reading the examples, use fig. 2 as a tool to see the different phases.

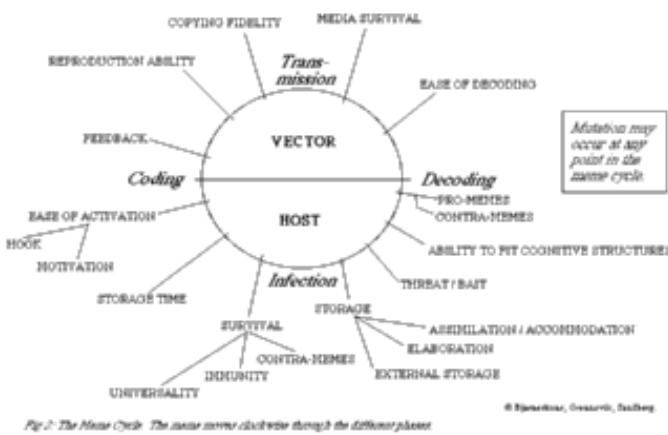


Fig. 2: The Meme Cycle. The meme moves clockwise through the different phases.

### 3. Three Illuminating Examples

#### 3.1. The "Kilroy Was Here" Meme

This meme originated during the Second World War, when wharf inspector James J. Kilroy of Quincy, Massachusetts used the slogan "Kilroy was here" to mark products he had tested and approved. The marked products appeared on many battlefields, and the signature that seemed to appear just about everywhere caught the imagination of many soldiers, who began to copy it on just about any writable surface (Funk 1950). Most likely others were intrigued by the slogan that appeared in unlikely places, so they copied it further to spread the myth.

While the meme spread well for several decades, it eventually went all but extinct in its active form. There seems to be several reasons for this:

- Competition from other forms of graffiti, with stronger ties to subcultures.
- The ageing of the most highly infected population. Since the tendency to scrawl graffiti is highly age-dependent, it seems to be likely that as the original cohort aged, they did not reproduce the meme as often as before, and the original context was gradually lost.
- A lack of novelty. Practically all people have been exposed to the meme, but part of its appeal was the surprise effect of a well placed "Kilroy was here" scrawl; once it has been seen and understood enough, the novelty wears off. A bit paradoxically, the meta-meme of knowing about the Kilroy meme inhibits the further spread of the Kilroy meme, which makes it in fact a contra-meme to the Kilroy meme. This is why today, such a large body of people have knowledge about the meme, i.e. the meta-memetal level, without actually being infected by the original meme (otherwise it would still appear on the city walls).

### **3.1.1. Properties of the Meme**

"Kilroy was here" is extremely well suited for the transmission phase, where it is encoded in a graffiti vector.

It is very easy to reproduce, and due to its brevity the copying fidelity can be very high. Its decodability is also high, since after the Second World War English became a lingua franca over a large part of the world and acquired a certain status. The meme was spread by English-speaking hosts, and would thus tend to end up in areas where English was understood at least by a part of the population.

The survival of graffiti is highly variable, but by its nature it is semi-permanent and intended to be highly visible, which ensures that more potential hosts notices it.

It is uncertain how well "Kilroy was here" can be abstracted. In its original form, the graffiti vector was an integral part of the meme and crucial to hint at that it should be reproduced. Later variants appeared, such as a cartoon figure and stickers, but they do not appear to have been as fertile, mostly because they were harder to copy.

The meme's intimate connection to its vector, e.g. walls, made it poorly fit to survive in other media. Also, the meme was very sensitive for mutations. It was enough that you changed one of its smallest parts, a letter, to seriously damage the meme.

It is of great help to understand this phrase, that is to know what the English words mean together. The problem of decoding the sentence is quite an easy one, but it is harder to decode what it *really* means. This is probably one of its strengths. One can find one's own explanation of its meaning. Furthermore it is small and simple. During a time when the meme is popular, the host also gets multiple chances to try to decode it.

Since the meme is without obvious meaning it is hard to contradict, so there should be no active defense against the meme. The meaninglessness of it can also invoke wandering thoughts about the meme, and actively elaborating is connected with better remembering.

What motivated people to spread the "Kilroy was here" meme? There was never any direct host-to-host contact in the case of this meme. This meant that no host received a direct positive feedback, which is a powerful reproduction booster. And there was no obvious hook accompanying the meme.

This is one of the great meme mysteries. Perhaps that was enough motivation to spread the meme, to become part of the mystery - and also, in the beginning, to share the joke of whom this much-talked-about Kilroy character was. Thus, the host created a bond with a community of Kilroy writers, most of which she would never meet, but could still belong to. A feeling of belonging may have served as a hook to motivate the conscious spread of the meme.

The Kilroy writers only way of confirming that there were others was the indirect feedback, but this is also the point. The Kilroy writers became invisible, even before each other, so that the

meme seemed to live its own life mysteriously reproducing on the walls and the writers themselves could feel as privileged members of a mysterious brotherhood.

### **3.2. Urban Legends**

Urban legends have been described (alt.folklore.urban FAQ) as stories that:

- Appears mysteriously and spreads spontaneously in various forms.
- Contains elements of humor or horror (the horror often "punishes" someone who flouts society's conventions)
- Makes good storytelling.

The first property suggests that they are memes, able to mutate and spreading with no link to the original creator (although some urban legends attribute the story to some proper authority to gain some measure of credibility; cf. Gross (1996)). The second property in part together with the third explains why they are replicated: they fill a psychological need for entertainment, emotion, reinforcement of attitudes and attention for the storyteller.

There has been much discussion about the links between traditional storytelling, urban legends, memorats [3] and rumors. Generally urban legends are apparently realistic stories but actually have a stylized content, with a simple plot which is often very visual and easily remembered and told (af Klintberg 1978, pp. 154). These contrasts to rumors, which are short (often just a simple statement with additional information) and lack epic structure. Both urban legends and rumors can be viewed as ways of spreading information in situations without official information and by releasing the tension of social uncertainty (Mullen 1972).

It is likely that rumors may evolve into urban legends. The classic study *The Psychology of Rumor* by Allport and Postman (1947) suggests that a rumor will become more stylized during spreading. This may be partially the result of the experimental set-up, which is based on unilinear spread; in collective spread variations tends to appear (Peterson and Gist 1951). These results are predicted by memetic theory: in a unilinear spreading situation only transmission ability will be relevant (since the number of hosts are small), while in a collective situation the increased number of transmissions will also lead to mutation and variation. Once a rumor or memorat taps into a good epic form due to a mutation or a deliberate change its spread will be highly enhanced, and it becomes an urban myth which will spread fast.

#### **3.2.1. Properties of the Meme**

Urban legends are usually transmitted as oral tradition; different legends thrive in different social groups. The mutation frequency is rather high due to the oral spread, but this is counteracted by comparatively simple and strong storylines which can be elaborated in various ways, often based on strongly interesting subjects such as sex, death, the supernatural and embarrassment. The subjects all tend to promote listener interest and hence replication.

To be accepted as anything other than a joke or pure horror story the legends need some measure of plausibility; often this is provided by referring to apparently real people and institutions

(Gross 1996), the story shows that the people involved are "normal" people or a real "friend of a friend" (thus extending the storyteller's credibility indirectly to the presumed source).

Another reason to accept or remember the story is by hearing it from several independent sources; this appears to confirm its veracity, and minor inconsistencies can be explained away as being errors in re-telling.

Memetic theory predicts that legends that fit in well with the social schemata and attitudes of their hosts will have a higher fitness than legends that do not conform, and this seems to be supported by the changes that occur in urban legends over time. Some urban legends have survived for many decades, changing to fit in with changes in popular attitude or society. The myths about people being drugged in subways originally involved white slave trade, but today warn of dealers seeking to make more people dependent on drugs.

Bengt af Klintberg points out (1978, pp. 153) that it is possible to partially classify urban myths by the way they spread. Many spread in the characteristic way of rumors: an exponential spread until saturation followed by a die-back (sometimes caused by an official denial). The same rumor or urban myth can recur in other places with a similar way of spreading; in a surprising number of cases newspapers acts as vectors. Note the similarity to epidemics. These urban myths often deal with things relating to the listener's life, something that he or she could experience. But there are also urban legends that spread in a less explosive manner. Their contents are less likely to be experienced by the listener, and are told more as entertainment than actual events. Typical examples are the horror stories told by teenagers (af Klintberg 1978, pp. 181).

To survive memetically the legends do not need to be believed, since they provide other incentives for being told, but if the storyteller believes in the legend its spread and credibility will be enhanced and can be motivated even if it lacks obvious entertainment value. This may explain the difference between the fast- and slow-spreading legends.

Another important factor is that believable legends involve the listeners much more by suggesting they could be the victims of the story; in the case of potential dangers (such as the "rat in the pizza" stories) there is an incentive of remembering the danger and passing on the knowledge.

### 3.3. Christianity

Religions represent some of the most powerful and elaborate meme complexes in existence today; they have evolved over millennia into countless variants and co-evolved with cultures. Making a complete memetic analysis of even a single religion is beyond the scope of this paper, so by necessity the following discussion will be rather general.

Religions tend to consist of some basic core memes (in the case of Christianity the belief in God and salvation through Christ) surrounded by symbiotic doctrinal memes (how salvation can be reached, ethical systems, the cosmology) and then an immense cloud of related memes (religious stories, doctrines, interpretations). These memes form a symbiotic whole; the core memes need symbiotic memes to provide hooks and baits, and the symbiotic memes reinforce each other and are given legitimacy by the core memes.

### **3.3.1. Properties of the Meme**

The Christianity meme complex has throughout history been transmitted in a multitude of forms: as oral stories, through books and art, through example and through upbringing. Due to its complexity the transmission takes time and is closely linked to cultural understanding. This either requires a relatively concentrated effort to transmit the complex (mission) or to spread it by cultural diffusion and imitation (upbringing). A frequent diffusion situation is when a child is brought up in a Christian home. The Christian meme complex is presented as the truth about how the world functions. Variants of the meme have increased their fitness by encouraging a high rate of reproduction and cultural transmission (Lynch 1996).

Most major religions rely on active transmission: one or more hosts actively supports the spread of the meme, often in an interactive and deliberate way. Efficient methods for mission have co-evolved with the religion and the situation; the best missionaries gained the most converts, among which were the next generation of missionaries (and missionary teachers) who would learn and spread some of their best methods.

Classically, Christianity have used the bait of salvation (freedom from fear, personal happiness and prosperity, spiritual fulfillment, eternal life or union with God have all been promoted at various times) combined with the threat of damnation to promote interest and infection. This is however just the explicit bait, it appears likely that many Christian movements have been spread by implicit factors such as a sense of belonging, social conformity and a consistent world-view. It is worth noting that the baits and threats are mostly based on the symbiotic memes and not the core memes of the complex, which means their relative prevalence can change to fit the situation (for example the ratio of hellfire threats to salvation baits used in sermons) or they can evolve while leaving the core memes unchanged.

Religions are often better than other meme complexes (such as science) at explaining how the world works on an emotional level. They provide answers to existential questions that are emotionally appealing, creating a satisfying world model (which then becomes intellectually satisfying regardless of its consistency due to cognitive dissonance). Because religions seldom try to empirically prove themselves they cannot be disproved, which further aids their stability. A religion can spread regardless of the truth or falsity of its claims.

The Christianity meme contains an entire world-view, and seeks to cause an accommodation in the schemata of the infected host; no other memes are allowed to influence high-level planning and behavior ("Ye cannot serve God and Mammon", Matthew 24). This is achieved by rejecting such memes or impulses as 'against God's will', 'sinful' or 'satanic'.

Like all the other major world religions, Christianity has a strong mission. It both exists as an explicit missionary order and in the form of an implicit altruistic hook (see the section about hooks and motivation). Christians are urged to set good examples to others, which also increases the likelihood of transmission through social learning.

## 4. Conclusion

The memetic approach is a tool that helps us to understand certain aspects of human behavior. As with all tools, it is not necessarily the best solution at all times. That is why we equip ourselves with a lot of different psychological analytical instruments, so we will be able to choose an efficient approach for each different setting.

Critics of memetics complain about the danger of transforming everything into memes and memetics, feeling that it somehow reduces the importance of the human mind and places focus elsewhere. While some memeticists tend to go overboard with explaining everything in terms of memes, the same could be said about researchers in the fields of psychoanalysis, cognitive science or sociobiology. The importance lies in realising that psychology is full of more or less fit different explanatory perspectives, and memetics is just one of them. But by memetics you can often explain very complex cognitive structures and/or social psychological phenomenon (like Christianity) in a very general, no-nonsense way without having to entangle yourself in a web of unnecessarily complex theories.

Before our paper ends, we will inform you of our until now secret sub-goal with this paper. It is our intention that by now you, by reading this text, have been infected with one of the strongest memes on the planet: The Meta-Meme, e.g. the meme about the theory of memes. It is our sincere hope that you will tell your friends about this (yes, transmission and further infection) or maybe even let them read this paper. In either case, unless you carry a very strong vaccine (see appendix), we have made you a host. And you didn't even flinch. You should be lucky we are not after your money...

## Footnotes

It should be noted that the idea of self-replicating ideas and their evolutionary struggles had come up before in the writings of R. Sperry (1965) and J. Monod (1970).

It is interesting to note that the real story eventually became a basis for a common urban legend, the "100th monkey effect": once 100 monkeys knew how to wash their food the ability suddenly became widespread among the monkeys, even outside Koshima. This claim was originally made in the 1979 book *Lifetide* by Lyall Watson where the author made up the story; since then it has become common myth in the peace and environmental movements (Amundson 1987). Compare this to the section about the evolution of urban legends.

Stories relating real personal experiences, which are interpreted through the collective tradition but lack the fixed intrigue of legends. The term was introduced during the 1930's by C.W. von Sydow. (af Klintberg 1978, pp. 152)

## References

Amundson, R. (1987) Watson and the Hundredth Monkey Phenomenon. *Sceptical Inquirer* Spring 1987, pp. 303-304.

- Allport, G. W., Postman L. J. (1947) *The Psychology of Rumour*  
 alt.folklore.urban FAQ <http://www.urbanlegends.com/afu.faq/>
- Bandura, A. (1977) *Social Learning Theory* Prentice-Hall, Englewood Cliffs, NJ.
- Biggs, N.L. (1985) *Discrete Mathematics*, Oxford: Oxford Science Publications
- Dawkins R, (1976) *The Selfish Gene*, Oxford: Oxford University Press
- Dawkins, R. (1993) Viruses of the Mind. *Free Inquiry*, summer 1993, vol. 13 nr 3  
<http://physserv1.physics.wisc.edu/~shalizi/Dawkins/viruses-of-the-mind.html>
- Dennet, D. (1990) Memes and the Exploitation of Imagination, *Journal of Aesthetics and Art Criticism*, 48, 127-35, Spring 1990 <http://www.tufts.edu/as/cogstud/papers/memeimag.htm>
- Eco, U. (1988) *Focaults Pendel*. Stockholm: Brombergs.
- Funk, W. (1950) *Word Origins and Their Romantic Stories*, Funk & Wagnall Publishing Co.
- Grant, G. (1990) Memetic Lexicon  
[http://www.aleph.se/Trans/Cultural/Memetics/meme\\_lex.html](http://www.aleph.se/Trans/Cultural/Memetics/meme_lex.html)
- Gross, D. (1996) The Blue Star Meme: Applying Natural Selection Thinking to Urban Legends  
<http://www.lycaeum.org/drugs/other/tattoo/meme.html>
- Hale-Evans, R. (1995) Memetics: A Systems Metabiology  
<http://www.apocalypse.org/pub/u/rwhe/memetics.html>
- Heylighen, F. (1994) Memetic Selection Criteria <http://pespmc1.vub.ac.be/MEMSELC.html>
- Hofstadter, D. R. (1985). *Metamagical Themas: Questing for the Essence of Mind and Pattern*. New York: Basic Books.
- Kahneman, D., Tversky, A., eds., (1982) *Judgment under uncertainty: Heuristics and biases*. Cambridge Univ. Press, NY.
- af Klintberg, B. (1982) *Harens Klagan: Studier i gammal och ny folktron*. Andra upplagan. Stockholm: Norstedt & Söner.
- Kelly, K. 1994 *Out of Control: the new Biology of Machines*, London: Fourth Estate Limited.
- Lynch, A. (1996) *Thought Contagion: How Belief Spreads Through Society: The New Science of Memes*. New York: Basic Books. <http://www.mcs.net/~aaron/thoughtcontagion.html>
- Monod, J. 1970 *Chance and Necessity*

Mullen, P. B. (1972) Modern Legend and Rumor Theory. *Journal of the Folklore Institute* 9:2-3, 95-109 1972.

Peterson, W., Gist, N. P. (1951) Rumor and Public Opinion. *American Journal of Sociology* 57, 159-167 1951.

Speel, H-C. (1996) Memetics: On a conceptual framework for cultural evolution. In: Heylighen, F. and Aerts, D. (ed.) *The Evolution of Complexity*. Kluwer, Dordrecht, in press.  
[http://www.sepa.tudelft.nl/~afd\\_ba/hcesmem.html](http://www.sepa.tudelft.nl/~afd_ba/hcesmem.html)

Sperry R. (1965) Mind, brain, and Humanist Values. In *New Views of the Nature of Man*, edited by John R. Platt. Chicago: University of Chicago Press.

Wright, R. (1996) Can Thor Make a Comeback? *Time* vol. 148 No. 25 December 16 1996

## **MEMES, MINDS, AND SELVES**

**By**

**Susan Blackmore**

LSE, Thursday November 28th 1996  
Seminar in the series ``About Biology''  
Department of Psychology  
University of the West of England  
Bristol BS16 2JP

“Memes”---when you hear that word, or worse still find yourself saying it, do you get an irresistible urge to raise your hands in scare quotes or giggle nervously to hide your embarrassment? If so you are not alone. The very idea of the meme seems to strike fear into even the most hardened evolutionist.

Some reject the meme outright as a ``meaningless metaphor" or empty analogy. However, my intention today is to argue that the idea of memes as independent replicators is a useful and powerful idea---and one that will prove essential to understanding the human mind.

I shall first outline the history of the idea of memes and then present four examples of how a theory of memetics can be used to explain specific phenomena; two of these I shall skip over quickly just to give an idea of the kind of reasoning involved. The others I shall consider in more detail. These are the origins of human altruism, and the size of our brains. I shall conclude with some thoughts on the nature of minds and selves.

In 1976 Dawkins published his best-selling *The Selfish Gene*. This book popularized the growing view in biology that natural selection proceeds not in the interest of the species or of the group, nor even of the individual, but in the interest of the genes. The genes are the true replicators and it is their competition that drives the evolution of biological design---or as he would now put it, *Climbing Mount Improbable*.

It is often allowed to elderly and famous men that they can get away with a few pages of wild speculation on forbidden topics at the very end of their great books. I well remember my tutor warning me off the last chapter of Eccles' book on the physiology of the giant squid axon in which he tackled the naughty subject of consciousness. Well, Dawkins was neither old nor (then) terribly famous, but he devoted his last few pages to the topic of memes and has been much derided for it.

Dawkins, clear and daring as always, suggested that all life everywhere in the universe must evolve by the differential survival of slightly inaccurate self-replicating entities. Furthermore, these replicators automatically band together in groups to create systems, or machines that carry them around and work to favor their continued replication. The gene, he claimed, is not the only replicator on our planet. Staring us in the face, though still drifting clumsily about in its primeval soup of culture, is another replicator---a unit of imitation. He gave it the name ``meme," and as

examples suggested ``tunes, ideas, catch-phrases, clothes fashions, ways of making pots or of building arches."

In just those few pages he laid the foundations for understanding the evolution of memes. He discussed their propagation by jumping from brain to brain, likened them to parasites infecting a host, treated them as living structures, and showed how mutually assisting memes will group together. He laid the basis for his later analysis of religions as co-adapted meme-complexes and argued that once a new replicator arises it will tend to take over and begin a new kind of evolution. Above all he treated them as independent replicators, chastising those of his colleagues who tended always to go back to ``biological advantage" to answer questions about human behavior. Yes, he agreed, we got our brains for biological (genetic) reasons but now we have them a new replicator has been unleashed and it need not be subservient to the old.

The meme meme has done fairly well. The word has even been considered for inclusion in the Oxford English Dictionary. However, the basic idea does not seem to have been either understood or much used and its two greatest proponents, Dawkins and Dennett, have both seemed to draw back from the idea after their initial enthusiasm.

Many writers prefer to avoid the term altogether. ``Meme" does not even appear in the index of many of the best recent books about human origins and language such as Pinker's *The Language Instinct*, Dunbar's *Grooming, Gossip and the Evolution of Language*, Mithen's *Prehistory of the Mind*, or Tudge's *The Day before Yesterday*, nor in those about human morality such as Ridley's *The Origins of Virtue*, or Wright's *The Moral Animal*.

There have been many attempts to develop theories of the origins of culture. Perhaps these use the same idea but call it something else. Such theories might roughly be seen as falling along a continuum from the outright rejection of cultural evolution, through most of sociobiology and Wilson's image of the genes holding culture on a leash, to schemes that treat cultural evolution as relatively independent. Among the latter, only Durham uses the word ``meme" and it may be no coincidence that he provides good examples of the memes acting against the interest of the genes or forcing the development of the genes in one direction rather than another. However, even he makes inclusive fitness the final arbiter in his theory of coevolution. As far as I can understand them, other authors do not really treat their unit of cultural exchange as an independent replicator. When they say ``adaptive" or ``maladaptive" they mean for the genes. In other words, when it comes to the crunch they always fall back on appeals to biological advantage, just as Dawkins complained that his colleagues did twenty years ago.

Dawkins is clear on this issue when he says ``There is no reason why success in a meme should have any connection whatever with genetic success." I agree.

The most direct support for this view comes from a philosopher, not a biologist. Dan Dennett has used the idea of memes in both his recent books. He is absolutely clear that the meme is a separate replicator and describes its ``replicator power" in terms of the fundamental algorithm of natural selection. He also goes further, describing a person as the entity created when ``A particular sort of animal is properly furnished by---or infested with---memes." His analysis is

thorough and fascinating, but he does not make specific empirical predictions and few seem to have taken up his ideas.

There is a flourishing Internet community communicating about memes, but as with all such groups the quality is mixed. Twenty years on it therefore seems necessary for me to defend the very idea that the meme is a replicator in its own right, as well as to prove its usefulness.

## The Meme as Replicator

It is widely agreed that for evolution to occur three things are needed; variation, heredity or replication, and differential survival of the copies made. There is enormous variety in the behaviors human beings emit, these behaviors are copied, more or less accurately by other human beings, and not all the copies survive. The meme therefore fits perfectly.

Think of tunes, for example. Millions of variants are sung by millions of people. Only a few get passed on and repeated and even fewer make it into the pop charts or the collections of classics. Scientific papers proliferate but only a few get long listings in the citation indexes. Only a few of the disgusting concoctions made in woks actually make it onto the TV shows that tell you how to Wok things and only a few of my brilliant ideas have ever been appreciated by anyone! In other words, competition to get copied is fierce.

Put another way, there needs to be a replicator that makes slightly inaccurate copies of itself in an environment in which not all the copies can survive. Whichever way you look at it, the meme seems to fit. However, there are some cogent objections to the notion of the meme as replicator. I shall consider three.

### 1: Memes are not like Genes

Unfortunately we only have one other well-known replicator with which to compare the meme. I say unfortunately because this tends to make us think that all replicators must be like genes. In fact genes may be just one example of many potential replicators.

So we need not reject the idea of the meme just because it works so differently from the gene. I suggest we should simply bear in mind the similarities and differences and wait to find out how important these are.

In definition genes and memes are comparable. The gene is an instruction for building proteins, stored in a cell and passed on by reproduction. The meme is an instruction for producing behavior, stored in a brain and passed on to other brains by imitation. However, there are many differences in the way the replication takes place. Genes use the cellular machinery to copy themselves rather accurately. Memes have to be copied by using the brain itself as the replicating machinery. One person has to observe another's behavior and work out somehow how to reproduce it. It is, if you like, a kind of reverse engineering. Say you snap your fingers above your head. I can copy this action relatively easily, yet the processes involved in my doing so must be fiendishly complex and we are nowhere near to understanding them. What we do know is that humans are supremely good at imitating each other and do so a great deal from very early

on in their lives. It is this fiendishly complicated process that makes the transmission of memes possible.

This fact also means that memetic transmission is in some sense Lamarckian. That is, I copy the actions you make, not the instructions your brains holds for making those actions. But don't forget that we can easily ignore the vagaries of environmental constraints on behavior. If I trip over when demonstrating my fantastic new combined garden hose and cat scarer, you will not copy my clumsiness when turning the invention on me. If I tell you a brilliant story when I have a sore throat you will not assume a hoarse voice when passing on the gossip to your friends. Indeed language appears remarkably resistant to the vagaries of individual voices and accents and language must be a major way of passing on memes.

Whether we see the process as Lamarckian also depends in part on what we consider the equivalent of the phenotype to be. If we follow some authors in persisting in seeing the organism as the phenotype then obviously the process is Lamarckian. If we follow Dawkins in treating meme products such as words, music, gestures, skills and fashions, as the equivalent of the phenotype, then the process still appears Lamarckian because these are the very things that are copied. However, we might follow his other suggestion that memes are still drifting clumsily about in their primeval soup. We can see the brain as the replicating machinery for behaviors which have not yet created clear phenotypes. In this case the process is not so obviously Lamarckian.

Does it matter that the copying of memes is so sloppy compared with meiosis? A topic for another lecture might be the way in which language, writing and reading, increase the accuracy of transmission and hence speed up the algorithm of memetic evolution. The creation of even more precise computer languages and software packages, not to mention hardware, may be another step in the same direction.

This relates to another important difference, often noted: the speed of replication. If there can be said to be a generation time for a meme, it can be as short as the time it takes to copy an action or a word or two. Most memes will hang around in various brains for weeks, months or years before being passed on, but many spread at fearsome rates. The news that Fergie spent a million pounds on a family holiday must have passed to millions of brains within a few hours of some journalist finding it out (or making it up).

I like to think of the question of speed in terms of Dennett's "Tower of Generate and Test." Darwinian creatures are generated by natural selection but have to die in the process. Skinnerian creatures can learn by trial and error and consequently can live on to repeat, or not, the behavior they generate. Popperian creatures can imagine behaviors and decide whether to do them or not according to an internal generate and test. And finally, Gregorian creatures can pass on their clever tricks to others, so not every individual has to try the ideas out. Each floor of the tower builds on the last and each speeds up still further the process of accumulating clever tricks. Although Dennett does not explicitly say so, the fourth floor of the tower is the home of the memes---and they are generated, tested and accumulated faster than anything that has gone before.

Just how important these differences will prove to be, we shall have to wait and see. In the mean time I suggest we take careful note of them and use extreme care in transferring the use of terms such as genotype and phenotype, allele and generation to the world of memes. We may be able to use them or we may need to invent entirely new terms.

## **2: What is the unit of the meme?**

This question is often asked and some people appear unwilling to proceed with memetics because the unit cannot be specified. However, we must remember that the same question can be asked of genes---indeed, Dawkins spells out in great detail the problems in choosing any particular size of unit. Nature does not specify the size of a gene.

I also suggest that when I have explained my four suggestions, you ask yourself whether it matters what size of unit I am referring to. I have done this often myself and concluded that for many purposes we may proceed with a memetic analysis without specifying the size of the unit.

## **3: We do not know how memes work**

No we do not---though we may speculate in terms of synaptic potentiation or variations in weights in neural networks. The fact is that when a person carries out any behavior there must be some kind of instruction stored in the brain, and when someone else copies and remembers an action they must also create some kind of neural change. We can get a very long way without knowing how it is done---just as Darwinian Theory got a very long way in the many decades before it was known that genes were the basis of heredity.

I don't think any of these objections is worth worrying about. They are excuses not to try---and I want to try. I want to take the meme as a true replicator, selfishly copying itself in the world of human brains, and see what happens.

The basic approach I take is this---*imagine a world full of hosts for memes (ie brains) and far more memes than can possibly find homes*. Now ask, which memes are more likely to find a safe home and get passed on again? It's that simple. In the process I must assiduously avoid lapsing into considering either the interests of the genes or those of the organisms they create. I am concerned here with the selfish replication of memes and memes do not care either about genes or people. Nor do memes have foresight! I must also be careful of short-hand statements like ``memes want x'' ``memes try to do y'' or ``z is a good meme strategy.'' These may be quite legitimate but I must always ensure that they can be translated back into the longer-winded version, such as ``memes that have the effect of producing x are more likely to survive than those that do not.'' I have done this as well as I can and I hope you may enjoy thinking about some of the consequences of this way of thinking. Some of them are startlingly obvious---once you see them.

I will now outline four consequences. I shall treat the first two very briefly, partly as exercises in thinking memically.

## **1: Why can't we stop thinking?**

*Imagine a world full of brains, and far more memes than can possibly find homes. Which memes are more likely to find a safe home and get passed on again?*

Now imagine a meme that encourages its host to keep on mentally rehearsing it, or a tune that is so easy to hum that it goes round and round in your head, or a thought that just compels you to keep thinking it.

Imagine in contrast a meme that buries itself quietly in your memory and is never rehearsed, or a tune that is too unmemorable to go round in your head, or a thought that is too boring to think again.

Which will do better? Other things being equal, the first lot will. Rehearsal aids memory, and you are likely to express (or even sing) the ideas and tunes that fill your waking hours. What is the consequence? The memosphere fills up with catchy tunes, and thinkable thoughts and we all think an awful lot.

The principle here is familiar. In a forest any tree that grows tall gets more light. So genes for growing tall become more common and the forest ends up being as high as the trees can make it. We can apply the same principle again.

## **2: Why do we talk so much?**

*Imagine a world full of brains, and far more memes than can possibly find homes. Which memes are more likely to find a safe home and get passed on again?*

Now imagine any meme that encourages talking. It might be an idea such as ``talking is a great idea'' or ``It's friendly to chat.'' It might be some urgent thought that just compels you to speak it aloud. It might be just something terribly easy to say.

Imagine in contrast any meme that discourages talking, such as the thought ``talking is a waste of time.'' It might be something you dare not voice aloud, or just something very difficult to say.

Which will do better? Put this way the answer is obvious. The first lot will be heard by more people and, other things being equal, simply must stand a better chance of being propagated. What is the consequence of this? The memosphere will fill up with memes that encourage talking and we will all talk an awful lot. An alternative way of looking at it is this---people who talk more will, on average, spread more memes. So any memes which thrive in chatterboxes are likely to spread.

This makes me see conversation in a new light. Is all that talking really founded on the interests of the genes---on biological advantage? Talking takes a lot of energy and we do talk about some daft and pointless things! And what about thinking? Any meditator will tell you that the mind is constantly full of rubbish---thoughts that come and come and come and won't go away. Do these trivial and stupid thoughts have some hidden biological advantage?

I would at least like to offer the suggestion that they do not. That we do all this talking and all this thinking merely because memes that make us do it are good survivors. This sets the stage for a more audacious suggestion.

### **3: Why are we so nice to each other?**

Of course we aren't always nice to each other, but human co-operation and altruism are something of a mystery. Despite the tremendous advances made in understanding kin selection and inclusive fitness, reciprocal altruism and evolutionarily stable strategies, we do seem to do some peculiarly altruistic things.

Everyone can probably think up their own favorite example. Richard Dawkins (1989 p230) calls blood doing ``a genuine case of pure, disinterested altruism.'' I am more impressed by charitable giving to people in faraway countries who probably share as few of our genes as anyone on earth and whom we are unlikely ever to meet. Why do we hand in wallets found in the street, clear away rubbish on picnics, support eco-friendly companies or recycle our bottles? Why do so many people want to be poorly paid nurses and counselors, social workers and psychotherapists? Many people believe all this must ultimately be explained in terms of biological advantage but I shall offer an alternative for consideration. We can use our, by now, familiar tactic.

***Imagine a world full of brains, and far more memes than can possibly find homes. Which memes are more likely to find a safe home and get passed on again?***

Imagine the sort of meme that encourages its host to be friendly and kind. They might be memes for throwing good parties, for being generous with the home-made marmalade, or being prepared to spend time listening to a friend's woes. Now compare this with memes for being unfriendly and mean---never giving dinners or buying drinks, and refusing to spare your time to listen. Which will spread more quickly?

The first type. People like to be with nice people. So those who harbor lots of friendliness memes will spend more time with others and have more chances to spread their memes. It is therefore in the interest of memes to get into nice people. In consequence many of us will end up harboring lots of memes for being nice to others. You may wish to challenge any of the above steps. It is therefore reassuring to learn from many experiments in social psychology, that people are more likely to adopt ideas from those they perceive as friendly or attractive. Whether this is a cause or a consequence of the above argument is debatable. It would be most interesting if psychological facts like this, or others such as cognitive dissonance, or the need for self esteem, could be derived from simple memetic principles---but that is a topic for another time!

For now we should consider whether the idea is testable. It predicts that people should act in ways that benefit the spread of their memes even at some cost to themselves. We are familiar with buying information and with buying a way into people's minds for the purposes of selling products, but this theory predicts that people will pay simply to spread the memes they hold---because the memes force them to. Many aspects of persuasion and conversion to causes may be turn out to involve this mechanism. Altruism may turn out to be yet another of the meme tricks

that religions (those most powerful of meme-complexes) have purloined. Almost all of them thrive on making their members believe they are doing good.

Of course, being generous is expensive. There will always be pressure against it and if memes can find alternative strategies for spreading they will. For example, powerful people may be able to spread memes without being altruistic as well! However, that does not change the basic argument---that altruism is a good way to spread memes. If memes have replicator power we should expect them to press their carriers into service.

You may have noticed that the underlying theme in all of these arguments is that the memes may act in opposition to the interest of the genes. Thinking all the time may not be very expensive in energy terms---the brain is still using energy even when not thinking. However, it must cost something. Talking is certainly expensive, as anyone who has been seriously ill will attest. And of course any altruistic act is, by definition, costly to the actor.

I would say that this is just what we should expect if memes are truly independent replicators. They do not care either about the genes or about the creatures the genes have created. Their only interest is self-propagation. The consequence is that if they can propagate by stealing resources from the genes, they will do so. And of course this is not a recipe for human happiness.

I want to take just one more step along this path. My final example shows the memes forcing the hand of the genes in a much more physical way.

#### **4: Why are our brains so big?**

Yes, I know this is an old chestnut, and there are lots and lots of good answers to the question. But are they good enough? Let us not forget how mysterious this issue really is. Brains are notoriously expensive both to build and to run. They take up about 2% of the body's weight but use about 20% of its energy. Our brains are three times the size of the brains of apes of equivalent body size. Compared to other mammals our encephalisation quotient is even higher, up to about 25. On many measures of brain capacity humans stand out alone. Brains are also dangerous organs to give birth to. The fact that such intelligence has arisen in an animal that stands upright may or may not be a coincidence but it certainly adds to the problem. Our pelvises are not ideally suited for giving birth to huge brains---yet we do it. Why?

The mystery was deepened for me by thinking about the size of the biological advantage required for survival. I was fascinated to read about a study addressing the question of the fate of the Neanderthals. Zubrow used computer simulations to determine the effect of a slight competitive edge and concluded that a 2% advantage could eliminate a competing population in less than a millennium. If we only need such a tiny advantage why do we have such a large one?

Several new answers have recently been proposed. For example, Robin Dunbar argues that we need large brains in order to gossip, and we need to gossip as a kind of verbal grooming to keep very large bands of people together. Christopher Wills argues that the runaway evolution of the human brain results from an increasingly swift gene-environment feedback loop. Miller proposes that our vast brains have been created by sexual selection; and Richerson and Boyd claim they

are used for individual and social learning, favored under increasing rates of environmental variation. What these authors all have in common is that their ultimate appeal is to the genes. Like Dawkins' bewailed colleagues, they always wish to go back to biological advantage. I am going to propose an alternative based on memetic advantage.

Imagine early hominids who, for good biological reasons, gained the ability to imitate each other and to develop simple language. Once this step occurred memes could begin to spread. And also--once this step occurred the genes would no longer be able to stop the spread! Presumably the earliest memes would be useful ones, such as ways of making pots or knives, ways of catching or dismembering prey, and names for people, events and tools. Let us assume that some people would have slightly larger brains and that larger brains are better copiers. As more and more people began to pick up these early memes, the environment would change so that it became more and more necessary to have the skills in order to survive. So these slightly larger brained people would have an advantage. That, I propose, is how we got our big brains.

The process is related to the Baldwin Effect. I like to use Dennett's "Tower of Generate and Test" again here.

On the ground floor are the Darwinian creatures. As these develop they change the environment in which they live, creating new selection pressures that lead to new design improvements. One result is larger brains capable of learning and the arrival of Skinnerian creatures. These again change their own environment, giving an advantage to the quicker learners. One aspect of quicker learning is internalization---thinking before you act. So Popperian creatures are born and again change their environment so that better thinkers are at an advantage. Finally the ability to copy actions appears, giving rise to the Gregorian creatures and the birth of the new replicators---the memes. Creatures of this kind again change their environment so that those most able to adopt the memes are at an advantage.

Although the process is similar to all the previous ones, this last step is a big one. Note, most importantly, that it depends not on learning or on cleverness *per se*, but on the ability to imitate. A second replicator has now appeared that spreads at a fantastic rate and changes the environment as it goes.

An early hominid who was incapable of mastering any of the new techniques of tool making, speaking or hunting would be at some disadvantage, and the importance of this disadvantage would increase as the memes spread. In a population with few available memes, brain size would not be very important, in a population with lots of memes it would. It seems to me that this fundamental change in selection pressures, spreading at the rate of meme propagation, provides for the first time a plausible reason why our brains are totally out of line with all other brains on the planet. They have been meme-driven. One replicator has forced the moves of another.

## Minds, Memes and Selves

We can now see the human mind as the creation of two replicators, one using the machinery created by the other for its replication. As Dennett pointed out, people are animals infested with memes. Our personalities, abilities and unique qualities derive from the complex interplay of

these replicators. What then of our innermost selves---the ``real me," the person who experiences 'my' life?

I would say that selves are co-adapted meme complexes---though only one of many supported by any given brain. Like religions, political belief systems and cults, they are sets of memes that thrive in each other's company. Like religions, political belief systems and cults, they are safe havens for all sorts of traveling memes and they are protected from destruction by various meme-tricks. They do not have to be true. In fact we know, of course, that selves are a myth. Look inside the brain and you find only neurons. You do not find the little person pulling the strings or the homunculus watching the show on an inner screen. You do not find the place where 'my' conscious decisions are made. You do not find the thing that lovingly holds all those beliefs and opinions. Most of us still persist in thinking about ourselves that way. But the truth is---there is no one in there!

We now have a radically new answer to the question ``who am I?" and a rather terrifying one. 'I am one of the many co-adapted meme-complexes living within this brain. No wonder people want to raise their hands in scare quotes and giggle nervously at the very idea of ``memes."

## References

Gould,S.J. (1996) *Life's Grandeur* (US Full House) London, Jonathan Cape---see chapter 15. He used the term ``meaningless metaphor" on BBC Radio 4's *Start the Week* November 11 1996.

Dawkins,R. (1976) *The Selfish Gene* Oxford, OUP (new edition with additional material 1989).

Dawkins,R. (1996) *Climbing Mount Improbable* London, Viking.

*The Selfish Gene* p322.

*The Selfish Gene* p192.

Dawkins,R. (1993) Viruses of the Mind. In B.Dahlbohm (ed) *Dennett and his Critics: Demystifying Mind* Oxford, Blackwell.

Dennett,D. (1995) *Darwin's Dangerous Idea* London, Penguin. On page 361 Dennett says of Dawkins ``He has since drawn in his horns slightly," and asks ``Why, indeed, is the meme meme so little discussed eighteen years after *The Selfish Gene* appeared?" Note that Dennett himself does not mention memes in his latest book *Kinds of Minds*, New York, Basic Books, (1996). The word ``meme" is not listed in the index of any of the following books:

Barkow,J.H., Cosmides,L. and Tooby,J. (Eds) (1992) *The Adapted Mind: Evolutionary psychology and the generation of culture* Oxford, OUP.

Dunbar,R. (1996) *Grooming, Gossip and the Evolution of Language* London, Faber & Faber.

Mithen,S. (1996) *The Prehistory of the Mind* London, Thames and Hudson.

- Pinker, S. (1994) *The Language Instinct* New York, Morrow.
- Ridley,M. (1996) *The Origins of Virtue* London, Viking.
- Tudge,C. (1996) *The Day before Yesterday*.
- Wills,C. (1993) *The Runaway Brain: The Evolution of human uniqueness* Basic Books.
- Wright,R. (1994) *The Moral Animal* Pantheon Books (and London, Abacus, 1996).
- (Cronin,H. (1991) *The Ant and the Peacock* Cambridge, CUP does have one entry)
- Gould,S.J. (1996) *Life's Grandeur*.
- Wilson,E.O. (1978) *On Human Nature* Cambridge, Mass. Harvard University Press.
- Durham,W.H. (1991) *Coevolution: Genes, culture and human diversity* Stanford, Ca., Stanford University Press.

For example, Cavalli-Sforza and Feldman have developed a detailed scheme based on cultural traits, including notions of cultural fitness and cultural selection pressures, yet they seem to expect natural selection to rein in cultural evolution.

Cavalli-Sforza,L.L. and Feldman,M.W. (1981) *Cultural Transmission and Evolution: A quantitative approach* Princeton, NJ, Princeton University Press.

Boyd and Richerson have analyzed the forces of cultural adaptation and yet still, when they say generation they mean genetic generation, and when they say ``adaptive'' or ``maladaptive'' they mean for the genes.

Boyd,R. and Richerson,P.J. (1985) *Culture and the Evolutionary Process* Chicago, University of Chicago Press.

Richerson,P.J. and Boyd,R. (1992) *Cultural inheritance and evolutionary ecology* In E.A.Smith and B.Winterhalder (Eds) *Evolutionary Ecology and Human Behaviour* 61-92.

Dennett,D. (1991) *Consciousness Explained* Boston and London, Little, Brown & Co.

Dennett,D. (1995) *Darwin's Dangerous Idea* London, Penguin.

[DDI341]  
*Darwin's Dangerous Idea* p341

See for example:  
[http://www.sepa.tudelft.nl/~afd\\_ba/mem.html](http://www.sepa.tudelft.nl/~afd_ba/mem.html)

<http://www.lucifer.com/virus/alt.memetics/>  
<http://www.lycaeum.org/~sputnik/memetics/>

Dawkins (1982) *The Extended Phenotype* Oxford, OUP. p110 discusses this in detail. Durham (1991) p187 calls this principle ``Campbell's Rule" quoting from Campbell (1965) ``The analogy to cultural accumulations (is not) from organic evolution per se: but rather from a general model... for which organic evolution is but one instance." See also Dennett (1991) p200.

For definitions of the meme see Dawkins (1976) p192 and (1982) p110.

Dawkins (1976) p192. R. Webb also takes this view. ``Fossil Soup: The evolution of technology: life as we know it or a pre-biotic soup?" presented at LSE, 27 June 1996

Dennett (1995) p373-381

Dawkins (1976) p28.

For good discussions of human altruism see:

Cronin,H. (1991) *The Ant and the Peacock* Cambridge, CUP.

Ridley,M. (1996) *The Origins of Virtue* London, Viking.

Wright,R. (1994) *The Moral Animal* Pantheon Books (and London, Abacus, 1996).

For recent discussions of brain size see Leakey,R. (1994) *The Origin of Humankind* London, Weidenfeld and Nicolson; Wills,C. (1993) *The Runaway Brain* London, Basic Books.

See Leakey (1994) p98. For further discussion of the mechanisms of competition see Tudge (1996).

Dunbar,R. (1996) *Grooming, Gossip and the Evolution of Language* London, Faber & Faber.

Wills,C. (1993) *The Runaway Brain* London, Basic Books. For discussion of

Miller see Mestel,R. Arts of Seduction, *New Scientist*, 23/30 December 1995 28-31. Richerson and Boyd (1992) p75.

Dennett (1991) describes this imaginary place as the ``Cartesian Theatre" and argues that we may reject Cartesian Dualism but still implicitly believe in the CT.

## **WAKING FROM THE MEME DREAM**

**By**

**Susan Blackmore**

Department of Psychology

University of the West of England

Bristol BS16 2JP

Paper presented at:

The Psychology of Awakening: International Conference on Buddhism, Science and Psychotherapy Dartington 7-10 November 1996

Wake up! Wake up!

Errrr, ummmmm, grrrggr, Oh yes, I'm awake now. Wow, that was a weird dream. I really thought I had to escape from the slurb, and it mattered terribly to get to the cupboard in time. How silly! Of course, now I see it wasn't real at all.

Wake up! Wake up!

What do you mean, "wake up", I'm already awake. This is real. This does matter. I can't wake up any more. Go away!

Wake up! Wake up!

But I don't understand - From what? And how?

These are the questions I want to tackle today. From what are we to awaken? And how? My answers will be "From the meme dream" and "By seeing that it is a meme dream". But it may take me some time to explain!

There is a long history, in spiritual and religious traditions, of the idea that normal waking life is a dream or illusion. This makes no sense to someone who looks around and is convinced there is a real world out there and a self who perceives it. However, there are many clues that this ordinary view is false.

Some clues come from spontaneous mystical experiences in which people "see the light!", realize that everything is one, and go "beyond self" to see the world "as it really is". They feel certain that the new way of seeing is better and truer than the old (though of course they could be mistaken!).

Other clues come from spiritual practice. Probably the first thing that anybody discovers when they try to meditate, or be mindful, is that their mind is constantly full of thoughts. Typically these are not wise and wonderful thoughts, or even useful and productive thoughts, but just endless chatter. From the truly trivial to the emotionally entangling, they go on and on. And what's more they nearly all involve "me". It is a short step to wondering who this suffering self is, and why "I" can't stop the thoughts.

Finally clues come from science. The most obvious (and scary) conclusion from modern neuroscience is that there is simply no one inside the brain. The more we learn about the way the brain functions the less it seems to need a central controller, a little person inside, a decider of decisions or an experiencer of experiences. These are just fictions - part of the story the brain tells itself about a self within (Churchland and Sejnowski, 1992; Dennett, 1991).

Some say there is no point in striving for an intellectual understanding of spiritual matters. I disagree. It is true that intellectual understanding is not the same as realization, but this does not mean it is useless. In my own tradition of practice, Zen, there is much room for intellectual struggle; for example, in the cultivation of the “don’t know mind”, or in working with koans. You can bring a question to such a state of intellectual confusion that it can be held, poised, in all its complexity and simplicity. Like “Who am I?”, “What is this?” or (one I have struggled with) “What drives you?”

There is also a terrible danger in refusing to be intellectual about spiritual matters. That is, we may divorce our spiritual practice from the science on which our whole society depends. If this society is going to have any spiritual depths to it, they must fit happily with our growing understanding of the workings of the brain and the nature of mind. We cannot afford to have one world in which scientists understand the mind, and another in which special people become enlightened.

So I make no apologies for my approach. I am going to try to answer my questions using the best science I can find. We seem to live in a muddle that we think matters to a self that doesn’t exist. I want to find out why.

## **Darwin’s Dangerous Idea**

There is one scientific idea which, to my mind, excels all others. It is exquisitely simple and beautiful. It explains the origins of all life forms and all biological design. It does away with the need for God, for a designer, for a master plan or for a purpose in life. Only in the light of this idea does anything in biology make sense. It is, of course, Darwin’s idea of evolution by natural selection.

The implications of natural selection are so profound that people have been awe-struck or maddened; fascinated or outraged, since it was first proposed in *The Origin of Species* in 1859. This is why Dennett (1995) calls it Darwin’s Dangerous Idea. Sadly, many people have misunderstood the idea and, even worse, have used it to defend indefensible political doctrines which have nothing to do with Darwinism. I therefore hope you will forgive me if I spend some time explaining it as clearly as I can.

All you need for natural selection to get started is a replicator in an appropriate environment. A replicator is something that copies itself, though not always perfectly. The environment must be one in which the replicator can create numerous copies of itself, not all of which can survive. That’s it.

Can it really be that simple? Yes. All that happens is this - in any one copying generation, not all the copies are identical and some are better able to survive in that environment than others are. In consequence they make more copies of themselves and so that kind of copy becomes more numerous. Of course things then begin to get complicated. The rapidly expanding population of copies starts to change the environment and that changes the selective pressures. Local variations in the environment mean different kinds of copy will do well in different places and so more complexity arises. This way the process can produce all the kinds of organized complexity we see in the living world - yet all it needs is this one simple, elegant, beautiful, and obvious process - natural selection.

To make things more concrete let's imagine a primeval soup in which a simple chemical replicator has arisen. We'll call the replicators "Bobbies". These bobbies, by virtue of their chemical constitution, just do make copies of themselves whenever they find the right chemicals. Now, put them in a rich chemical swamp and they start copying, though with occasional errors. A few million years go by and there are lots of kinds of bobbies. The ones that need lots of swampon have used up all the supplies and are failing, so now the sort that can use isoswampin instead, are doing better. Soon there are several areas in which different chemicals predominate and different kinds of blobby appear. Competition for swamp chemicals gets fierce and most copies that are made die out. Only those that, by rare chance, turn out to have clever new properties, go on go on to copy themselves again.

Clever properties might include the ability to move around and find the swampon, to trap isoswampin<sup>3-7</sup> and hang onto it, or to build a membrane around themselves. Once bobbies with membranes appear, they will start winning out over free-floating ones and super-bobbies are made.

Another few million years go by and tricks are discovered like taking other bobbies inside the membrane, or joining several super-bobbies together. Super-dooper-bobbies appear, like multi-celled animals with power supplies and specialized parts for moving about and protecting themselves. However, these are only food to even bigger super-dooper-bobbies. It is only a matter of time before random variation and natural selection will create a vast living world. In the process billions and billions of unsuccessful bobbies have been created and died, but such a slow, blind process produces the goods. "The goods" on our planet includes bacteria and plants, fish and frogs, duck-billed platypuses and us.

Design appears out of nothing. There is no need for a creator or a master plan, and no end point towards which creation is heading. Richard Dawkins (1996) calls it "Climbing Mount Improbable". It is just a simple but inexorable process by which unbelievably improbable things get created.

It is important to remember that evolution has no foresight and so doesn't necessarily produce the "best" solution. Evolution can only go on from where it is now. That is why, among other things, we have such a daft design in our eyes, with all the neurons going out of the front of the retina and getting in the way of the light. Once evolution had started off on this kind of eye it was stuck with it. There was no creator around to say "hey, start again with that one, let's put the

wires out the back". Nor was there a creator around to say "Hey, let's make it fun for the humans". The genes simply do not care.

Understanding the fantastic process of natural selection we can see how our human bodies came to be the way they are. But what about our minds? Evolutionary psychology does not easily answer my questions.

For example, why do we think all the time? From a genetic point of view this seems extremely wasteful - and animals that waste energy don't survive. The brain uses about 20% of the body's energy while weighing only 2%. If we were thinking useful thoughts, or solving relevant problems there might be some point, but mostly we don't seem to be.

So why can't we just sit down and not think?

Why do we believe in a self that does not exist? Someone may yet explain this in evolutionary terms, but at least superficially it appears pointless. Why construct a false idea of self, with all its mechanisms protecting self-esteem and its fear of failure and loss, when from the biological point of view it is the body that needs protecting. Note that if we thought of ourselves as the entire organism there would be no problem, but we don't - rather, we seem to believe in a separate self; something that is in charge of the body; something that has to be protected for its own sake. I bet if I asked you "Which would you rather lose - your body or your mind?" you wouldn't spend long deciding.

Like many other scientists I would love to find a principle as simple, as beautiful and as elegant as natural selection that would explain the nature of the mind.

I think there is one. It is closely related to natural selection. Although it has been around for twenty years, it has not yet been put fully to use. It is the theory of memes.

## A Brief History of the Meme Meme

In 1976 Richard Dawkins wrote what is probably the most popular book ever on evolution - *The Selfish Gene*. The book gave a catchy name to the theory that evolution proceeds entirely for the sake of the selfish replicators. That is, evolution happens not for the good of the species, nor for the good of the group, nor even for the individual organism. It is all for the good of the genes. Genes that are successful spread and those that aren't don't. The rest is all a consequence of this fact.

Of course the main replicator he considered was the gene - a unit of information coded in the DNA and read out in protein synthesis. However, at the very end of the book he claimed that there is another replicator on this planet; the meme.

The meme is a unit of information (or instruction for behavior) stored in a brain and passed on by imitation from one brain to another. Dawkins gave as examples; ideas, tunes, scientific theories, religious beliefs, clothes fashions, and skills, such as new ways of making pots or building arches.

The implications of this idea are staggering and Dawkins spelt some of them out. If memes are really replicators then they will, inevitably, behave selfishly. That is, ones that are good at spreading will spread and ones that are not will not. As a consequence the world of ideas - or memosphere - will not fill up with the best, truest, most hopeful or helpful ideas, but with the survivors. Memes are just survivors like genes.

In the process of surviving they will, just like genes, create mutually supportive meme groups. Remember the blobbies. In a few million years they began to get together into groups, because the ones in groups survived better than loners. The groups got bigger and better, and a complex ecosystem evolved. In the real world of biology, genes have grouped together to create enormous creatures that then mate and pass the groups on. In a similar way memes may group together in human brains and fill the world of ideas with their products.

If this view is correct, then the memes should be able to evolve quite independently of the genes (apart from needing a brain). There have been many attempts to study cultural evolution, but most of them implicitly treat ideas (or memes) as subservient to the genes (see e.g. Cavalli-Sforza and Feldman, 1981; Crook, 1995; Durham, 1991; Lumsden and Wilson, 1981). The power of realizing that memes are replicators is that they can be seen as working purely and simply in their own interest. Of course to some extent memes will be successful if they are useful to their hosts, but this is not the only way for a meme to survive - and we shall soon see some consequences of this.

Since he first suggested the idea of memes Dawkins has discussed the spread of such behaviors as wearing baseball caps back to front (my kids have recently turned theirs the right way round again!), the use of special clothing markers to identify gangs, and (most famously) the power of religions. Religions are, according to Dawkins (1993), huge co-adapted meme-complexes; that is groups of memes that hang around together for mutual support and thereby survive better than lone memes could do. Other meme-complexes include cults, political systems, alternative belief systems, and scientific theories and paradigms.

Religions are special because they use just about every meme-trick in the book (which is presumably why they last so long and infect so many brains). Think of it this way. The idea of hell is initially useful because the fear of hell reinforces socially desirable behavior. Now add the idea that unbelievers go to hell, and the meme and any companions are well protected. The idea of God is a natural companion meme, assuaging fear and providing (spurious) comfort. The spread of the meme-complex is aided by exhortations to convert others and by tricks such as the celibate priesthood. Celibacy is a disaster for genes, but will help spread memes since a celibate priest has more time to spend promoting his faith.

Another trick is to value faith and suppress the doubt that leads every child to ask difficult questions like "where is hell?" and "If God is so good why did those people get tortured?" Note that science (and some forms of Buddhism) does the opposite and encourage doubt.

Finally, once you've been infected with these meme-complexes they are hard to get rid of. If you try to throw them out, some even protect themselves with last-ditch threats of death, ex-communication, or burning in hell-fire for eternity.

I shouldn't get carried away. The point I want to make is that these religious memes have not survived for centuries because they are true, because they are useful to the genes, or because they make us happy. In fact I think they are false and are responsible for the worst miseries in human history. No - they have survived because they are selfish memes and are good at surviving - they need no other reason.

Once you start to think this way a truly frightening prospect opens up. We have all become used to thinking of our bodies as biological organisms created by evolution. Yet we still like to think of ourselves as something more. We are in charge of our bodies, we run the show, we decide which ideas to believe in and which to reject. But do we really? If you begin to think about selfish memes it becomes clear that our ideas are in our heads because they are successful memes. American philosopher Dan Dennett (1995) concludes that a "person" is a particular sort of animal infested with memes. In other words you and I and all our friends are the products of two blind replicators, the genes and the memes.

I find these ideas absolutely stunning. Potentially we might be able to understand all of mental life in terms of the competition between memes, just as we can understand all biological life in terms of the competition between genes.

What I want to do now, finally, is apply the ideas of memetics to the questions I asked at the beginning. What are we waking up from and how do we do it?

Why is my head so full of thoughts?

This question has a ridiculously easy answer once you start thinking in terms of memes. If a meme is going to survive it needs to be safely stored in a human brain and passed accurately on to more brains. A meme that buries itself deep in the memory and never shows itself again will simply fizzle out. A meme that gets terribly distorted in the memory or in transmission will also fizzle out. One simple way of ensuring survival is for a meme to get itself repeatedly rehearsed inside your head.

Take two tunes. One of them is tricky to sing, and even harder to sing silently to yourself. The other is a catchy little number that you almost can't help humming to yourself. So you do. It goes round and round. Next time you feel like singing aloud this tune is more likely to be picked for the singing. And if anyone is listening they'll pick it up too. That's how it became successful, and that's why the world is so full of awful catchy tunes and advertising jingles.

But there is another consequence. Our brains get full up with them too. These successful memes hop from person to person, filling up their hosts' minds as they go. In this way all our minds get fuller and fuller.

We can apply the same logic to other kinds of meme. Ideas that go round and round in your head will be successful. Not only will they be well remembered, but when you are next talking to someone they will be the ideas "on your mind" and so will get passed on. They may get to this position by being emotionally charged, exciting, easily memorable or relevant to your current

concerns. It does not matter how they do it. The point is that memes that get themselves repeated will generally win out over ones that don't. The obvious consequence of this fact is that your head will soon fill up with ideas. Any attempt to clear the mind just creates spare processing capacity for other memes to grab.

This simple logic explains why it is so hard for us to sit down and "not think"; why the battle to subdue "our" thoughts is doomed. In a very real sense they are not "our" thoughts at all. They are simply the memes that happen to be successfully exploiting our brain-ware at the moment.

This raises the tricky question of who is thinking or not thinking. Who is to do battle with the selfish memes? In other words, who am I?

Who am I?

I suppose you can tell by now what my answer to this one is going to be. We are just co-adapted meme-complexes. We, our precious, mythical "selves", are just groups of selfish memes that have come together by and for themselves.

This is a truly startling idea and, in my experience, the better you understand it, the more fascinating and weird it becomes. It dismantles our ordinary way of thinking about ourselves and raises bizarre questions about the relationship of ourselves to our ideas. To understand it we need to think about how and why memes get together into groups at all.

Just as with blobbies or genes, memes in groups are safer than free-floating memes. An idea that is firmly embedded in a meme-complex is more likely to survive in the memosphere than is an isolated idea. This may be because ideas within meme-groups get passed on together (e.g. when someone is converted to a faith, theory or political creed), get mutual support (e.g. if you hate the free-market economy you are likely also to favor a generous welfare state), and they protect themselves from destruction. If they did not, they would not last and would not be around today. The meme-complexes we come across are all the successful ones!

Like religions, astrology is a successful meme-complex. The idea that Leos get on well with Aquarians is unlikely to survive on its own, but as part of astrology is easy to remember and pass on. Astrology has obvious appeal that gets it into your brain in the first place; it provides a nice (though spurious) explanation for human differences and a comforting (though false) sense of predictability. It is easily expandable (you can go on adding new ideas forever!) and is highly resistant to being overturned by evidence. In fact the results of hundreds of experiments show that the claims of astrology are false but this has apparently not reduced belief in astrology one bit (Dean, Mather and Kelly, 1996). Clearly, once you believe in astrology it is hard work to root out all the beliefs and find alternatives. It may not be worth the effort. Thus we all become unwitting hosts to an enormous baggage of useless and even harmful meme-complexes.

One of those is myself.

Why do I say that the self is a meme-complex? Because it works the same way as other meme-complexes. As with astrology, the idea of "self" has a good reason for getting installed in the

first place. Then once it is in place, memes inside the complex are mutually supportive, can go on being added to almost infinitely, and the whole complex is resistant to evidence that it is false.

First the idea of self has to get in there. Imagine a highly intelligent and social creature without language. She will need a sense of self to predict others' behavior (Humphrey, 1986) and to deal with ownership, deception, friendships and alliances (Crook, 1980). With this straightforward sense of self she may know that her daughter is afraid of a high ranking female and take steps to protect her, but she does not have the language with which to think "I believe that my daughter is afraid ... etc.". It is with language that the memes really get going - and with language that "I" appears. Lots of simple memes can then become united as "my" beliefs, desires and opinions.

As an example, let's consider the idea of sex differences in ability. As an abstract idea (or isolated meme) this is unlikely to be a winner. But get it into the form "I believe in the equality of the sexes" and it suddenly has the enormous weight of "self" behind it. "I" will fight for this idea as though I were being threatened. I might argue with friends, write opinion pieces, or go on marches. The meme is safe inside the haven of "self" even in the face of evidence against it. "My" ideas are protected.

Then they start proliferating. Ideas that can get inside a self - that is, be "my" ideas, or "my" opinions, are winners. So we all get lots of them. Before we know it, "we" are a vast conglomerate of successful memes. Of course there is no "I" who "has" the opinions. That is obviously a nonsense when you think clearly about it. Yes, of course there is a body that says "I believe in being nice to people" and a body that is (or is not) nice to people, but there is not in addition a self who "has" the belief.

Now we have a radically new idea of who we are. We are just temporary conglomerations of ideas, molded together for their own protection. The analogy with our bodies is close. Bodies are the creations of temporary gene-complexes: although each of us is unique, the genes themselves have all come from previous creatures and will, if we reproduce, go on into future creatures. Our minds are the creations of temporary meme-complexes: although each of us is unique, the memes themselves have come from previous creatures and will, if we speak and write and communicate, go on into future creatures. That's all.

The problem is that we don't see it this way. We believe there really is someone inside to do the believing, and really someone who needs to be protected. This is the illusion - this is the meme-dream from which we can wake up.

## Dismantling the Meme-Dream

There are two systems I know of that are capable of dismantling meme-complexes (though I am sure there are others). Of course these systems are memes themselves but they are, if you like, meme-disinfectants, meme-eating memes, or "meme-complex destroying meme-complexes". These two are science and Zen.

Science works this way because of its ideals of truth and seeking evidence. It doesn't always live up to these ideals, but in principle it is capable of destroying any untruthful meme-complex by putting it to the test, by demanding evidence, or by devising an experiment.

Zen does this too, though the methods are completely different. In Zen training every concept is held up to scrutiny, nothing is left uninvestigated, even the self who is doing the investigation is to be held up to the light and questioned. "Who are you?"

After about 15 years of Zen practice, and when reading The Three Pillars of Zen by Philip Kapleau, I began working with the koan "Who...?" The experience was most interesting and I can best liken it to watching a meme unzipping other memes. Every thought that came up in meditation was met with "Who is thinking that?" or "Who is seeing this?" or "Who is feeling that?" or just "Who...?" Seeing the false self as a vast meme-complex seemed to help - for it is much easier to let go of passing memes than of a real, solid and permanent self. It is much easier to let the meme-unzipper do its stuff if you know that all it's doing is unzipping memes.

Another koan of mine fell to the memes. Q. "Who drives you?" A. "The memes of course." This isn't just an intellectual answer, but a way into seeing yourself as a temporary passing construction. The question dissolves when both self and driver are seen as memes.

I have had to take a long route to answer my questions but I hope you can now understand my answers. "From what are we to awaken? From the meme dream of course. And how?" "By seeing that it is a meme dream".

And who lets the meme-unzipper go its way? Who wakes up when the meme-dream is all dismantled? Ah, there's a question.

## References

- Cavalli-Sforza,L.L. and Feldman,M.W. (1981) Cultural transmission and Evolution: A quantitative approach. Princeton NJ, Princeton University Press.
- Churchland,P.S. and Sejnowski,T.J. (1992) The Computational Brain. Cambridge, Mass. MIT Press
- Crook, J.H. (1980) The Evolution of Human Consciousness. Oxford, Clarendon Press.
- Crook,J.H. (1995) Psychological processes in cultural and genetic coevolution. In Survival and Religion: Biological Evolution and Cultural Change. Ed. E.Jones and V.Reynolds. London, Wiley. 45-110
- Darwin, C. (1859) On the Origin of Species by Means of Natural Selection. London, Murray.
- Dawkins,R. (1976) The Selfish Gene, Oxford University Press.
- Dawkins,R. (1993) Viruses of the Mind. In Dennett and his Critics. Ed. B.Dahlbom, Oxford, Blackwell. 13-27.

- Dawkins,R. (1996) Climbing Mount Improbable. London, Viking
- Dean,G., Mather,A. and Kelly,I.W. (1996) Astrology. In The Encyclopedia of the Paranormal, Ed. G.Stein, New York, Prometheus, 47-99.
- Dennett,D.C. (1991) Consciousness Explained. London, Little, Brown & Co.
- Dennett,D. (1995) Darwin's Dangerous Idea, London, Penguin
- Durham,W.H. (1991) Coevolution: Genes, culture and human diversity. Palo Alto, Ca., Stanford University Press.
- Humphrey,N. (1986) The Inner Eye London, Faber and Faber
- Kapleau,P. (1965) The Three Pillars of Zen. John Weatherhill (and Doubleday, 1989)
- Lumsden,C.T. and Wilson,E.O. (1981) Genes, Mind and Culture: The Co-evolutionary Process. Cambridge, Mass., Harvard University Press.

# IMITATION AND THE DEFINITION OF A MEME

Susan Blackmore

University of the West of England  
St. Matthias College  
Bristol BS16 2JP

## Abstract

The dictionary definition, and Dawkins's (1976) original conception of the meme, both include the idea that memes are copied from one person to another by imitation. We therefore need to be clear what is meant by imitation. Imitation is distinguished from contagion, individual learning and various kinds of non-imitative social learning such as stimulus enhancement, local enhancement and goal emulation. True imitation is extremely rare in animals other than humans, except for birdsong and dolphin vocalization, suggesting that they can have few or no memes. I argue that more complex human cognitive processes, such as language, reading, scientific research and so on, all build in some way on the ability to imitate, and therefore all these processes are, or can be, memetic. When we are clear about the nature of imitation, it is obvious what does and does not count as a meme. I suggest that we stick to defining the meme as that which is passed on by imitation.

## 1. Introduction

There are many ways of defining the meme but there are two that we should perhaps take particularly seriously. First, Dawkins, who coined the term meme, described memes as units of cultural transmission which "*propagate themselves in the meme pool by ... a process which, in the broad sense, can be called imitation*" (Dawkins, 1976 p 192). Second, the Oxford English Dictionary defines a meme as follows: "*meme (mi:m), n. Biol. (shortened from mimeme ... that which is imitated, after GENE n.). An element of a culture that may be considered to be passed on by non-genetic means, esp. imitation*". Both these definitions include the critical point that memes are cultural information that is copied, and that it is copied by imitation. The OED is arguably the most important dictionary of the English language and is, as far as I know, the first to include the word 'meme'. It would be unfortunate if memeticists began to use definitions of the meme that were incompatible with the dictionary definition, unless there were good reasons for doing so.

Some technical definitions are quite unlike the dictionary one, but are nonetheless perfectly compatible with it. A good example is Wilkins's: "*A meme is the least unit of sociocultural information relative to a selection process that has favorable or unfavourable selection bias that exceeds its endogenous tendency to change.*" This is useful because it emphasizes, first, that the size of the relevant unit is not fixed but can vary in different contexts and, second, the importance of fidelity (i.e. the stability or resistance to change of the information). In this way Wilkins's (1998) definition of the meme is similar to Williams's (1966) definition of the gene. This definition may be useful for theoretical purposes but is too complicated for more general use, or for popular treatments of memetics. However this is not a problem since Wilkins's definition does not conflict with the OED definition.

Some other definitions clearly are incompatible with the OED. Some, for example, imply that almost everything we know or experience can count as a meme, whether acquired by imitation or not (e.g. Brodie, 1996; Gabora, 1997; Lynch, 1996). Brodie includes operant conditioning, and indeed all conditioning, as memetic. Gabora goes even further and includes ideas, perceptions, emotions, attitudes, and indeed "*anything that can be the subject of an instant of experience*". According to this broadest definition a garden frog would have a mass of memes (even though it is totally incapable of imitation or any kind of culture) because it has perceptions and emotions, and is capable of many kinds of learning.

I shall argue that these broader definitions are deeply confusing. They take away the idea of the meme as a replicator (which was the original reason for its invention, and provides its context within evolutionary theory), ignore the idea that memes must be passed on by some kind of copying, and merely add confusion to the already difficult problem of understanding consciousness. I suggest we are better to stick to the original definition of the meme as transmitted by imitation.

## 1.1 Defining Imitation

What is imitation? In this paper I want to tackle the question from two directions. First I will consider simpler kinds of learning and cognitive processes which may, or may not, be counted as imitation. Second I will consider whether imitation includes all higher-order human cognitive processes, such as speech, reading, teaching and instruction, upon which much of our cultural life depends. I will argue that when we have clarified these issues we will no longer have serious problems in defining the meme.

There is a long history of research on imitation in both animal behavior and human social psychology (for review see Whiten and Ham, 1992). In the nineteenth century Darwin collected many examples of what he took to be imitation in animals, as did Romanes (1882, 1883) but they did not define what they meant by imitation. Baldwin (1902) gave imitation a central role in his theories of evolution, pointing out that all adaptive processes can be seen as imitative - perhaps foreshadowing the universal Darwinism that today enables comparisons between biological evolution and memetic evolution (e.g. Dawkins, 1976; Plotkin, 1993). The psychologist, Thorndike (1898), was possibly the first to provide a clear definition of imitation as "*learning to do an act from seeing it done*".

Thorndike's definition (though confined to visual information) captures the essential idea that in imitation a new behavior is learned by copying it from someone else. One hundred years later we can see the importance of this point in distinguishing imitation from simple contagion and from other kinds of learning. These other kinds of learning can be divided roughly into individual learning and non-imitative social learning.

## 2. Contagion

The term "contagion" is often associated with memetics. We may say that certain memes are contagious, or more contagious than others. We may treat the spread of memes as comparable with the spread of infectious or contagious diseases and use models derived from epidemiology

(Lynch, 1996). The term social contagion is often used to include phenomena that are certainly memetic, such as the spread of fads, hysterical reactions (Showalter, 1997), or even suicide (Marsden, 1997). However, the term is used in confusing ways (Levy & Nail, 1993) and there is one kind of contagion that we must clearly distinguish from imitation.

This is what has variously been called instinctive imitation, imitative suggestion, social facilitation, coaction, and (simply) contagion (Whiten & Ham, 1992). Examples in humans include the spread of yawning, coughing or laughter. All these behaviors are extremely contagious. Indeed it can be difficult not to laugh if everyone around you is already laughing. This kind of contagion probably relies on specific stimulus feature detectors which detect laughing or yawning in someone else and then trigger the same innate behavior as the response. In other animals there are many examples of contagious vocalizations, such as alarm calls. Most vertebrates yawn and some animals, such as chimpanzees, laugh in response to tickling and play, but contagious laughter appears to be limited to humans (Provine, 1996).

This kind of contagion is not true imitation. We can see why by considering Thorndike's simple definition. Yawning, coughing and laughing are innate behaviours. When we start laughing because everyone else is laughing we have not learned how to do an act. We already knew how to laugh, and the kind of laugh we make is not modelled on the laugh we hear. So this kind of contagion is not imitation and should not be counted as memetic.

### **3. Individual Learning**

In individual learning a person or animal learns something by itself, without anyone else necessarily being involved. There are traditionally two major types of learning in psychology - classical conditioning and operant conditioning.

#### **3.1 Classical conditioning**

Classical conditioning is when two stimuli become associated by repeated pairing. In the best known experiments Pavlov paired sounds with the smell of meat and found that dogs then salivated to the sounds even without any meat. Classical conditioning is widespread in the animal kingdom, for example when animals learn to distinguish palatable foods from poisonous foods, or learn other important facts about their environment. It occurs in humans whenever we associate two things together because they have previously been paired, whether those things are sights, sounds, tunes, ideas or pain. Behavior is changed by the process but nothing is passed on by imitation from one person to another, so the process is not memetic.

You may say that Pavlov, in setting up the experiments in the first place, was passing on something to his dogs. However, the dogs were not imitating him. There is no replication or copying of behavior from Pavlov to his dogs. Similarly if you have a dog or cat, it probably starts salivating at the sound of the fridge opening or the knife hitting the food dish. Or maybe it gets frightened at the sight of a new flea collar. This is classical conditioning at work. You have certainly trained the animals by the contingencies you set up, but there is no imitation involved.

### **3.2 Operant conditioning**

Operant conditioning is when a behavior made by an animal is either rewarded or punished and therefore either increases or decreases in frequency. A hundred years ago Thorndike studied this kind of trial and error learning in cats trying to escape from specially designed boxes. In the 1930s Skinner famously provided animals with levers which, when pressed, caused food to be delivered. Rats, pigeons and other animals quickly learn to press the levers and their subsequent behavior depends on the schedules of reinforcement used.

Skinner (1953) pointed out the similarity between operant conditioning and natural selection - some behaviors are positively selected and others weeded out. In this way learning can be seen as an evolutionary system in which the behaviors are the replicators. Several selection theories of learning and of brain development have since been proposed (e.g. Calvin, 1987; Edelmann, 1989; Plotkin, 1993), and are important for understanding the breadth of application of Darwinian processes. However, as long as the behaviors cannot be passed on to someone else by imitation then they do not count as memes and the selection is not memetic selection.

Much of human learning is Skinnerian and not memetic. Whether consciously or not, parents shape their children's behavior by the way they reinforce them. The most effective reward for children is attention and rewards work better than punishment. So if parents pay lots of attention to their children when they are behaving well, and act uninterested when they scream or have tantrums, then behaving well is in the best interests of the children and they will do it. The parents' behavior can be seen as part of the environment in which the children learn, or as part of a complex pattern of social learning (discussed below). Either way if the children are not imitating the parents then the process is not memetic.

We learn many things by trial and error, such as the physical skills of walking without falling over or riding a bike, or general ways of interacting with other people and the world. For example, people who are generally rewarded for hard work and persistence will behave differently from people whose efforts are met with arbitrary results. Of course memes may be involved - such as the very idea of riding a bike in the first place - but whenever we repeat actions that led to successful outcomes and suppress actions that led to pain or failure, then we are learning for ourselves by operant conditioning. We are not acquiring new memes.

### **3.3 Non-memetic learning**

There are other kinds of learning that are also not memetic such as the formation of cognitive maps. Many animals develop complex mental maps of their environment without which they could not live at all, whether they are cats, rats, insects or birds. Some have complex territorial systems in which boundaries are carefully guarded, some (like squirrels for example) hide large numbers of food items are able accurately to find them again, while others use well known paths to explore and find food. The information in the maps is learned by exploration and conditioning. There is no imitation involved. Similarly we develop complex cognitive maps of our own house and garden, the city we live in, and the places we go for our holidays. We can find our way around these places and conjure them up in our imagination. All this is individual learning and not memetic.

In practice we may not always be able to tease out those things we have individually learned by conditioning from those we have learned by imitation, and very often both are involved - but in principle the two are different. We know lots of things that are not memes.

## 4. Social Learning

Social learning means learning something from other people, (or, more generally, from conspecifics). Very often classical and operant conditioning are the basic processes involved, but something is learned in a way that involves other people or animals.

Social learning includes true imitation, but there are other kinds of social learning as well. According to Heyes (1993) the difference is best explained like this. Imitation means learning something about the form of behavior through observing others, while other kinds of social learning are learning about the environment through observing others. I like to think of the difference in terms of copying. In true imitation something about the action is copied from actor to imitator, while in other forms of social learning nothing is copied. This is important because evolutionary processes depend on there being something that is copied or replicated. Note that this fits well with Thorndike's definition of imitation as "*learning to do an act from seeing it done*". If memes depend on imitation for their transmission then we need to be clear about how to distinguish true imitation, which involves learning by copying a new form of behavior, from other kinds of social learning which do not.

### 4.1 Stimulus enhancement, local enhancement, and goal emulation

In 1921, in the south of England, some small garden birds called tits were seen prising open the wax-board tops of milk bottles left on the doorstep. The habit later became widespread across England and some parts of Scotland and Wales, with other species of bird joining in, and foil tops being pecked as well. It seemed that the tits learned from each other because this new behavior spread gradually from village to village, and across different areas, although it was obviously independently reinvented many times (Fisher & Hinde, 1949). The spread of milk-bottle pecking was a simple cultural phenomenon but purists would argue that it was based not on imitation, but on social learning (Sherry & Galef, 1984). Imagine that one bird learned, by trial and error, that there was cream under the bottle top. Then imagine another bird came by and saw the pecking and the obviously pecked top. Pecking is a natural action for tits and their attention can easily be drawn to something like a pecked top by the actions of other birds. So now the second bird is more likely to land on the bottle and peck too. Reinforcement from the cream now means that this bird is likely to repeat the action. It may then be seen by other birds. So this learning involves operant conditioning (the effect of the cream), is a kind of social learning (because another bird is involved), but is not true imitation (because the pecking is not actually copied). The fact that the birds used lots of different methods for opening the bottles also suggests they did not learn by direct imitation.

This kind of social learning is sometimes called "stimulus enhancement" - the stimulus, in this case the bottle top, has become more readily noticed by the birds. Another form of social learning that is not imitation is called "local enhancement". This is when attention is directed towards a specific place. For example, animals learn from each other which objects or places to

fear or ignore. Birds and rabbits learn not to fear trains by following others who are not afraid and therefore become used to the frightening noise. Rabbits can therefore establish warrens on railway embankments even though a naive rabbit from somewhere else would run in terror from the sound of a train. Yet another kind of social learning has been called "goal emulation", when one animal copies the goals or outcomes of another animal's behavior but without copying the form of that behavior itself. An example might be when an ape sees another ape getting food from a container and then uses a different method of its own invention for getting at that food. We can now see that none of these processes is true imitation because no new behaviors are copied from one animal to another (for reviews of social learning and imitation see Heyes and Galef, 1996; Whiten and Ham, 1992; Zentall and Galef, 1988). The behavior of one animal comes to be similar to that of another animal, but not by copying it.

Other famous examples that look like true cultural learning based on imitation include the troop of Japanese macaques that learned to wash sweet-potatoes, and chimpanzees that learned how to fish for termites by poking sticks into the mounds. However, both of these appear to depend on individual learning and the kinds of social learning described above, not on true imitation (Galef, 1992). So if you want to stick to the definition of memes as transmitted by imitation then you have to say that bottle-top pecking, termite-fishing and potato-washing are probably not memes.

## 5. True Imitation

The comparison with other forms of social learning raises the question whether true imitation occurs at all in non-human animals.

### 5.1 Vocal imitation in birds and dolphins

There is no doubt that there are examples of vocal imitation in birds and cetaceans. Song birds have been treated as a special case since research on imitation first began a hundred years ago (Bonner, 1980; Delius, 1989; Thorndike, 1898; Whiten and Ham, 1992). This is partly because imitation in birds is generally confined to sounds, and to rather specific kinds of sound at that (with the possible exception of parrots who may be able to imitate simple gestures). Many songbirds have long traditions. The young learn what to sing by imitating their parents or neighbors. In chaffinches, for example, the nestling may hear its father sing long before it is capable of singing itself. A few months later it begins to make a wide variety of sounds, gradually narrowing down to the song it heard as a chick. Experiments show that there is a critical period for learning and that the bird has to hear its own song and match it to the remembered song it is imitating. Hand-raised birds can learn songs from tape recorders and adopted birds sing songs more like their adopted, not biological, parents. Some species learn many songs from neighbors and a few, like parrots and mynahs, can imitate human speech.

Dolphins can also copy vocalizations, and young dolphins produce a wide variety of sounds which they later cut down. Captive bottlenose dolphins have been shown to easily imitate artificial signals relating to specific objects and use these new signals in spontaneous play (McCowan & Reiss, 1997). So we can count dolphin whistles and bird songs as memes, and indeed the cultural evolution of chaffinch song has been studied in terms of the mutation, flow

and drift of song memes (Lynch, Plunkett, Baker and Jenkins, 1989). These very specific kinds of imitation are therefore unlike the examples of social learning we were considering before.

## **5.2 Imitation in humans and other animals**

As for other species, the picture is unclear. There have been claims of imitation (other than vocalization) in parrots, budgerigars, pigeons and rats, though all the claims have also been disputed (see Heyes & Galef, 1996). Dolphins can apparently mimic familiar behaviors but have not so far been shown to copy novel behaviors (Bauer & Johnson, 1994). Chimpanzees and gorillas that have been brought up in human families occasionally imitate in ways that their wild counterparts do not (Tomasello, Kruger & Ratner, 1993). However, when apes and human children are given the same problems, only the children readily use imitation to solve them (Call and Tomasello, 1995).

Humans, therefore, seem to differ considerably from all other species. They are "*the consummate imitative generalist*" (Meltzoff, 1988, p 59). Human infants are able to imitate a wide range of vocal sounds, body postures, actions on objects, and even completely arbitrary actions like bending down to touch your head on a plastic panel. By 14 months of age they can even delay imitation for a week or more (Meltzoff, 1988), and they seem to know when they are being imitated by adults (Meltzoff, 1996). Unlike any other animals we readily imitate almost everything and anything.

If we define memes as transmitted by imitation then we must conclude that only humans are capable of extensive memetic transmission. Some other theorists have included all forms of social learning in their definitions of cultural evolution (e.g. Boyd and Richerson, 1985; Delius, 1989; Plotkin, 1996) and their mathematical models may usefully apply to all. However, I suggest that it will be better for memetics to stick to the original definition of memes.

## **5.3 Only imitation sustains a true evolutionary process.**

One might argue that both social learning and imitation allow information about behavior to be transmitted and that the difference is only one of fidelity, longevity or fecundity. Indeed Heyes (1994) does just this, arguing that the difference lies primarily in longevity. However, I suggest that the other forms of social learning do not support a replication system with true heredity.

Although new behaviors can be passed on by other kinds of social learning, the process is cumbersome. For example, one animal must invent a new behavior during individual learning and then somehow lead a second animal into such a situation that it is likely to learn the same new behavior - or perhaps the first can behave in such a way as to change the contingencies of learning for the second animal so that it learns the same (or a similar) new behavior. Most importantly, in these cases, the behavior must be created anew each time by the learner. The social situation and the behavior of the other animal plays a role, but the details of the first behavior are not transmitted and therefore cannot be built upon and refined by further selective copying. In this sense, then, there is no true heredity. This means there is no new replicator, no true evolution, and therefore the process should not be considered as memetic.

By contrast, the skill of generalized imitation means that humans can invent new behaviors of almost unlimited kinds and pass them on to each other by a kind of copying. If we define memes as transmitted by imitation then whatever is passed on by this copying process is a meme. Memes fulfill the role of replicator because they exhibit all three of the necessary conditions; that is, heredity (the form and details of the behavior are copied), variation (they are copied with errors, embellishments or other variations), and selection (only some behaviors are successfully copied). This is a true evolutionary process.

## 6. Human Cultural Learning

We learn about our culture in many ways, including reading and writing, watching television, being deliberately taught by parents and school teachers, and by listening to the conversations of others. In any consideration of memetics, from its origins in Dawkins's work, right through to the present, we count as memes all of the cultural behaviors passed on in these various ways, including everything from fashions and habits, to political ideologies and scientific theories. We would be daft to redefine the meme in such a way that any of these was excluded, but this naturally raises the question of whether all these forms of learning can really be counted as imitation.

There have been many attempts to classify the kinds of learning that underlie human cultural transmission. For example Tomasello et al (1993) describe three forms of cultural learning; imitative learning, instructed learning, and collaborative learning. They argue that all these involve some kind of inter-subjectivity or perspective taking, with increasingly complex kinds of social-cognitive concepts and processes involved. However, some prefer not to speculate at all about the intentionality or hidden cognitive processes involved in cultural learning (Zentall, 1996) while others take up the substantial questions of just what kinds of perspective-taking, intentionality, mind-reading, or other complex cognitive processes are involved in just which kinds of learning and teaching (see the extensive commentaries following Tomasello et al, 1993). There is clearly no consensus here yet, and it will not help us with defining the meme to launch into these tricky issues, even though in the future of memetics these issues may be very important.

Heyes (1994) takes an entirely different route and argues that human cultural learning is different from animals' because it involves instruction and not imitation. However, this distinction will not work, partly because animals do on occasion use instruction without imitation (as in the examples of manipulating contingencies described above) and partly because human instruction often entails imitation (as in learning to write or cook). Heyes further argues that the creation of language and artifacts decouples cultural accumulation from the process of imitation, but this means denying an important role to imitation in either language or the creation of books, art, buildings and ideologies.

We may take a simpler position - that all these kinds of learning and teaching require at least the ability to imitate. Language learning is a good example. Although there are many arguments about just how much of language depends on having an innate grammar module (Pinker, 1994), there is no doubt that human language learning involves the imitation of sounds. Chinese children do not spontaneously start using French words, and German children do not suddenly

burst into Hindi. The sounds of words are acquired by imitating others. Reading and writing are also learned, at least to some extent, by imitation, as when the shape of a letter is meticulously copied.

When we read a story and then tell it to someone else is this imitation? I would say that it is. The skills involved may be far more complex than the kinds of imitation I have described above, but they have a basis in imitation and are of the same general form. Something about the story is internalized in the listener and then reproduced when she or he tells the story again. The same can be said of passing on religious or scientific ideas - the reader or hearer of the ideas must internalize them in some way and then reproduce them for another reader or listener.

I have no doubt that in the future memetics will become involved in discovering just what the cognitive processes are that are involved in teaching, learning, instruction, and any other kinds of cultural transmission, but for the moment I suggest that we may consider all of them as in some way being forms of imitation, or based on the ability to imitate. I assume that this is what Dawkins meant by "... *a process which, in the broad sense, can be called imitation*". He meant to include reading, writing, conversation, and academic study as "*in the broad sense... imitation*" and I think we should have no hesitation in continuing to do so.

## 7. Conclusion

My argument has been that the definition of the meme depends on, and should depend on, the concept of imitation. Therefore, only those things that can be passed on by imitation should count as memes.

This means we can immediately exclude many things that a few authors have confusingly included as memes, such as perceptions, emotional states, cognitive maps, experiences in general, or "*anything that can be the subject of an instant of experience*". Furthermore we can build on the long history of research in animal behavior to distinguish imitation from contagion, and from individual and social learning, and so to eliminate from memetics the catching of yawns or all the many things we each learn for ourselves, by ourselves.

This, I suggest, leaves us with a simple definition of the meme that not only makes it easy to decide what is and is not a meme, but also shows why it is that humans alone have produced complex culture. Humans are fundamentally unique not because they are especially clever, not just because they have big brains or language, but because they are capable of extensive and generalized imitation. I think we will discover that it is imitation that gave rise to our cleverness, big brains and language - and it is imitation that makes culture possible, for only imitation gives rise to a new replicator that can propagate from brain to brain, or from brain to artifact and back to brain. For all these reasons I suggest that we stick with the dictionary, and define the meme as that which is passed on by imitation.

## References

Baldwin, J.M. (1902). *Development and Evolution*. New York, Macmillan.

Bauer, G.B. and Johnson, C.M. (1994). Trained motor imitation by bottlenose dolphins (*Tursiops Truncatus*). *Perceptual and Motor Skills*, 79, 1307-1315

Bonner, J.T. (1980). *The Evolution of Culture in Animals*. Princeton, NJ, Princeton University Press.

Boyd, R. and Richerson, P.J. (1985). *Culture and the Evolutionary Process*. Chicago, University of Chicago Press.

Brodie, R. (1996). *Viruses of the Mind: The New Science of the Meme*. Seattle, Integral Press.

Call, J. and Tomasello, M. (1995). Use of social information in the problem solving of orangutans (*Pongo pygmaeus*) and human children (*Homo sapiens*). *Journal of Comparative Psychology*, 109, 308-320

Calvin, W. (1987). The brain as a Darwin machine. *Nature*, 330, 33-44

Dawkins, R. (1976). *The Selfish Gene*. Oxford, Oxford University Press (new edition with additional material, 1989)

Delius, J. (1989). Of mind memes and brain bugs, a natural history of culture. In W.A. Koch (Ed), *The Nature of Culture*. Bochum, Germany, Bochum Publications. 26-79

Edelman, G.M. (1989). *Neural Darwinism: The Theory of Neuronal Group Selection*. Oxford, Oxford University Press.

Fisher, J. and Hinde, R.A. (1949). The opening of milk bottles by birds. *British Birds*, 42, 347-357.

Gabora,L. (1997). The origin and evolution of culture and creativity. *Journal of Memetics - Evolutionary Models of Information Transmission*, 1. [http://jom-emit.cfpm.org/1997/vol1/gabora\\_1.html](http://jom-emit.cfpm.org/1997/vol1/gabora_1.html)

Galef,B.G. (1992) The question of animal culture. *Human Nature*, 3, 157-178

Heyes ,C.M. (1993) Imitation, culture and cognition. *Animal Behaviour*, 46, 999-1010

Heyes ,C.M. (1994) Imitation and culture: Longevity, fecundity and fidelity in social transmission. In B.G.Galef, M.Mainardi and P.Valsecchi (Eds) *Behavioural Aspects of Feeding*. Chur, Switzerland, Harwood. 271-287

Heyes ,C.M. and Galef ,B.G. (Eds) (1996). *Social Learning in Animals: The roots of culture*. San Diego, CA, Academic Press.

Levy ,D.A. and Nail ,P.R. (1993). Contagion: A theoretical and empirical review and reconceptualization. *Genetic Social and General Psychology Monographs*. 119, 235-284

- Lynch, A. (1996). *Thought Contagion: How Belief Spreads Through Society*. N.Y. Basic Books.
- Lynch, A., Plunkett,G.M., Baker,A.J. and Jenkins,P.F. (1989). A model of cultural evolution of chaffinch song derived with the meme concept. *The American Naturalist*, 133, 634-653
- Marsden, P. (1997). Crash contagion and the death of Diana: Memetics as a new paradigm for understanding mass behaviour. Paper presented at the 'Death of Diana' conference. University of Sussex. 14 November 1997.
- McCowan, B and Reiss, D. (1997) Vocal learning in captive bottlenose dolphins: comparison with humans and nonhuman animals. In C.T.Snowdon and M.Hausberger (Eds) *Social Influences on Vocal Development*. Cambridge, Cambridge University Press, 178-207
- Meltzoff,A.N. (1988) Imitation, objects, tools, and the rudiments of language in human ontogeny. *Human Evolution*, 3, 45-64
- Meltzoff, A.N. (1996) The human infant as imitative generalist: A 20-year progress report on infant imitation with implications for comparative psychology. In C.M. Heyes and B.G. Galef (Eds) *Social Learning in Animals: The roots of culture*. San Diego, CA, Academic Press. 347-370.
- Pinker, S. (1994). *The Language Instinct*. New York, Morrow
- Plotkin, H. (1993). *Darwin Machines and the Nature of Knowledge*. Cambridge, Mass., Harvard University Press.
- Plotkin, H.C. (1996). Non-genetic transmission of information: Candidate cognitive processes and the evolution of culture. *Behavioural Processes*, 35, 207-213
- Provine, R.R. (1996). Contagious yawning and laughter: Significance for sensory feature detection, motor pattern generation, imitation, and the evolution of social behaviour. In Heyes,C.M. and Galef,B.G. (Eds) *Social Learning in Animals: The roots of culture*. San Diego, CA, Academic Press, 179-208.
- Romanes, G.J. (1882). *Animal Intelligence*. London, Kegan Paul Trench
- Romanes, G.J. (1883). *Mental Evolution in Animals*. London, Kegan Paul Trench
- Sherry, D.F. and Galef, B.G. (1984). Cultural transmission without imitation: milk bottle opening by birds. *Animal Behaviour*, 32, 937-938
- Showalter, E. (1997). *Hystories: Hysterical Epidemics and Modern Culture*. New York, Columbia University Press
- Skinner, B.F. (1953). *Science and Human Behaviour*. New York, Macmillan.

Thorndike, E.L. (1898). Animal intelligence: An experimental study of the associative processes in animals. *Psychological Review Monographs*, 2, No 8.

Tomasello, M., Kruger, A.C. and Ratner, H.H. (1993). Cultural Learning. *Behavioral and Brain Sciences*, 16, 495-552.

Whiten, A. and Ham, R. (1992). On the nature and evolution of imitation in the animal kingdom: Reappraisal of a century of research. In *Advances in the Study of Behavior* Vol. 21. Ed. Slater,P.J.B., Rosenblatt,J.S., Beer,C. and Milinski,M. San Diego, CA, Academic Press.

Wilkins, J.S. (1998). What's in a meme? Reflections from the perspective of the history and philosophy of evolutionary biology. *Journal of Memetics - Evolutionary Models of Information Transmission*, 2. [http://jom-emit.cfpm.org/1998/vol2/wilkins\\_js.html](http://jom-emit.cfpm.org/1998/vol2/wilkins_js.html)

Williams, G.C. (1966). *Adaptation and Natural Selection*. Princeton, Princeton University Press.

Zentall, T.R. (1996). An analysis of imitative learning in animals. In. Heyes,C.M. and

Galef, B.G. (Eds) *Social Learning in Animals: The roots of culture*. San Diego, CA, Academic Press, 221-243

Zentall, T.R. and Galef, B.G. (Eds) (1988). *Social Learning: Psychological and Biological Perspectives*. Hillsdale, N.J., Erlbaum.

## THE FORGET-MEME-NOT THEORY

By  
**Susan Blackmore**

This piece was published in the Times Higher Education Supplement 26th February 1999, with responses to it published 5th March 1999 (p19)

*Why do some ideas evolve through history while others are lost in the annals of time? Memes, of course. Susan Blackmore outlines a theory that has split the academic world.*

We humans are odd creatures. Like other animals we are born and die, and enjoy food and sex, but unlike other species we wear clothes and cook our food, watch movies, go hang gliding, and read newspapers. Why? Is it just that we are more intelligent, or more conscious?

A completely new answer is provided by the theory of memes. Memes are ideas, skills, habits, stories or inventions, which are passed from person to person by imitation. Like genes they compete to get copied, but unlike genes their competition is for space in our memories, and for the chance to get into books, magazines and television programs. The survivors in this game are the ones we see all around us. Just as genes have created our bodies, so memes have created our minds and our cultures.

This explains, so the theory goes, our incurably religious nature, our unusual forms of cooperation and altruism, our use of language, and our ability to defy our genes with birth control and genetic engineering. We humans, alone on this planet, are meme machines. The term 'meme' (to rhyme with cream or dream) was coined in 1976 by Richard Dawkins, in *The Selfish Gene*. The purpose of his book was to explain the power and generality of Darwin's great insight. What Darwin had realized was that a simple mindless process can account for evolution---without a designer. If you have creatures that vary, and if only some of them can survive, and if the survivors pass on to their offspring whatever it was that helped them survive, then the next generation must be better adapted than the first---and so the process goes on. In more modern terms, if you have variation, heredity and selection, then you must get, as philosopher Dan Dennett puts it "Design out of Chaos without the aid of Mind". And this inevitable process works on anything that is copied---not just genes.

To push home his point, Dawkins wanted another example, and so he invented the meme. He included "tunes, ideas, catch-phrases, clothes fashions, ways of making pots or of building arches." He mentioned scientific ideas and religions, fashions in dress or diet, ceremonies, customs and technologies---all of which are spread by one person copying another. Memes are inherited every time they are passed on. They vary all the time, as when you tell the story slightly differently, or mix up two ideas to produce a new one. And they undergo enormous selection pressure. Just think how many ideas you have come across in this newspaper. Only a tiny fraction will survive in your memory, and even fewer will be passed on again to someone else. These successful memes are the survivors, and are the ones that shape our minds and cultures.

If we could understand what makes for a successful meme, then we would have a whole new way of understanding our minds. This is the promise that memetics holds out---but can it really work?

At the heart of memetics is the idea that memes compete with genes, causing effects that could not be predicted by biology alone. This is nowhere more obvious than in the arena of sex. Sex sells. People pay attention, get aroused, and pay money for sex. The genes made us that way, for their own propagation. But this gives sex memes a boost. They get all over the place, and some of them do not benefit our genes. Take birth control. Pills and condoms and coils thrive because they promise all the pleasure of sex without the responsibility of childbirth. They get an unfair advantage too, by a simple bit of logic.

Compare two women---one with one child and one with six. Which of them is likely to spread more of her memes? The woman with one child has more time to meet people and chat, more time for a career, more time to become a politician, television presenter, or writer. So her memes spread far and wide, including those for birth control and a small family. This simple inequality helps the memes along.

Other gene-challenging memes include celibacy, martyrdom and, most recently, genetic engineering. We all appreciate the threat of genetic modification getting out of control but memetics allows us to see just how and why that can happen. If the memes can spread they will, and given human need and greed they will.

One of the strengths of memetics is to explain not only why we like good ideas, but why false and even harmful ones have such power over us--- because they are good replicators. Take religions. Dawkins has made himself extremely unpopular in some circles by calling religions "viruses of the mind". They sneak into our minds, avoid our memetic defenses, and get themselves propagated all around the world, even though their claims are false. Millions of people believe in God, the afterlife, the transubstantiation, and the virgin birth, without a shred of evidence. Why? Because these memes use clever tricks to get themselves copied, just like viruses do. Believers, once infected, will fight for their beliefs, build great cathedrals to inspire others, bring up their children to fear God and go to church, and spend inordinate amounts of time singing, praying, and reading holy books---in other words, passing on the memes. And if they start to doubt, they will be told that the faithful go to heaven---an invisible place you cannot visit until you are dead. No wonder religions spread.

Alternative therapies do too. You visit the homeopath who gives you a long explanation and some water to take. Like everyone else you are susceptible to the placebo effect, and feel better from the attention. So you keep taking the water and are convinced by the homeopathy memes as well. Or you consult your stars, or a Tarot reader, or a psychic. The psychic homes in on your personal troubles using feedback and subtle body cues. If you are lucky she is gifted and empathic and can really help. So you go away, not just feeling better and poorer, but more convinced by her memes and likely to tell your friends.

Like these superstitions, science itself is a system of memes, and inevitably includes false memes and viruses, but the difference is this. Right at the heart of science lies the principle of testing all

new memes against the facts. If they fail they are thrown out. This is the ultimate arbiter against which memetics itself will be judged. Unlike God, if it helps us explain and predict human nature, then it will thrive. Otherwise---out it goes.

But we had better start finding out soon. For the memes are rapidly taking off. We might think we humans designed all those computers and phone links for our own pleasure, but from the meme's-eye-view we are just their copying machines, and they are using us to design a vast planet-wide system for their own propagation. The mechanism is frighteningly simple. Like genes, memes have no foresight and no plans, they just multiply when they can. Memes that get onto computers are successful and easily copied. Memes that jump from city to city and continent to continent, in multiple accurate copies, do even better, and take with them the idea of the Internet itself. We can't stop this evolutionary process and we are not, and never have been, in control. We may have been the first meme machines on this planet, but we won't be the last.

## What is a Meme?

"meme /mi:m/ n. L20.(f. Gk mimema. that which is imitated, after GENE.) Biol. An element of a culture or system of behaviour that may be considered to be passed from one individual to another by non-genetic means, esp. imitation". (Shorter Oxford English Dictionary) Derivatives: memetic a., memetics n. the science of memes. memplex n. abbreviated from 'co-adapted meme complex'; a group of memes passed on together, e.g. religions, political ideologies and belief systems. meme pool n. all the memes in a population at a particular time. memetic engineering n. manipulating memes, as in psychotherapy, advertising or education.

Before 'meme' entered the OED in 1997 there were many other definitions, from the unhelpfully encompassing "anything that can be the subject of an instant of experience" to the awkwardly narrow "constellations of activated and non-activated synapses within neural memory networks". But the basic principle of memes is that people copy each other, and copying lets loose a new evolutionary process. From this perspective the best definition must be "A meme is that which is passed on by imitation".

## The Tricks Memes Play

*Dear Friend, pass on this letter and you will have good luck for a year. Have you ever fallen for a chain letter? Perhaps you got one urging you to send money or postcards or stamps, or the more sinister---Pass this on to ten of your friends or you will fall ill, die, and suffer a fate worse than death...?*

These memes succeed because they use crude tricks to get themselves copied, together with threats or promises. Their modern versions are winging their way around the Internet, using up valuable resources for their own selfish propagation. A long-running warning from "Pen pal Greetings" begins "received this morning from IBM". It tells you (at great length) to warn all your friends not to open the dangerous "Trojan Horse" virus. So you pass on the friendly warning, not realizing, unless you are very wily, that the warning itself is the virus---this is what gets into millions of e-mail boxes all over the world.

Fortunately most memes succeed because we want to copy them, from useful inventions like cars and central heating to beautiful paintings, stories and music. There are languages and customs, political ideologies and scientific theories. All these count as memes because they are copied from person to person. But what we must remember is that memes succeed because we copy them---some for good reasons---others not.

## **What the Experts Say about Memetics**

"I think memes provide a neat way of explaining some of the paradoxes of cultural transmission and, if they help with explaining something that important, there had better be a science of them." Nicholas Humphrey, senior research fellow, London School of Economics

"The idea of memes is a meaningless metaphor". Stephen Jay Gould, professor of zoology, Harvard University

"A surprising number of people these days even talk about 'memes' ... Absurd as it may seem to imagine the seamless web of culture being disaggregated and transmitted between minds by gene-like replication, even some sociologists and philosophers have become captivated by the metaphor." Steven Rose, biologist, Open University

"It is an empty and misleading metaphor to call religion, science, or any other human activity a 'virus' or 'parasite'". Memes are a "useless and essentially superstitious notion." Mary Midgley, philosopher, Newcastle.

"Memetics needs to come up with supported, unique predictions and/or an existence proof to become valuable. The challenge to our speakers is to provide some support---either theoretical or empirical---for the meme hypothesis." Robert Aunger, anthropologist, Cambridge.

*Susan Blackmore is Senior Lecturer in Psychology at the University of the West of England, Bristol, and author of *The Meme Machine*, Oxford University Press. The first academic conference on memes will be held at King's College Cambridge, June 3 and 4. Speakers include Dan Sperber, Dan Dennett, John Maynard Smith and Susan Blackmore*

# THE POWER OF THE MEME MEME

By  
**Susan Blackmore**

Department of Psychology  
University of the West of England  
Bristol BS16 2JP

For *Skeptic* June 1997, Vol 5, #2, pp43--49

Without the theory of evolution by natural selection nothing in the world of biology makes much sense. Without Darwin and neo-Darwinism, you cannot answer questions like “Why do bats have wings? Why do cats have five claws? or Why do our optic fibres cross in front of our retinas?” You can only fall back on appeals to an imaginary creator.

I am going to make a bold claim.

Without the theory of evolution by memetic selection nothing in the world of the mind makes much sense. Without memetics you cannot answer questions like “Why can’t I get that thought out of my mind? Why did I decide to write this article and not that one? Who am I?” Without memetics you can only fall back on appeals to an imaginary conscious agent.

In this article I want to lay the groundwork for a theory of memetics and see how far we can get. I shall outline the history and origins of the idea, explore how it has been used, abused, and ignored, and how it has provided new insight into the power of religions and cults. I shall then take on a meme’s eye view of the world and use this to answer five previously unanswered questions about human nature. Why can’t we stop thinking? Why do we talk so much? Why are we so nice to each other? Why are our brains so big? And, finally, what is a self?

I have tried to write the sections to stand alone. If you only want to read some of them I suggest you read the section Taking the meme’s eye view, and pick any others that take your fancy.

## A History of the Meme Meme

In 1976 Dawkins published his best-selling *The Selfish Gene*. This book popularised the growing view in biology that natural selection proceeds not in the interest of the species or of the group, nor even of the individual, but in the interest of the genes. Although selection takes place largely at the individual level, the genes are the true replicators and it is their competition that drives the evolution of biological design.

Dawkins, clear and daring as always, suggested that all life everywhere in the universe must evolve by the differential survival of slightly inaccurate self-replicating entities; he called these “replicators”. Furthermore, these replicators automatically band together in to groups to create systems, or machines, that carry them around and work to favour their continued replication. These survival machines, or “vehicles” are our familiar bodies - and those of cats, e-coli and cabbages - created to carry around and protect the genes inside them.

Right at the end of the book he suggests that Darwinism is too big a theory to be confined to the narrow context of the gene. So he asks an obvious, if provocative, question. Are there any other replicators on our planet? Yes, he claims. Staring us in the face, though still drifting clumsily about in its primeval soup of culture, is another replicator - a unit of imitation. He gave it the name "meme" (to rhyme with "dream" or "seem") and as examples suggested "tunes, ideas, catch-phrases, clothes fashions, ways of making pots or of building arches." Memes are stored in human brains and passed on by imitation.

In just those few pages he laid the foundations for understanding the evolution of memes. He discussed their propagation by jumping from brain to brain, likened them to parasites infecting a host, treated them as physically realised living structures, and showed how mutually assisting memes will group together just as genes do. He argued that once a new replicator arises it will tend to take over and begin a new kind of evolution. Above all he treated memes as replicators in their own right, chastising those of his colleagues who tended always to go back to "biological advantage" to answer questions about human behaviour. Yes, he agreed, we got our brains for biological (genetic) reasons but now we have them a new replicator has been unleashed and it need not be subservient to the old. In other words, memetic evolution can now proceed without regard to its effects on the genes.

A few years later Douglas Hofstadter wrote about viral sentences and self-replicating structures in his Scientific American column Metamagical Themas. Readers replied, with examples of text using bait and hooks to ensure its own replication. They suggested viral sentences from the simplest instruction, such as "Copy me!", through those with added threats ("Say me or I'll put a curse on you") or promises ("I'll grant you three wishes"), to examples of virulent chain letters (Hofstadter, 1985, p 53). One reader suggested the term memetics for the discipline studying memes. Yet memetics did not really take off.

Why not? The basic idea is very simple. If Dawkins is right then everything you have learned by imitation from someone else is a meme. This includes all the words in your vocabulary, the stories you know, the skills and habits you have picked up from others and the games you like to play. It includes the songs you sing and the rules you obey. So, for example, whenever you drive on the right (and I on the left!), eat a hamburger or a pizza, whistle "Happy Birthday to You" or "Mama I love you" or even shake hands, you are dealing in memes. Memetics is all about why some memes spread and others do not.

The greatest proponent of memetics since Dawkins has been the philosopher Dan Dennett. In his books Consciousness Explained (1991) and Darwin's Dangerous Idea (1995) he expands on the idea of the meme as replicator.

In The Origin of Species, Darwin (1859) explained how natural selection must happen if certain conditions are met. If there is heredity from parent to offspring, variation among the offspring, and not all the offspring can survive - then selection must happen. Individuals who have some useful advantage "have the best chance of being preserved in the struggle for life" (Darwin, 1859, p 127, and see Dennett, 1995, p 48) and will then pass on this advantage to their offspring. Darwin clearly saw how obvious the process of natural selection is once you have grasped it. It just must happen.

Dennett describes evolution as a simple algorithm - that is, a mindless procedure that when carried out must produce a result. For evolution you need three things - heredity, variation and selection - then evolution is inevitable. You need not get us, of course, or anything remotely like us; for evolution has no plans and no foresight. Nevertheless, you must get something more complex than what you started with. The evolutionary algorithm is “a scheme for creating Design out of Chaos without the aid of Mind” (Dennett, 1995, p 50). This, says Dennett, is Darwin’s Dangerous Idea.

No wonder people have been terrified of it, and fought so hard against it. It is outrageously simple and terrifyingly powerful.

If evolution is an algorithm then it should be able to run on different substrates. We tend to think of evolution as depending on genes because that is the way biology works on this planet, but the algorithm is neutral about this and will run wherever there is heredity, variation and selection. Or -as Dawkins puts it - a replicator. It doesn’t matter which replicator. If memes are replicators then evolution will occur.

## **So are memes replicators?**

There is enormous variety in the behaviours human beings emit, these behaviours are copied, more or less accurately by other human beings, and not all the copies survive. The meme therefore fits perfectly with the scheme of heredity, variation and selection. Think of tunes, for example. Millions of variants are sung by millions of people. Only a few get passed on and repeated and even fewer make it into the pop charts or the collections of classics. Scientific papers proliferate but only a few get long listings in the citation indexes. Only a few of the disgusting concoctions made in woks actually make it onto the TV shows that tell you how to Wok things and only a few of my brilliant ideas have ever been appreciated by anyone! In other words, competition to get copied is fierce.

Of course memes are not like genes in many ways and we must be very careful in applying terms from genetics to memes. The copying of memes is done by a kind of “reverse engineering” by one person copying another’s behaviour, rather than by chemical transcription. Also we do not know just how memes are stored in human brains and whether they will turn out to be digitally stored, like genes, or not. However, the important point is that if memes are true replicators, memetic evolution must occur.

Dennett is convinced they are and he explores how memes compete to get into as many minds as possible. This competition is the selective force of the memosphere and the successful memes create human minds as they go, restructuring our brains to make them ever better havens for more memes. Human consciousness, claims Dennett, is itself a huge meme-complex, and a person is best understood as a certain sort of ape infested with memes. If he is right then we cannot hope to understand the origins of the human mind without memetics.

This makes it all the more fascinating that most people interested in the human mind have ignored memetics or simply failed to understand it. Mary Midgley (1994) calls memes “mythical entities” that cannot have interests of their own; “an empty and misleading metaphor”. In a

recent radio debate, Stephen Jay Gould called the idea of memes a “meaningless metaphor” (though I am not sure one can actually have a meaningless metaphor!). He wishes “that the term “cultural evolution” would drop from use.” (Gould, 1996, p 219-20).

The word “Meme” does not even appear in the index of important books about human origins and language (e.g. Donald, 1991; Dunbar, 1996; Mithen, 1996; Pinker, 1994; Tudge, 1995; Wills, 1993), in an excellent collection on evolutionary psychology (Barkow, Cosmides and Tooby, 1992), nor in books about human morality (Ridley, 1996; Wright, 1994). Although there are many theories of the evolution of culture, almost all make culture entirely subservient to genetic fitness, as in Wilson’s (1978) metaphor of the genes holding culture on a leash or Lumsden and Wilson’s claim that “the link between genes and culture cannot be severed” (1981, p 344). Cavalli-Sforza and Feldman (1981) treat “cultural activity as an extension of Darwinian fitness” (p 362) and even Durham (1991), the only one to use the word “meme”, sticks to examples of cultural features with obvious relevance to genetic fitness, such as color naming, dietary habits and marriage customs. Perhaps Boyd and Richerson (1990) come closest to treating the cultural unit as a true replicator. However, they still view “genetic and cultural evolution as a tightly coupled coevolutionary process in humans” (Richerson & Boyd, 1992, p 80).

As far as I can understand them, no one except Cloak (1975) and Dawkins treats their unit of cultural exchange as a true replicator. If there is a continuum from Gould’s outright rejection at one end, to Dawkins and Cloak at the other, then most lie in between. They accept cultural evolution but not the idea of a second replicator. When they say “adaptive” or “maladaptive” they mean for the genes. When it comes to the crunch they always fall back on appeals to biological advantage, just as Dawkins complained that his colleagues did twenty years ago. Dawkins is clear on this issue when he says “there is no reason why success in a meme should have any connection whatever with genetic success”. I agree. I am going to propose a theory of memetics that lies at the far end of this continuum. I suggest that once genetic evolution had created creatures that were capable of imitating each other, a second replicator was born. Since then our brains and minds have been the product of two replicators, not one. Today many of the selection pressures on memes are still of genetic origin (such as whom we find sexy and what food tastes good) but as memetic evolution proceeds faster and faster, our minds are increasingly the product of memes, not genes. If memetics is true then the memes have created human minds and culture just as surely as the genes have created human bodies.

## **Religions as Co-Adapted Meme-Complexes**

Dawkins (1976) introduced the term co-adapted meme-complex. By this he meant a group of memes that thrive in each others’ company. Just as genes group together for mutual protection, leading ultimately to the creation of organisms, so we might expect memes to group together. As Dawkins (1993) puts it “there will be a ganging up of ideas that flourish in one another’s presence”.

Meme-complexes include all those groups of memes that tend to be passed on together, such as political ideologies, religious beliefs, scientific theories and paradigms, artistic movements, and languages. The most successful of these are not just loose agglomerations of compatible ideas,

but well structured groups with different memes specialising as hooks, bait, threats, and immune system. (Memetic jargon is still evolving and these terms may change but see Grant's "memetic lexicon" (Grant, 1990)).

When I was about ten years old I received a post card and a letter that contained a list of six names and instructed me to send a post card to the first name on the list. I was to put my own name and address at the bottom and send the new list to six more people. It promised me I would receive lots of postcards.

This was a fairly innocuous chain letter as these things go, consisting just of a bait (the promised postcards) and a hook (send it to six more people). Threats are also common (send this on or the evil eye will get you) and many have far worse consequences than a waste of stamps. What they have in common is the instruction to "duplicate me" (the hook) along with co-memes for coercion. These simple little groups can spread quite well.

With the advent of computers viral meme-groups have much more space to play in and can leap from disk to disk among "unhygienic" computer users. Dawkins (1993) discusses how computer viruses and worms use tricks to get themselves spread. Some bury themselves in memory only to pop up as a time bomb; some infect only a small proportion of those they reach, and some are triggered probabilistically. Like biological viruses they must not kill their host too soon or they will die out. Their final effect may be quite funny, such as one that makes the Mackintosh's loudspeaker say "Don't Panic!", but some have clogged up entire networks and destroyed whole doctoral theses. My students have recently encountered a virus in WORD6 that lives in a formatting section called "Thesis" - tempting you to get infected just when your year's work is almost finished. No wonder we now have a proliferation of anti-virus software - the equivalent of medicine for the info-sphere.

Internet viruses are a relatively new arrival. Last week I received a very kind warning from someone I've never met. "Do not download any message entitled "Penpal Greetings"" it said - and went on to warn me that if I read this terrible message I would have let in a "Trojan Horse" virus that would destroy everything on my hard drive and then send itself on to every e-mail address in my mail box. To protect all my friends, and the world-wide computer network, I had to act fast and send the warning on to them.

Have you spotted it? The virus described does not make sense - and does not exist. The real virus is the warning. This is a very clever little meme-complex that uses both threats and appeals to altruism to get you - the silly, caring victim - to pass it on. It is not the first - "Good Times" and "Deeyenda Maddick" used a similar trick - and it probably won't be the last. However, as more people learn to ignore the warnings these viruses will start to fail and perhaps that will let in worse viruses, as people start to ignore warnings they ought to heed. So Watch Out! What does this have to do with religions? According to Dawkins, a great deal. The most controversial application of memetics is undoubtedly his treatment of religions as co-adapted meme-complexes (Dawkins 1976, 1993). He unashamedly describes religions as "viruses of the mind" and sets about analysing how they work.

They work because human brains are just what info-viruses need; brains can soak up information, replicate it reasonably accurately, and obey the instructions it embodies. Dawkins uses the example of Roman Catholicism; a gang of mutually compatible memes that is stable enough to deserve a name. The heart of Catholicism is its major beliefs; a powerful and forgiving God, Jesus his son who was born of a virgin and rose again from the dead, the holy spirit, and so on. If these aren't implausible enough you can add belief in miracles or the literal transubstantiation of wine into blood. Why should anyone believe these things? Dawkins explains.

Threats of hell-fire and damnation are an effective and nasty technique of persuasion. From an early age children are brought up by their Catholic parents to believe that if they break certain rules they will burn in hell forever after death. The children cannot easily test this since neither hell nor God can be seen, although He can see everything they do. So they must simply live in life-long fear until death, when they will find out for sure - or not! The idea of hell is thus a self-perpetuating meme.

And did I say "test" the idea? Some religious beliefs could be tested, such as whether wine really turns into blood, or whether prayer actually helps; hence the need for the anti-testing meme of faith. In Catholicism, doubt must be resisted, while faith is nurtured and respected. If your knowledge of biology leads you to doubt the virgin birth, - or if war, cruelty and starvation seem to challenge the goodness of God - then you must have faith. The story of Doubting Thomas is a cautionary tale against seeking evidence. As Dawkins puts it "Nothing is more lethal for certain kinds of meme than a tendency to look for evidence" (Dawkins, 1976, p 198) and religions, unlike science, make sure they discourage it. Also unlike science, religions often include memes that make their carriers violently intolerant of new and unfamiliar ideas so protecting themselves against being ousted in favour of a different religion - or none at all.

Finally the meme-complex needs mechanisms to ensure its own spread. "Kill the infidel" will dispose of the opposition. "Go forth and multiply" will produce more children to pass itself onto. So will forbidding masturbation, birth control or inter-faith marriages. If fear of going blind doesn't work, there are prizes in heaven for missionaries and those who convert unbelievers (Dawkins, 1993; Lynch, 1996).

Catholicism generally spreads from parent to child but celibate priests play a role too. This is particularly interesting since celibacy means a dead end for the genes, but not for the memes. A priest who has no wife and children to care for has more time to spread his memes, including that for celibacy. Celibacy is another partner in this vast complex of mutually assisting religious memes.

Dawkins (1993) gives other examples from Judaism, such as the pointlessness of Rabbis testing for the kosher-purity of food, or the horrors of Jim Jones leading his flock to mass suicide in the Guyana jungle. Today he might add "Heaven's Gate" to the catalogue. "Obviously a meme that causes individuals bearing it to kill themselves has a grave disadvantage, but not necessarily a fatal one. .... a suicidal meme can spread, as when a dramatic and well-publicised martyrdom inspires others to die for a deeply loved cause, and this in turn inspires others to die, and so on." (Dawkins, 1982, p 111).

He might equally have chosen Islam; a faith that includes the concept of the jihad or holy war, and has particularly nasty punishments for people who desert the faith. Even today the author, and heretic, Salman Rushdie lives in fear of his life because many Muslims consider it their holy duty to kill him. Once you have been infected with powerful memes like these you must pay a high price to get rid of them.

Lynch (1996) explores in depth some tricks used by religions and cults. “Honour thy father and mother” is an excellent commandment, increasing the chance that children will take on beliefs from their parents, including the commandment itself. As a secular meme it might not succeed very well, since kids would surely reject it if they thought it came straight from the parents. However, presented as an idea from God (who is powerful, all-seeing and punishes disobedience) it has a much better chance - a good example of memes “ganging up”.

Dietary laws may thrive because they protect against disease, but may also keep people in the faith by making it harder for them to adapt to other diets outside. Moral codes may enhance effective cooperation and survival but may also be ways of punishing lapses of faith. Observing “holy days” ensures lots of time for spreading the memes, and public prayers and grace at meals ensure that lots of people are exposed to them. Learning sacred texts by heart, and setting them to inspiring or memorable music ensures their longevity.

In the long history of religions most of them have spread vertically - that is from parent to child. Even today the best predictor of your religion is your parent’s religion - even if you think you rationally chose the “best” or “truest” one! However, today more and more new religions and cults spread horizontally - from any person to any other person. The two types use different meme tricks for their replication.

As an example of the first type Lynch (1996) gives the Hutterites. They average more than ten children per couple, a fantastic rate that is possibly helped by the way they distribute parental responsibility, making each extra child only a slightly greater burden for its natural parents. Other religions put more effort into conversion, like the evangelical faiths which thrive on instant rewards and spiritual joy on conversion.

In case I seem to be implying that people have deliberately manufactured religions this way, that is not at all what I mean. Look at it this way - imagine in the long, long history of human religious endeavour, all the millions and millions of different statements, ideas, and commandments that must have been uttered at some time or another. Which would you expect to have survived through to the present? The answer is, of course, the ones that just happened to have included clever tricks or come together with other ideas they could gang up with. The countless millions of other ideas have simply been lost. This is memetic evolution.

## Taking the Meme’s Eye View

We are now ready to take on the meme’s eye view. The basic approach I take is this - imagine a world full of hosts for memes (e.g. brains) and far more memes than can possibly find homes. Now ask - which memes are more likely to find a safe home and get passed on again? It’s that simple.

In doing this I try to follow some simple rules.

First, remember that memes (like genes) do not have foresight!

Second, consider only the interests of the memes, not of the genes or the organism. Memes do not care about genes or people - all they do is reproduce themselves. Short-hand statements like "memes want x" or "memes try to do y" must always be translatable back into the longer version, such as "memes that have the effect of producing x are more likely to survive than those that do not."

Third, memes, by definition, are passed on by imitation. So learning by trial and error or by feedback is not memetic, nor are all forms of communication. Only when an idea, behaviour or skill is passed on by imitation does it count as a meme. Now, remembering these rules, we can ask the question and see where it leads.

Imagine a world full of brains, and far more memes than can possibly find homes. Which memes are more likely to find a safe home and get passed on again?

Some of the consequences are startlingly obvious - once you see them. And some are frighteningly powerful.

I shall start with two simple ones, partly as exercises in thinking memetically.

### 1. Why can't we stop thinking ?

Can you stop thinking? If you have ever meditated you will know just how hard this is - the mind just seems to keep blithering on. If we were thinking useful thoughts, practising mental skills, or solving relevant problems there might be some point, but mostly we don't seem to be. So why can't we just sit down and not think? From a genetic point of view all this extra thinking seems extremely wasteful - and animals that waste energy don't survive. Memetics provides a simple answer.

Imagine a world full of brains, and far more memes than can possibly find homes. Which memes are more likely to find a safe home and get passed on again?

Imagine a meme that encourages its host to keep on mentally rehearsing it, or a tune that is so easy to hum that it goes round and round in your head, or a thought that just compels you to keep thinking it.

Imagine in contrast a meme that buries itself quietly in your memory and is never rehearsed, or a tune that is too unmemorable to go round in your head, or a thought that is too boring to think again.

Which will do better? Other things being equal, the first lot will. Rehearsal aids memory, and you are likely to express (or even sing) the ideas and tunes that fill your waking hours. What is the consequence? The memosphere fills up with catchy tunes, and thinkable thoughts. We all come across them and so we all think an awful lot.

The principle here is familiar from biology. In a forest, any tree that grows tall gets more light. So genes for growing tall become more common in the gene pool and the forest ends up being as high as the trees can make it.

We can apply the same principle again.

## 2. Why do we talk so much?

Imagine a world full of brains, and far more memes than can possibly find homes. Which memes are more likely to find a safe home and get passed on again?

Imagine any meme that encourages talking. It might be an idea like “talking makes people like you” or “It’s friendly to chat”. It might be an urgent thought that you feel compelled to share, a funny joke, good news that everybody wants to hear, or any meme that thrives inside a talkative person.

Imagine in contrast any meme that discourages talking, such as the thought “talking is a waste of time”. It might be something you dare not voice aloud, something very difficult to say, or any meme that thrives inside a shy and retiring person.

Which will do better? Put this way the answer is obvious. The first lot will be heard by more people and, other things being equal, simply must stand a better chance of being propagated. What is the consequence of this? The memosphere will fill up with memes that encourage talking and we will all talk an awful lot. And we do!

A simpler way of putting it is this:- people who talk more will, on average, spread more memes. So any memes which thrive in chatterboxes are likely to spread.

This makes me see conversation in a new light. Is all that talking really founded on biological advantage? Talking takes a lot of energy and we do talk about some daft and pointless things! Do these trivial and stupid thoughts and conversations have some hidden biological advantage? I would at least like to offer the suggestion that they do not. That we do all this talking and all this thinking merely because the memes that make us do it are good survivors. The memes seem to be working against the genes.

This sets the stage for a more audacious suggestion.

## 3. Why are we so nice to each other?

Of course we aren’t always nice to each other, but human co-operation and altruism are something of a mystery - despite the tremendous advances made in understanding kin selection and inclusive fitness, reciprocal altruism and evolutionarily stable strategies (see e.g. Wright, 1994; Ridley, 1996). Human societies exhibit much more cooperation than is typical of vertebrate societies, and we cooperate with non-relatives on a massive scale (Richerson and Boyd, 1992). As Cronin puts it, human morality “presents an obvious challenge to Darwinian theory” (Cronin, 1991, p325).

Everyone can probably think up their own favourite example. Richard Dawkins (1989 p 230) calls blood donation "a genuine case of pure, disinterested altruism". I am more impressed by charitable giving to people in faraway countries who probably share as few of our genes as anyone on earth and whom we are unlikely ever to meet. And why do we hand in wallets found in the street, rescue injured wildlife, support eco-friendly companies or recycle our bottles? Why do so many people want to be poorly paid nurses and counsellors, social workers and psychotherapists, when they could live in bigger houses, attract richer mates, and afford more children if they were bankers, stock brokers or lawyers?

Many people believe all this must ultimately be explained in terms of biological advantage. Perhaps it will, but I offer an alternative for consideration; a memetic theory of altruism. We can use our, by now, familiar tactic.

Imagine a world full of brains, and far more memes than can possibly find homes. Which memes are more likely to find a safe home and get passed on again?

Imagine the sort of meme that encourages its host to be friendly and kind. It might be a meme for throwing good parties, for being generous with the home-made marmalade, or just being prepared to spend time listening to a friend's woes. Now compare this with memes for being unfriendly and mean - never cooking people dinners or buying drinks, and refusing to give your time to others. Which will spread more quickly?

The first type, of course. People like to be with nice people. So those who harbour lots of friendliness memes will spend more time with others and have more chances to spread their memes. In consequence many of us will end up harbouring lots of memes for being nice to others.

A simpler way of putting it is this:- people who are altruistic will, on average, spread more memes. So any memes which thrive in altruistic people are likely to spread - including the memes for being altruistic.

You may wish to challenge any of the above steps. It is therefore reassuring to learn from many experiments in social psychology, that people are more likely to adopt ideas from people they like (Eagly and Chaiken, 1984). Whether this is a cause or a consequence of the above argument is debatable. It would be most interesting if psychological facts like this, or others such as cognitive dissonance, or the need for self esteem, could be derived from simple memetic principles - but that is a topic for another time!

For now we should consider whether the idea is testable. It predicts that people should act in ways that benefit the spread of their memes even at some cost to themselves. We are familiar with buying useful information, and with advertisers buying their way into people's minds for the purposes of selling products, but this theory predicts that people will pay (or work) simply to spread the memes they hold - because the memes force them to. Missionaries and Jehovah's Witnesses seem to.

Many aspects of persuasion and conversion to causes may turn out to involve meme-driven altruism. Altruism is yet another of the meme tricks that religions (those most powerful of meme-complexes) have purloined. Almost all of them thrive on making their members work for them and believe they are doing good.

Of course, being generous is expensive. There will always be pressure against it, and if memes can find alternative strategies for spreading they will. For example, powerful people may be able to spread memes without being altruistic at all! However, that does not change the basic argument - that altruism spreads memes.

You may have noticed that the underlying theme in all these arguments is that the memes may act in opposition to the interest of the genes. Thinking all the time may not use much energy but it must cost something. Talking is certainly expensive, as anyone who has been utterly exhausted or seriously ill will attest. And of course any altruistic act is, by definition, costly to the actor. I would say that this is just what we should expect if memes are true replicators. They do not care about the genes or the creatures the genes created. Their only interest is self-propagation. So if they can propagate by stealing resources from the genes, they will do so.

In the next example we see the memes forcing the hand of the genes in a much more dramatic way.

#### 4. Why are our brains so big?

Yes, I know this is an old chestnut, and there are lots and lots of good answers to the question. But are they good enough? Let us not forget how mysterious this issue really is. Brains are notoriously expensive both to build and to run. They take up about 2% of the body's weight but use about 20% of its energy. Our brains are three times the size of the brains of apes of equivalent body size. Compared to other mammals our encephalisation quotient is even higher, up to about 25 (Jerison, 1973; Leakey, 1994; Wills, 1993). On many measures of brain capacity humans stand out alone. The fact that such intelligence has arisen in an animal that stands upright may or may not be a coincidence but it certainly adds to the problem. Our pelvises are not ideally suited for giving birth to huge brains and so childbirth is a risky process for human beings - yet we do it. Why?

The mystery was deepened for me by thinking about the size of the biological advantage required for survival. In a study concerned with the fate of the Neanderthals, Zubrow (Leakey, 1994) used computer simulations to determine the effect of a slight competitive edge. He concluded that a 2% advantage could eliminate a competing population in less than a millennium. If we needed only such a tiny advantage why do we have such a huge one? Several answers have recently been proposed. For example, Dunbar (1996) argues that we need large brains in order to gossip, and we need to gossip as a kind of verbal grooming to keep very large bands of people together. Christopher Wills (1993) argues that the runaway evolution of the human brain results from an increasingly swift gene-environment feedback loop. Miller (1993) proposes that our vast brains have been created by sexual selection; and Richerson and Boyd (1992) claim they are used for individual and social learning, favoured under increasing rates of environmental variation.

What these authors all have in common is that their ultimate appeal is to the genes. Like Dawkins' bewailed colleagues, they always wish to go back to biological advantage. I propose an alternative based on memetic advantage.

Imagine early hominids who, for good biological reasons, gained the ability to imitate each other and to develop simple language. Once this step occurred memes could begin to spread, and the second replicator was born. Remember - once this happened the genes would no longer be able to stop the spread! Presumably the earliest memes would be useful ones, such as ways of making pots or knives, or ways of catching or dismembering prey. Let us assume that some people would have slightly larger brains and that larger brains are better copiers. As more and more people began to pick up these early memes, the environment would change so that it became more and more necessary to have the new skills in order to survive.

A person who could quickly learn to make a good pot or tell a popular story would more easily find a mate, and so sexual selection would add to the pressure for big brains. In the new environment larger brained people would have an advantage and the importance of the advantage would increase as the memes spread. It seems to me that this fundamental change in selection pressures, spreading at the rate of meme propagation, provides for the first time a plausible reason why our brains are totally out of line with all other brains on the planet. They have been meme-driven. One replicator has forced the moves of another.

## 5. Who am I?

We can now see the human mind as the creation of two replicators, one using for its replication the machinery created by the other. As Dennett pointed out, people are animals infested with memes. Our personalities, abilities and unique qualities derive from the complex interplay of these replicators. What then of our innermost selves - the "real me", the person who experiences "my" life?

I would say that selves are co-adapted meme complexes - though only one of many supported by any given brain (Blackmore, 1996). Like religions, political belief systems and cults, they are sets of memes that thrive in each other's company. Like religions, political belief systems and cults, they are safe havens for all sorts of travelling memes and they are protected from destruction by various meme-tricks. They do not have to be true.

In fact we know that selves are a myth. Look inside the brain and you find only neurons. You do not find the little person pulling the strings or the homunculus watching the show on an inner screen (Dennett, 1991). You do not find the place where "my" conscious decisions are made. You do not find the thing that lovingly holds all those beliefs and opinions. Most of us still persist in thinking about ourselves that way. But the truth is - there is no one in there!

We now have a radically new answer to the question "Who am I?", and a rather terrifying one. "I" am one of the many co-adapted meme-complexes living within this brain. This scary idea may explain why memetics is not more popular. Memetics deals a terrible blow to the supremacy of self.

## The Future for Memes

The memes are out! For most of human history memes have evolved alongside genes. They were passed on largely vertically - from parent to child - and therefore evolved at much the same rate as genes. This is no longer true. Memes can leap from brain to brain in seconds - even when the brains are half a planet apart.

While some memes hang around in brains for weeks, months or years before being passed on, many now spread in multiple copies at the speed of light. The invention of the telephone, fax machine and e-mail all increase the speed of propagation of memes. As high speed, accurate, horizontal copying of memes increases we can expect some dramatic developments in the memosphere.

First, the faster memes spread the weaker is the hold of natural (genetic) selection. This relative uncoupling of genes and memes may mean that more than ever before memes will spread that are detrimental to their carriers. We may be seeing this already with some of the dangerous cults, fads, political systems, copy-cat crimes and false beliefs that can now spread so quickly. Second, we may expect memes to build themselves ever better vehicles for their own propagation. Genes have built themselves organisms to carry them around in. What is the memic equivalent? Artifacts such as books, paintings, tools and aeroplanes might count (Dennett, 1995) but they are feeble compared with computers or the Internet. Even these recent inventions are still largely dependent on humans for their functioning, and on the genes those humans are carrying - after all, sex is the most popular topic on the internet. So can the second replicator ever really break free? It might if ever we construct robots that directly imitate each other. Fortunately this is such a difficult task that it will not be achieved very soon and perhaps by then we will have a better understanding of memetics and be in a better position to cope with our new neighbours.

## Conclusion

I have shown how a theory of memetics provides new answers to some important questions about human nature. If I am right, then we humans are the product of two replicators, not just one. In the past hundred years we have successfully thrown off the illusion that a God is needed to understand the design of our bodies. Perhaps in the next millennium we can throw off the illusion that conscious agents are needed to understand the design of our minds.

## References

- Barkow,J.H., Cosmides,L. and Tooby,J. (Eds) (1992) *The Adapted Mind: Evolutionary psychology and the generation of culture*. Oxford, OUP.
- Blackmore,S.J. (1996) Waking from the Meme Dream The Psychology of Awakening: An International Conference on Buddhism, Science and Psychotherapy, Dartington, 9 November 1996.

- Boyd,R. and Richerson,P.J. (1990) Group selection among alternative evolutionarily stable strategies. *Journal of Theoretical Biology*, 145, 331-342
- Cavalli-Sforza,L.L. and Feldman,M.W. (1981) *Cultural Transmission and Evolution: A quantitative approach*. Princeton, NJ, Princeton University Press.
- Cloak,F.R. (1975) Is a cultural ethology possible? *Human Ecology*, 3, 161-182
- Cronin,H. (1991) *The Ant and the Peacock*. Cambridge, Cambridge University Press.
- Darwin, C. (1859) *On the Origin of Species by Means of Natural Selection*. London, Murray.
- Dawkins,R. (1976) *The Selfish Gene* Oxford, Oxford University Press (new edition with additional material, 1989)
- Dawkins,R. (1982) *The Extended Phenotype*. Oxford, Oxford University Press.
- Dawkins,R. (1993) Viruses of the mind. In B.Dahlbohm (ed) *Dennett and his Critics: Demystifying Mind*. Oxford, Blackwell.
- Dennett,D. (1991). *Consciousness Explained*. Boston, Little, Brown.
- Dennett,D. (1995) *Darwin's Dangerous Idea*, London, Penguin
- Donald, M (1991) *Origins of the Modern Mind: Three Stages in the Evolution of Culture and Cognition*. Cambridge, Mass, Harvard University Press.
- Dunbar,R. (1996) *Grooming, Gossip and the Evolution of Language*. London, Faber & Faber.
- Durham,W.H. (1991) *Coevolution: Genes, Culture and Human Diversity*. Stanford, Ca., Stanford University Press.
- Eagly,A.H. and Chaiken,S. (1984) Cognitive theories of persuasion. In L.Berkowitz (Ed), *Advances in Experimental Social Psychology*, 17, 267-359. New York, Academic Press.
- Gould,S.J. (1996) *Full House*. Harmony Books (Published in the UK as *Life's Grandeur*, London, Jonathan Cape.)
- Grant,G. (1990) Memetic lexicon. [http://pespmc1.vub.ac.be/\\*memes.html](http://pespmc1.vub.ac.be/*memes.html)
- Hofstadter (1985) *Metamagical Themas: Questing for the essence of mind and pattern*. N.Y. Basic Books
- Jerison,H.J. (1973) *Evolution of the Brain and Intelligence*. N.Y. Academic Press.
- Leakey,R. (1994) *The Origin of Humankind* London, Weidenfeld and Nicolson

- Lumsden,C.J. and Wilson,E.O. (1981) Genes, Mind and Culture. Cambridge, Mass., Harvard University Press.
- Lynch,A. (1996) Thought Contagion: How Belief Spreads Through Society. N.Y. Basic Books.
- Midgley,M. (1994) Letter to the Editor, New Scientist, 12 Feb, 50.
- Miller,G. (1993) Evolution of the Human Brain through Runaway Sexual Selection. PhD Thesis, Stanford University Psychology Department
- Mithen,S. (1996) The Prehistory of the Mind. London, Thames and Hudson.
- Pinker, S. (1994) The Language Instinct New York, Morrow
- Richerson,P.J. and Boyd,R. (1992) Cultural inheritance and evolutionary ecology. In E.A.Smith and B.Winterhalder (Eds) Evolutionary Ecology and Human Behaviour 61-92.
- Ridley,M. (1996) The Origins of Virtue. London, Viking.
- Tudge,C. (1995) The Day before Yesterday: Five Million Years of Human History, London, Jonathan Cape.
- Wills,C. (1993) The Runaway Brain: The Evolution of human uniqueness. N.Y., Basic Books.
- Wilson, E.O. (1978) On Human Nature. Cambridge, Mass., Harvard University Press.
- Wright,R. (1994) The Moral Animal. Pantheon Books.

Skeptic Magazine contact details:

Web: <http://www.skeptic.com/>

Email: SKEPTICMAG@aol.com, also lippard@skeptic.com

Tel: +1 818 794 3119

Fax: +1 818 794 1301

Post: *Skeptics Society*, PO Box 338, Altadena, CA 91001

# THE EVOLUTIONARY ECOLOGY OF SCIENCE

By  
**Marion Blute**

Department of Sociology  
University of Toronto at Mississauga  
marion.blute@utoronto.ca

## Abstract

In the study of sociocultural/memetic evolution, approaches modelled on population genetics (e.g. Cavalli-Sforza and Feldman 1981) and on systematics/taxonomy (e.g. Hull 1988) have been prominent but the influence of evolutionary ecology has been slight. In the tradition of Toulmin (1972) and Hull (1988) this paper is about cultural evolution in science. In particular, it applies some principles of evolutionary ecology to the scientific process. The effects of density, scale, frequency and heterogeneity on various strategies of research and teaching in science are considered. In the future, the analysis should be extended to the sociobiology of science including the relationship between supervisors and their graduate and postdoctoral students and among peers e.g. publishing and other forms of social collaboration.

**Keywords:** sociocultural evolution, cultural evolution, memetics, sociology of science, philosophy of science, science studies, evolutionary ecology

## 1 Introduction: Sociocultural/Memetic Evolution, Science and Evolutionary Ecology

Since Carl Degler's prize winning book (1991) documented the decline and subsequent revival of Darwinism in social thought, Darwinian-style evolution-based theorizing and research has continued to grow rapidly in a whole host of humanistic and social scientific disciplines and fields. It comes in three broad forms. These are: human behavioural ecology/evolutionary psychology/sociobiology emphasizing biological (gene-based) evolution, sociocultural evolutionism emphasizing sociocultural (social learning or meme-based) evolution, and gene-culture coevolution emphasizing their interaction (gene-culture coevolution). While not denying the utility of the others, this article is concerned exclusively with the second -- with sociocultural evolution. Among these, some find it useful or even necessary (after Dawkins 1976) to use the concept of memes, analogous to genes, for the basic units of hereditary information or replication involved, others are indifferent, and still others find it misleading. While those who find the meme concept misleading may be underestimating the similar ambiguity, even multiplicity, of gene concepts in biology, those who find it indispensable may be underestimating how much useful theorizing and research has gone on historically in biology including much taxonomy and ecology largely ignoring genes while working purely at the phenotypic level. Because, as Whiten put it: "that there are just two major forms of behavioural evolution, occurring through genetic and cultural transmission respectively must rank among the most exciting and fundamental discoveries of biology achieved over the last century and a half" (2001:359), I find myself in agreement with those who think it is more important to just get on with it. [note 1]

Social scientists and even humanists in a variety of fields have been getting on with it. These include not only epistemology (Campbell 1988: Part V, Plotkin 1994, Buskes 1998), anthropology (Sperber 1996, Cullen 2000, Ehrlich 2000) and sociology (Runciman 1989, Lenski et. al. 1994, Fog 1999) generally, but also more specific fields such as institutional economics (Boulding 1981, Nelson and Winter 1982), linguistics (Ruhlen 1994, Croft 2000), organizations (Aldrich 1999) and technology studies (Basalla 1988, Ziman 2000) in addition to memetics of course (Brodie 1996, Lynch 1996, Blackmore 1999 and Aunger 2000) to mention only some books.

On the topic of sociocultural evolution in science in particular which is the subject of this paper, Darwinian-style evolutionary approaches emerged more from the philosophy than from the sociology of science. Although Popper's (1963) falsificationism was evolutionary at least in the negative sense of selection against, in retrospect, it seems almost inevitable that once Kuhn (1962) started his revolution making philosophers of science more historically and sociologically aware, that a full-blown Darwinian-style evolutionary approach to science would emerge. The pioneer was Stephen Toulmin (1972) but eventually, more than anyone else, it was developed by David Hull. Hull's sophisticated Science as a Process (1988) is at once a treatise on the philosophy of science, the sociology of science, and evolutionary theory (whether instantiated biologically or socioculturally). Not immodestly, but accurately in my opinion, Hull views himself as the heir of Kuhn and Toulmin. The major elements of his evolutionary theory of science are "curiosity" (mutation), "credit" (descent) and "checking" (selection). While this paper is in the tradition of Hull's theory, there are some differences in emphasis. It places an equal emphasis on the acquisition of knowledge in science and its dissemination, it deals almost exclusively with dissemination through graduate and postdoctoral students rather than through peers by means of paper publishing etc. (the sociobiology of science is beyond the scope of this paper), and emphasizes more selection over descent (the reason for which follows).

In any evolutionary process, the currently existing array of forms is explained by constraints (most obviously the laws of physics and chemistry but there may be others), chance (mutation and sampling error in finite populations), history and selection. The three most important research programmes in evolutionary biology are population genetics (which deals with some of these formally), taxonomy which emphasizes history by building trees, and evolutionary ecology which emphasizes the forces of, and responses to, selection. In previous work on sociocultural evolution, approaches modelled on population genetics (particularly Cavalli-Sforza and Feldman 1981) and on taxonomy (particularly Hull, not surprisingly given that the science he studied was taxonomy) have been prominent but the influence of evolutionary ecology has been slight. While this paper is about cultural evolution in science then, it is particularly about evolution in science under different ecological conditions.

Darwin's theory of evolution was a theory 'that' natural selection, rather than 'of' natural selection. The Origin of Species contains no general theory of the circumstances under which selection favours this or that property of organisms. Evolutionary ecology is the discipline that attempts to fill this gap (Cockburn 1991, Bulmer 1994, Fox, Roff and Fairbairn 2001) see also the relevant parts of general ecology texts including (Begon, Harper and Townsend 1996, Pianka 1999, and Ricklefs and Miller 2000). Evolutionary ecology seeks a theoretical halfway house between the near-universal tautology of the fitness-selection nexus and the near-complete

historical specificity of the myriad details of what is adaptive in locally prevailing circumstances. Its research programme is in turn divided into a number of subbranches, each of which combines evolution and ecology with some other traditional field of biological study: physiology (Sibly and Calow 1986), development i.e. life histories (Roff 1992, Stearns 1992, Charnov 1993), behaviour (Krebs and Davies 1993, Krebs and Davies 1997) including foraging (Stephens and Krebs 1986), and extends even beyond populations to metapopulations (Rhodes, Chesson and Smith 1996, Hanski and Gilpin 1997, Hanski 1999). The appropriate terminology becomes long as in evolutionary physiological (or developmental or behavioural or metapopulation) ecology. In addition, an intersecting set of divisions pertain to the group studied microbial (Andrews 1991), plant (Bazzaz 1996), animal, and human (Smith and Winterhalder 1992). Despite this complexity, it is often recognized that similar principles should apply to all divisions (Andrews 1991) and at all scales (Bazzaz 1996).

The most important ecological conditions which control evolution are density (low versus high discussed in sections two and three) and scale (small versus large discussed in section five). These reflect the optimization approach (in the narrow sense) to evolutionary ecology. In both cases, alternatively or in addition, negative frequency dependence may obtain (it being advantageous to do the opposite of what others are doing). This is the game theory approach discussed with respect to density dependence in section four and with respect to scale dependence included in section five. In both cases as well, heterogeneity of conditions may obtain as discussed in section six. In all cases, the evolutionary-ecological theoretical approach is to ask, and try to answer logically, "Under alternative ecological conditions, what adaptations, or 'strategies' as they are called, would selection favour?" Given that the influence of evolutionary ecology on the study of sociocultural evolution and memetics has been slight as noted, the concepts employed may be largely unfamiliar to readers of this paper. If so, it might be helpful to follow the summaries in Table 1a (on density dependence and its complications) and in Table 1b (on scale dependence and its complications) as you go. For more interested or knowledgeable readers, a discussion of how the evolutionary ecology employed here differs from others where it does is included in the appendix.

## **2 Density Dependence: Primary Versus Secondary Research/Teaching**

Much of modern evolutionary ecology can be understood as flowing from the premise that the most fundamental physical properties of the ecological environment energy content, time and space (whether literal or metaphorical, i.e. niche time/space) are what organisms are most liable to evolve adaptations to. Albeit controversial in some respects, the most general theory proposed to date and hence the most convenient to use to organize this discussion is the verbal theory of density-independent ( $r$ ) versus density-dependent ( $K$ ) selection (MacArthur 1962, MacArthur and Wilson 1967). [note 3] Consider the following specific interpretation of the former for individual growth in a homogeneous environment. When population density (measured for now in size) relative to resources [note 4] is low, selection favours spending on growth. This entails maximizing  $r$  (the intrinsic growth rate of the population measured in size) by acquiring as much in the way of resources as possible and engaging in intermediary (degradative) metabolism. When density is high on the other hand, selection favours investing in maintenance. This entails maintaining the population at or near the carrying capacity of the environment,  $K$ , by means of

intermediary (biosynthetic) metabolism, building up structures with some of the substances derived and excreting others. [note 5]

**Table 1. Some Principles of the Evolutionary Ecology of Science** [note 2]

**Table 1a) Density-Dependent Selection**

Homogeneous Environment Growth OR 3M's	
At Low Density or History of Catastrophes growth	At High Density or History of Bonanzas maintenance, motility, mutability
Heterogeneous Environment Flexibility	
Fine-grained Heterogeneity (environment changes/varies between low and high densities with high frequency relative to period/distance) morphological, physiological, behavioural change/differentiation	Coarse-grained Heterogeneity (environment changes/varies between low and high densities with long period/distance relative to frequency) phenotypic plasticity

**Table 1b) Scale-Dependent Selection**

Homogeneous Environment r/d OR t/s Growth	
In or History of Small-scale Environment rate/density growth	In or History of Large-scale Environment time/space growth
Heterogeneous Environment Flexibility	
Fine-grained Heterogeneity (environment changes/varies between small & large scales with high frequency relative to period/distance) morphological, physiological, behavioural change/differentiation	Coarse-grained Heterogeneity (environment changes/varies between small & large scales with long period/distance relative to frequency) phenotypic plasticity

There are many kinds of scientific research. New empirical research can yield knowledge, new theories, understanding, new combinations of these, explanations and predictions. New methods can alter the ways in which any of these kinds of research are conducted. Whatever classification of kinds of research one prefers, the analysis here is intended to apply to all of them but for convenience, the word "knowledge" will be used to apply to all. It has been a fundamental principle of the sociology of science since Merton (1973) virtually invented the subject, that scientists compete for status rather than wealth, income or power. They compete to have their work incorporated into that of others (through their direct cultural heirs -- graduate and postdoctoral students as well as indirectly through peers via publications with their influence or ancestry acknowledged in citations). In Hull's terminology, they compete for "credit". Of course just as organisms must grow and develop in order to reproduce, scientists must acquire knowledge i.e. do research in order to have something to communicate and receive credit for having done so. Imagine then members of a population of scientists competing to acquire knowledge from nature by means of research on a new topic or in a new field -- a low density or uncrowded environment. Because the topic or field is new, none of the competitors know very much and a whole world lies waiting to be discovered. Given the objective of maximizing personal knowledge, when under such benign conditions it pays to engage in as much primary research as possible, data is collected empirically, and axioms or postulates (or in multi-paradigm sciences theories themselves) are added. Complex inputs must additionally be analysed (broken down). In empirical research outliers must be identified, statistically significant results distinguished from the insignificant, repeatable results distinguished from the non-repeatable, and so on. Similarly in theoretical research analysis is necessary. Not only are axioms or postulates required, but theorems must be proven. As the topic or field matures however, competitors already possess a lot of knowledge and there is not much new left to be discovered.

That is not the end of the task however as has sometimes been thought (Horgan 1996). Knowledge must be maintained by what has sometimes been called "secondary research". The task of synthesis is primarily one of assimilating the new into the context of old ("adding new bricks to the wall" as the expression has it) but the positively misleading must be weeded out as well. At least three studies have shown that even articles which have been retracted continue to be cited (Campanario 2000), an extreme example of the necessity of such work. Other kinds of "repairs" are required -- for example errors corrected and findings rearranged in relation to one another. Much of the work of secondary research is performed by referees and writers of review articles, meta-analyses, textbooks, and other "secondary" literature with the goal of building durable knowledge. In science, as in life, low densities favour spending on growth particularly by acquisition and analysis; high densities favour investing in maintenance particularly by synthesis and rejection.

What applies individually in evolutionary ecology applies similarly demographically. When population density (measured now in numbers) is low relative to resources [note 6], selection favours spending on maximizing  $r$  (the growth rate of the population measured in numbers) by producing as many offspring as possible and "breaking down" by assessing quality. When density is high, selection favours investing in maintaining the population at or near the carrying capacity of the environment,  $K$ , by means of intermediary (biosynthetic) metabolism -- "building up" fewer offspring with parental care and "excreting" by practising active discrimination in favour of those assessed to be of high rather than low quality. In short, low densities favour

growth by quantity and assessment; high densities favour maintenance by quality and active discrimination. Again imagine members of a population of scientists competing in a low density environment, measured now in numbers. Competitors working on the new topic or field are few and a whole world lies waiting to be populated with researchers. Given the objective of maximizing public knowledge (Ziman 1968), under such benign conditions it pays to spend on what might be called "primary" teaching -- training as many graduate and post-doctoral students as possible and assessing quality. As the topic or field matures however, competitors are numerous and there is little room for more. Under such harsh conditions, it pays to invest in what might be called "secondary" teaching. Demographic success is better served by mentoring fewer students and pruning the "brood" to improve quality.

In summary, when densities are low, selection favours spending on maximizing  $r$  by acquiring knowledge and analysing it (primary research) and by teaching and assessing large numbers of graduate and postdoctoral students (primary teaching). When densities are high, it favours investing in maintaining  $K$  by synthesizing knowledge and weeding out inaccuracies (secondary research) and mentoring fewer students along with active discrimination (secondary teaching). Note that neither kind of strategy is intrinsically better than the other, rather, both should be similarly successful in their appropriate circumstances. A population that begins small and grows in size or numbers to become large and one which, already large, manages to remain so rather than declining are, at equal costs, equally successful. While often not so considered, scientists who add and those who maintain personal or public knowledge are contributing equally.

### **3 Density Dependence: Primary Research/Teaching Versus Moving or Innovating**

A second alternative to maintenance at high densities is motility. Where (rather than when) population density measured in size is low, selection favours spending on growth, but where it is high, selection favours investing in motility -- dispersing in space to where circumstances are more favourable. This distinction is commonly drawn in evolutionary ecology for example "handle versus search", "exploit versus explore", or "assimilate versus migrate" in foraging theory and "competitive ability" versus dispersal in the study of life histories. Even in the absence of motility as conventionally understood, this distinction can be applicable. In branching growth forms, laterals are often devoted to consuming (e.g. leaves in plants eat sunlight) while leaders are often devoted to extending the plant's reach. Stems in plants photosynthesise very little but extend the plant's reach, functioning as a form of motility in sessile organisms (Andrews 1991). All that applies to individual growth applies equally to demographic growth. Where densities are low, selection favours a large "clutch size" -- offspring quantity. Where it is high, selection favours motile offspring which, following the logic of trade-offs, would be expected to be fewer initially.

In their research and teaching, scientists often have opportunities to cope with diminishing returns to primary research in a topic or field by escaping to a different, more promising one: from molecular genetics to molecular neurobiology, from the war on cancer to the war on aids, from youth in the 60's to the ageing population of the 90's, etc. Individual scientists prepared to translocate knowledge acquired in one context to another can be successful. Hull's own work applying concepts and theories acquired in evolutionary biology to the sociology of science is a

case in point. Not only do scientists engage in such strategies, they sometimes prepare their students to do so. For example, many mathematicians, physicists and computer scientists moved into finance a decade or so ago (Chance and Peterson 1999) and some of the former kinds of graduate programmes specifically trained at least some of their students to do so.

In addition to maintenance and motility, high densities should also favour mutability. In 1982, Walter Fitch a pioneer of the study of molecular evolution clearly described and predicted the existence of the phenomena that have come to be known as "adaptive" and "directed" mutation respectively. "The organism might be better off if it could vary its mutation rate upward in stressful times and downward in favourable times ... If the organism needs to change only a few of its genes, one would prefer to increase the mutation rate in those genes specifically ... I predict that a mechanism for just this will be found" (quoted in Hall 1998). In 1988 in *Nature* and in 1991 in *Genetics* Cairns et. al. and Cairns and Foster created something of a sensation by reporting just such results for an experimental system in *E. coli*. Eventually the dust settled and many of the mechanisms even of adaptive although not necessarily directed mutation in bacteria and yeast are understood (for some reviews see Hall 1998, Foster 1999 and Rosenberg 2001). At least two mutator genes (which code for DNA polymerases with high error rates) have homologues in eubacteria, archaea and eukaryotes.

The concept of investing in innovation by means of research and development under conditions of scarcity is well known to economists many of whom see it as the solution to environmental problems. Indeed, it is well known to all of us. In recent decades, rather than acquiring more fossil fuels, we have to some degree squeezed more mileage for our vehicles and heat for our homes per unit of fuel input by increasing the efficiency of cars and furnaces (a response to the energy crisis of the 1970's suggested by Lovins in 1976). Historically, for heating, wood became scarce and expensive. We moved to coal and successively to oil and to natural gas. In the human economic world, normally innovations which increase efficiency or change niches require some investment in research and development. Depending upon how K, the carrying capacity of the environment is defined, "R&D" can be thought of as a strategy for raising carrying capacity or for realizing capacity unutilised for what evolutionists call historical reasons. The necessary innovations were simply not previously available in the population. R&D is not prescient but neither is it entirely random -- it is normally focused on the relevant technology. Similarly, its success is not guaranteed; much of the investment goes for naught, but successes often make it worthwhile. That does not mean however that economic change is not an evolutionary process. For example, Nelson and Winter (1982) and Nelson (1996) include search processes, including for information, in their evolutionary model of economic change. It is not surprising therefore that although risky, natural selection sometimes favours genes or the expression of existing inducible genes that search for genetic solutions to difficult problems or take advantage of opportunities -- genes that, in effect, invest in genetic R&D. It hardly needs to be said that there is nothing "Lamarckian" about the phenomenon of adaptive mutation. Even if somewhat, although obviously never perfectly directed, (e.g. toward genes coding for enzymes in a biochemical pathway whose substrate is in short supply or toward those making organisms capable of utilizing a recently unutilised resource) mutator genes themselves must evolve by normal Darwinian processes. In eukaryotes, the logic of the evolution of such genes would be similar to that of genes for recombination i.e. genes which have their effect indirectly through other genes on which they then hitchhike. In the broadest sense, all of science is a form of R&D.

In a narrower sense however, it seems likely that, as well as favouring investing in the first "2M's", stalled progress should favour a 3<sup>rd</sup> M -- mutability, i.e. individuals with heretical tendencies or the expression of such, disposed towards breaking out of the mould of existing theories, methods, etc., in their research/teaching searching for ways to reach new goals or to break through historical obstacles to achieving existing ones.

## 4 Frequency Dependence: Avoiding Competition

To be favoured, normally the high-density strategies of maintenance, motility or mutability under discussion in the last two sections require what anthropologists call an uncircumscribed or unbounded environment (Ehrlich 2000:239 after Carneiro 1970). Such an environment is renewable, colonizable or with historically unutilised carrying capacity and hence crowding is only temporary, local, or niche-specific respectively. Maintenance is of no use if resources are not renewable; motility is of no use if there is not space available to move into; and mutability is of no use if there is no carrying capacity unutilised for historical reasons. However, ecological environments not only affect the strategies which evolve but are in turn affected by them, i.e. there are evolutionary and ecological interactions. Low densities favour growth strategies, which cause densities to rise depleting and degrading the environment. These high (temporary/local/niche-specific) densities in turn favour the spread of maintenance/motility/mutability strategies which do not so much cause densities to decline as they do carry population members to a time/space/niche in which densities are low favouring the spread of growth strategies once again.

Given that growth strategies do normally have direct ecological consequences in the form of depletion and degradation, it is not surprising that members of some populations, instead of evolving in response to density, evolve in direct response to the strategies present in the population, i.e. to what the others are doing. Rather than ecological optimization in the narrow sense, the principle is the game-theory one of negative frequency dependence. If some individuals (or all of them some of the time -- either at random or with some fixed probability) are engaging in primary research/teaching, then it pays others to do secondary research/teaching, to change topics or fields, or to innovate and vice-versa. This negative frequency-dependent principle is to avoid competition. Much science involves not so much the kind of intensely competitive race made famous by Watson (1969), as it does precisely avoiding such races: with A studying X, B studying Y, C studying Z etc. However, negative frequency dependence can be added to, rather than simply substituted for, density dependence. Under the simplest set of assumptions, at equal costs, pure density dependence predicts strategies should be present in proportion to density. These range from all growth at one extreme, through equal frequencies at intermediate densities, to all one or more of the 3M's at the other extreme. At equal costs, pure frequency dependence predicts the two should always be present at equal frequencies. At equal costs, combined density and frequency dependence still predicts equal frequencies at intermediate densities, but other equilibria are possible. For example at equal costs, with resources depleted by 20%, growth and maintenance/motility/mutability strategies should be present in a ratio of 2:1 [note 7]. In short, in science as in life, selection should be sensitive not only to the ecological environment but also to the social environment in the sense of what strategies others are pursuing.

## 5 Scale Dependence: Fast Specialists Versus Longer, Slower Generalists

Thus far it has been assumed that spatio-temporal boundaries are fixed (now versus then, here versus there, this niche versus that) and appropriate strategies have been considered. Now shift controls. Instead of assuming boundaries are fixed while energy content varies, allow boundaries to vary while energy content is fixed. Environments may differ because the same energy content is more concentrated or dispersed in time/space.

Resources may be devoted to doing research rapidly but for a shorter period of time in a field, yielding a higher rate (mass per unit time, so to speak) of knowledge. Alternatively, resources may be devoted to doing research at a slower rate but for longer (yielding knowledge over a longer time period). Biologically (but not sociologically), trade-offs between growth rate and time are well known. Mammals that develop more rapidly by many criteria lead faster, but shorter lives. "Live fast, die young" is the rule (Promislow and Harvey 1990). More familiar sociologically is the spatial equivalent. Resources may be devoted to doing research more intensively but in a smaller area of a field, (yielding a higher mass per unit area of knowledge so to speak) i.e. as a "specialist". Alternatively, resources may be devoted to doing research less intensively but spanning a larger area i.e. as a "generalist". Similar distinctions apply to teaching. Students with graduate degrees may be produced at short "interbirth" intervals or alternatively, at a slower rate but for longer. They may be produced in a single topic or alternatively, teaching effort may be spread more thinly but broadly across a wider range of topics. It was once thought that spatiotemporal properties of organisms including lifespan and size could be attributed to density. Now it seems obvious that they should instead be attributed to environmental scale. Organisms (and scientists) should evolve to be adapted to the spatiotemporal scale of environmental opportunities, whether small (narrow/short) or large (broad/slow but long).

Moreover there is a correlation between spatial and temporal strategies so that smaller organisms (e.g. measured in length) and more specialized tend to have fast life cycles, while the larger (measured similarly) and more generalized tend to have longer, slower life cycles. The reason for such correlations is not obvious. To attribute them to ecology simply moves the need for an explanation one step back. They may be attributable to cytoplasmic and other nongenetic forms of inheritance or to genetic correlations as follows.

Working rapidly (and by analogy) intensively is expensive in energy, i.e. in ATP or the mitochondria that generate it -- in theoretical economics, roughly labour power. Working more slowly, for longer may be more expensive in enzymes (tools), building blocks (inventories of supplies), ribosomes (protein factories), and mRNA (managerial knowledge and expertise) -- all in theoretical economics, roughly capital. Or to choose a technological analogy, Formula 1 racing cars are basically portable gas tanks (and small in area or volume) while vehicles that compete in long distance rallies carry tools, spare parts, repair manuals and so on (and are larger). In short, it may not just take a long time to grow large as has often been thought. It may also require a large area or volume to store all the things required to keep going for a long time. Hence organisms that cytoplasmically inherit more ATP or mitochondria themselves may have a material comparative advantage for both rapid growth and intense specialization. Those that inherit more "capital" may have such an advantage for both longer, slower growth and

generalization. Alternatively, the preadaptations may be genetic in which genes for the former complement and mutually favour each other while genes for the latter do the same. As is well known, selection is not just ecologically-dependent, but also own condition-dependent (including, but not exclusively, genetic condition).

Examples of such contrasts and their correlations are easy to come by in the history of science. Contrast contemporary examples of competitions with their narrow objectives to find the "gene for" whatever (which tend to develop and be concluded rapidly), with Darwin's work on evolution or even Hull's own work. The broad scope of the former is well known. So too is the length of time it took to complete. The scope of Hull's book is also truly vast in it are woven together a treatment of most major issues in the philosophy of science, the sociology of science, and evolutionary theory. According to the preface, his decision to study the systematics community sociologically was made in 1973 (although he probably became convinced of the relevance of the kin-selection concept to the sociology of science even earlier). While much of the material in his book appeared previously in journal articles, the entire package was not published until 15 years later in 1988. If the reasons for such correlations are endogenous, it means that scientists need to assess not only the scale of opportunities available in the field in which they are working but also their own suitability for different strategies. Are they personally more suited for fast, specialized races or for longer, slower work that is broader in scope?

There exists some sentiment that contemporary science is missing out on opportunities for the acquisition of knowledge in the form of too much support for specialized, short-term projects and not enough for broader, longer-term projects. However, in the past decade, there have been some moves in that direction. For example, *Nature* editorialized on the desirability of support for longer projects (Maddox 1993) and published a commentary decrying the loss of breadth in scientific education and practice (e.g. Greene 1997). The U.S. National Science Foundation moved toward longer, bigger grants around the same time (Mervis 1993). The applied message here is the importance of adapting science and its funding to the opportunities available in a particular field and to the different kinds of comparative advantage possessed by researchers rather than the intrinsic superiority of one strategy over the other. And of course evolutionary and ecological interactions and frequency dependence may obtain with scale-dependent strategies just as they can with density-dependent strategies. Fast specialists deplete small-scale resources but may give large-scale ones an opportunity to recover and vice-versa. If some are or are mainly, fast specialists, then it can pay others to be longer, slower generalists and vice-versa. It is also worthwhile to point out that given similar energy content in the environment, rate/density and time/space strategies may have similar payoffs, that is only on average. Concentrating as they do in small spatiotemporal ranges, r/d strategies represent high risk, high potential return alternatives (like market timing and stockpicking) while t/s strategies represent low risk, low potential return alternatives (like index funds). Scientists who are highly specialized and work rapidly in their research and teaching may hit the jackpot, but more commonly they come up empty-handed -- the variance in outcomes should be high. Not surprisingly given the emphasis on specialized, short-term research, there is some evidence that error rates in the published literature have been increasing since the early 1970's (Hawkins 1999). Choosing exactly what to specialize in and when to exhaust resources in working at high speed is a risky business. Generalists who work more slowly but for longer may be less likely to

make a revolutionary discovery, but they should also be less likely to come up empty-handed. The variance in outcomes should be lower.

## 6 Heterogeneous Environments: Flexible Strategies

Whether considering density or scale dependence, we have largely assumed that environments are homogeneous and hence so too strategy populations. Densities were either low or high not both, scale was either small or large not both. When we did admit of heterogeneity (introducing frequency dependence in Section 4), it was assumed to be endogenous to the system. For example, if densities changed from low to high, it was assumed to be because a growing population depleted resources; if they changed from high to low, it was assumed to be because maintenance/motility/mutability strategies carried population members to a time/place/niche in which they were plentiful again. However, environments can be exogenously heterogeneous -- with periods/patches of low and others of high densities for reasons unrelated to internal interactions within the system. Similarly, resources can be exogenously distributed on more than one spatio-temporal scale (fruit on trees, trees in stands etc.) Levins (1968) distinguished between fine and coarse-grained environments. Here the terms heterogeneous versus homogeneous will be used initially because heterogeneous environments themselves may differ in grain. Given a homogeneous environment, selection favours spending in several senses (e.g. whether on growth directly or via the 3M's, or whether r/d or t/s). Given a heterogeneous environment however, selection favours investing in morphological, physiological, and behavioural change/differentiation or phenotypic plasticity -- both of which for convenience can be lumped together as 'flexibility'.

It has long been known that development is a process not only of growth but also of morphological, physiological and behavioural change and differentiation. In individual development, life cycles pass through a sequence of stages and different kinds of cells, tissues and organs emerge, becoming different from each other -- processes largely achieved at the molecular genetic level through change and variation in control of gene expression. A genotype can give rise to individuals, each of which is internally heterogeneous in time/space. At the opposite extreme, on an evolutionary rather than a developmental scale, populations, species and higher taxa change and diverge from each other -- processes largely achieved through genetic change and variation. Between these, cyclical change can take place between generations in dual or even multi-generational life cycles (common in the complex life cycles of many parasites), and different kinds of individuals can become different from each other (morphs, genders, morphs within genders etc.). This between as opposed to within generations/individuals but still intrapopulational change and variation is sometimes achieved genetically -- gender differences in birds and mammals are genetic. Sometimes it is achieved through change and variation in control of gene activity. In many turtles and reptiles gender is phenotypically plastic. A genotype can give rise to heterogeneous generations/kinds of individuals (the former presumably by means of epigenetic inheritance mechanisms). In either the within or between generations/individuals case, heterogeneous environments favour flexibility -- which type being dependent upon patch size. If small, morphological, physiological or behavioural change/variation tends to occur within generations/individuals (change/differentiation) -- creating phenotypic checkerboards. If large, they can obtain between generations/individuals (phenotypic plasticity) -- creating a checkerboard of phenotypes. If uncertainty prevails but reliable signals are available,

change/variation of either kind may be condition-dependent. Gender differences in many turtles and lizards are dependent upon environmental temperature. Investing in flexibility is favoured in heterogeneous environments because normally specialist stages/structures are more efficient in the range of a niche they specialize in than are generalists in that range. Presumably the cost of flexibility must initially be traded off against growth, i.e. there are fixed or set-up costs of switching/varying in order to adapt to a heterogeneous environment. Just as with growth versus the 3M's, under the right conditions, the advantages are more than sufficient to compensate for these costs.

Flexibility in scale that takes place both within (differentiation) and between (plasticity) individuals is illustrated by organisms with a branching architecture. Such architecture has evolved independently in organisms many times among the unicellular (some algae), the multicellular (some algae as well as fungi and plants), and the colonial. Colonies of many marine invertebrates in which the individual elements are multicellular animals develop such an architecture. It is found even among colonies of eusocial insect colonies as in polygyne (multiple queen) fire ants (*Solenopsis invicta*). Supercolonies of these insects spread in the American south in vast branching networks in which individual colonies, connected by underground tunnels, arise by a budding process (Mann 1994). There is a large literature on possible evolutionary explanations of the growth habit and other features of branching organisms much of it on whether the function of branching relates to resources, antagonists or competitive interactions (for a discussion and references see Buss and Blackstone 1991). Here, for convenience, we shall assume resources are at issue.

The simplest of such organisms may be viewed as being composed of what, for simplicity's sake, we shall call and "laterals" and "leaders". A branching architecture is correlated with sessility and as mentioned in Section 3, where only laterals consume and hence the internal flow of resources is unidirectional, the function of leaders is to extend the organism's reach. Leaders can function as a form of motility in sessile organisms (Andrews, 1991). Often however, both laterals and leaders consume as in a fungal hypha; hence the internal flow of resources is bi-directional. Assuming equal profitability at equilibrium, the normally shorter laterals must be consuming a higher mass per unit area or volume (and presumably more rapidly), i.e. are more r/d selected. The normally longer leaders must be consuming resources through a larger area or volume of space (and presumably for longer), i.e. are more t/s selected. The entire organism is a mixed r/d and t/s strategist, adapted to more than one ecological scale. Despite the mixed nature of their strategy, individuals and populations do not necessarily devote equal resources to the two. If laterals arise frequently, which in turn, functioning as leaders, also give rise to their own laterals frequently the overall effect is a short but dense, bushy structure, a pattern long called "phalanx" like. This is still a mixed strategy, but one tilted towards the r/d end of the spectrum. Conversely, if leader(s) only give rise to laterals infrequently, which in turn, functioning as leaders, also only give rise to their own laterals infrequently, the overall effect is a long but sparse, tree-like structure, a pattern long called "guerrilla" like. This is still a mixed strategy, but one tilted towards the t/s end of the spectrum (for diagrams and examples see Andrews 1991, Buss and Blackstone 1991). Hence the same character can exhibit both kinds of flexibility -- differentiation within and plasticity between individuals.

Branching organisms are often constructed modularly whether somatically, reproductively, or both (whether severally or jointly). Somatic modules commonly arise from buds in the upper axes between laterals and leaders. These are a physical manifestation of resources invested in differentiation. They give rise to both leaders and laterals; they permit environmental exploitation to take place on more than one scale.

The parallel in science is clear. In heterogeneous environments, scientists can be called upon to not confine themselves to a single strategy. They may be required to pay the set-up costs of varying and changing among strategies -- working on one project in one way and another in another way. At various times and places they may be required to engage in both primary and secondary research, in primary research and changing topics or fields, in being conformist and heretical, in working as a fast specialist and as a longer slower generalist -- even if there is likely to be an over-all emphasis on one or the other. Not only are such mixed strategies possible, in fields in which opportunities are available on more than one scale, such flexibility lies at the very heart of the scientific process. Scientists commonly describe what they do in their research as reasoning from the specific to the general and from the general to the specific and importantly, believe that each complements the other. Philosophers are more inclined to talk of induction and deduction, or in more modern terminology observation and theory, with explanations emerging from their interaction (Kosso 1992). Such a formulation is consistent with what is observed in nature. Resources spent on maximizing the acquisition of specialized knowledge rapidly may yield a return which can be invested in the acquisition of general knowledge, over a long period of time. Science can proceed from the specific to the general and from the short to the long run. Equally, resources spent on the acquisition of general knowledge over a long period of time may yield a return that can be invested in the acquisition of specific knowledge, rapidly. Science can proceed from the general to the specific, and from the long to the short run.

Commonly philosophers argue that for circularity to be avoided in the scientific process, observations that nourish a particular theory should be independent of those used to confirm it. Hull builds his evolutionary sociology of science from a series of narratives about the history of biology and then tests it against evidence from the contemporary systematics community. Because observation is "theory laden" the requirement may be put differently: the theory used to support particular observations should be independent of the theory for which the observations serve as evidence (Kosso 1992, 155-8). However it is put, branching organisms proceed in such a manner. Hence laterals that nourish a particular leader are distinct from the laterals that the leader subsequently supports. The part of a leader that supports particular laterals is distinct from the part of the leader that the laterals subsequently nourish. Science, like branching growth, is an iterative process. In highly developed sciences, developing new subtopics or specialties (modules) capable of giving rise to both is one of the most valuable activities of all.

For heuristic purposes I slipped from scientists' "reasoning from the particular to the general and from the general to the particular" to philosophers' induction and deduction or observation and theory. I emphasize however that the original terminology more adequately expresses the evolutionary-ecological strategy distinction drawn here. That distinction is not the traditional philosophical one rather it is between specialized research performed rapidly, and more generalized research performed slowly but for longer. Empirical, theoretical, or methodological

research can be carried out in either style. The scientific literature is replete with theoretical research for example, mathematical models, of narrow scope and short lifespan.

At the same time, some caution is in order. There is something of a cult of variation and change in science today which from a bench perspective often appears to have more to do with the careers of administrators and the whims of granting agencies than with the real needs of research and teaching. Scientists and departments are told to change and distinguish themselves from one another. It sometimes appears to scientists that some days and places administrators want more research/teaching and others better research/teaching; some staying put and becoming established in a field and others pursuing the latest fad in topics, methods and theories; some being conservative and others radically innovative; some narrow specialists and others broad generalists. As long as the demand for such flexibility is imposed by real heterogeneity in the environment of research and education it should be profitable, but such change/variation does impose a fixed or set-up cost which, if unjustified, will detract from, rather than add to, the acquisition and dissemination of knowledge.

## 7 Conclusion

This paper has followed some principles of evolutionary ecology into some of the highways of the sociology and philosophy of science, not all of the byways. One major highway it has omitted is the whole complementary set of principles involved when evolution takes place in response to antagonists like predators and parasites instead of resources which, generally speaking, reverses things. For example if low densities relative to resources are conditions you take advantage of and high densities ones you flee, then the reverse is the case with respect to antagonists. How effort should be allocated between acquiring resources and defending against antagonists depends upon what densities, scale etc. relative to resources and relative to antagonists are, relative to each other. Science as a cultural activity has its antagonists. These include: pseudoscientists, extreme animal rights activists, and some creationists (which at times and in places it is wise to oppose). Additionally, throughout we have discussed ecological conditions and strategies appropriate to both individual growth (somatic functions, in science research) and demographic growth (reproductive or replicative functions, in science graduate and postdoctoral teaching). A second major highway omitted then is the relationship between the two. $r$  and K selection theory for example argued that low densities favour the latter and high densities the former. This is both a complicated issue in evolutionary theory and a contentious issue in science. Since in a sense it pertains to parent-offspring relations, it may be viewed as part of the sociobiology rather than the evolutionary ecology of science. In the future however, the evolutionary approach to science should be extended to the sociobiology of science including not only this, but also paper publishing and other forms of social interaction among peers.

The general principles of evolutionary ecology applied to science here are not the kind which are normally directly tested empirically in evolutionary ecology. Rather, they are general principles which logically must be true if the general theoretical framework, that culture, including that of science evolves adaptively, that it does so under ecological control, and that energy content, time and space (including niche time and space) are fundamental dimensions of ecological control. However, they do lead into a multitude of byways in which additional assumptions can be added, among which there are choices to be made, which therefore lead to different testable predictions.

While concentrating on highways rather than byways then, this paper has covered enough ground to make the case that introducing more evolutionary-ecological thinking into sociocultural evolutionism/memetics in general, and science studies in particular, is liable to be productive. Evolutionary-ecological concepts including density, scale, frequency and heterogeneity; appropriate strategies; and correlations among strategies including the role of comparative advantage suggest the direction in which explanations should be sought for such basic phenomena in science as primary and secondary research/teaching, changing topics or fields, and innovating; fast specialized and longer, slower more generalized research/teaching; avoiding competition; and the need for flexibility.

## Notes

1. Actually there are more than two. These others include the adaptive immune response in vertebrates and learning by operant conditioning (Hull, Langman and Glenn 2001). However, most would also consider these "derived" evolutionary processes as Donald Campbell once called them -- programmed at least in broad outlines by biological/sociocultural evolution.
2. Strategies included in these tables apply similarly to the individual and the demographic. In addition to density/scale conditions relative to resources summarized in these tables, strategies can be favoured if at low frequency or if their user possesses a material/genetic comparative advantage (preadaptation) for them.
3. Fog's 1997, 1999 cultural r and K selection are different, mechanism-free concepts as he indicates, see (1999), Chapter 4.
4. How resources in the denominator should be measured depends upon the details. For predators, numbers of prey might matter most while for parasites, the size of hosts might matter most for example.
5. With simple inputs for which the decision to accept or reject can be made as a whole, it might be possible to acquire and excrete at low densities (likely using a single entry and exit) and to break down and build up at high densities. Commonly however resources are more complex and need to be processed for the decision to be made. In these circumstances low densities favour growth by acquiring and breaking down; high densities favour maintenance by building up and excreting.
6. See note 4.
7. This follows from a simple (linear) equilibrium model of density and frequency dependence. Assume that strategies of growth and maintenance/motility are present in frequencies of N and (1-N) respectively at costs of C and (1-C) respectively. Let E be the amount of resources which remain freely available in the environment and (1-E) be the amount which has been absorbed into the population. For pure density dependence we set the product of the frequency and cost of one strategy divided by its total benefits equal to that of the other, i.e. at equilibrium:

$$(1) \frac{NC}{E} = \frac{(1-N)(1-C)}{1-E}$$

For pure frequency dependence the total benefits of one strategy are proportional to the frequency and cost of the other so that at equilibrium:

$$(2) \frac{NC}{(1-N)(1-C)} = \frac{(1-N)(1-C)}{NC}$$

For combined density and frequency dependence then:

$$(3) \frac{NC}{E(1-N)(1-C)} = \frac{(1-N)(1-C)}{(1-E)NC}$$

## Appendix: r And K Selection Revisited

The mathematics of density-independent and density-dependent selection are not in doubt (see any ecology text). However, what they mean in the sense of what properties we should expect of organisms adapted to these different ecological conditions, particularly the latter, has been the source of some confusion. The eat versus (roughly) assimilate/search of foraging theory and the growth versus maintenance/motility of life history theory are, each on their own scale of course, essentially the same theory. They are also essentially the same theory as that of r (density-independent) versus K (density-dependent) selection. This is because the former alternative in each of the first two cases should be favoured under low (temporary/local) densities and the latter under high. However, they also make it obvious that there is no reason to expect that low densities favour small size and fast life cycles while high densities favour large size and long life cycles as the original interpretations of r versus K selection theory had it (Pianka 1970, 1999). If anything, initially they lead one to expect sizes and life spans to be negatively associated because of the trade-offs involved. Resources spent on growth cannot be invested in maintenance (similarly for motility) and vice versa. Ultimately of course, at equal costs and allocated under their appropriate conditions, sizes should be equivalent. Similarly, the parallel growth versus maintenance (or motility) distinction with regard to the production of offspring does not correspond, as is sometimes thought, to the distinction between more numerous smaller, versus less numerous larger offspring. Quality is quite simply quality, not quantity. It does however correspond, as is also commonly thought, to a high fertility regime versus one in which discriminating parental care (or provisioning of motility) permits a higher proportion of those fewer offspring that are produced to survive. Quality students do not know more, they know better -- their knowledge is better organized, contains fewer errors and as a consequence is more enduring/useful in other fields.

In essence, the original r versus K selection theory of what was to be expected under low and high density conditions did not distinguish between uncircumscribed and circumscribed environments and tried to explain the spatial and temporal characteristics of organisms by paying attention only to the energy content of the environment and ignoring its spatial and temporal features. In section 5, the problem of explaining why size and life span are commonly positively

correlated was addressed beginning with the concepts of small versus large environmental scale and rate/density versus time/space selection.

Stably low densities were thought to be unlikely because populations were expected to normally eat and breed their way up to the carrying capacity of the environment. Hence the original theory of r selection proposed that low or a history of unstable densities favour growth strategies. A population with a history of exponential growth punctuated by catastrophes was expected to be dominated by such strategies, which is reasonable. However, it also needs to be added, whether with reference to strategies of individual or demographic growth or both, that catastrophes of similar magnitude may occur with high frequency relative to period putting a premium on the rate of individual/demographic growth. Alternatively, a long period relative to frequency puts a premium on the time through which growth is sustained. Hence the distinction was drawn between r and t selection, attributable not to density but to (temporal) environmental scale. Similar distinctions apply spatially between d (density, i.e. in the sense of mass or numbers per unit area or volume) versus s (space, i.e. area or volume) selection. Ecological correlations between temporal and spatial scales, or alternatively, correlated comparative advantages (material or genetic) are what must explain positive correlations between life span and size -- whether physical or metaphoric, i.e. niche span and size. Scientists in their research and teaching, no less than organisms in their consumption and production of offspring, should evolve to conform to the spatio-temporal scale small versus large of resources available in their environment.

The original theory of K selection proposed that maintenance strategies were favoured at high densities or with a history of stably high densities. However, just as stably low densities are unlikely because populations are liable to eat and breed their way to the ceiling; stably high densities are unlikely because populations are liable to be eroded to the floor by their parasites and predators. On the other hand, just as catastrophes can strike growing populations, bonanzas can give a boost to declining ones. Hence it is reasonable to conclude that high densities or a history of instability in the form of bonanzas should favour maintenance strategies. Bonanzas of similar magnitude, like catastrophes, can take place with high frequency relative to period/distance, putting a premium on speed of motility/density of maintenance. Or they can take place with a long period/distance relative to frequency, putting a premium on the time/space through which such are sustained.

In addition to claiming that low densities favour growth strategies while high densities favour maintenance strategies (a form of which has been embraced here for uncircumscribed environments, although motility and in some cases mutability are also options), and that low density selected life cycles should be small and fast while the high density selected should be large and slow but long (rejected here explaining organismic scale requires concepts of environmental scale), the original theory of r versus K selection claimed that the low density (r) selected devote more resources to demographic growth (e.g. in science, teaching) while the high density (K) selected devote more to individual growth (e.g. in science, research). This question which can be viewed as belonging properly to the sociobiology of science is beyond the scope of this paper.

Finally on classical evolutionary ecology, r and K selection theory had nothing to say about the other half of the life history story beyond growth/some of the 3M's, i.e. about flexibility. Levins (1968) introduced the subject of evolution in changing/varying environments. Densities are not necessarily low because of catastrophes or high because of bonanzas -- they may be both at different times/places. The scale of environmental resources is not necessarily small or large -- resources may be distributed and utilized on more than one scale. Such heterogeneity calls for effort to be devoted to flexibility which itself may be fine (change/differentiation) or coarse-grained (phenotypic plasticity).

## Acknowledgements

The author would like to acknowledge helpful suggestions from Gail Greer, from three referees and an editor, and from attendees at sessions on the sociology of science of the International Sociological Association, on evolving science and on the evo-devo problem at successive meetings of the International Association for the History, Philosophy and Social Studies of Biology, and the Duke University special conference on the evo-devo relationship to whom various aspects of this paper have been presented in recent years.

## References

- Aldrich, H. 1999. *Organizations Evolving*. Sage Publications.
- Andrews, J.H. 1991. *Comparative Ecology of Microorganisms and Macroorganisms*. SpringerVerlag, New York.
- Aunger, R. (ed.) 2000. *Darwinizing Culture: The Status of Memetics as a Science*. Oxford University Press, New York.
- Basalla, G. 1988. *The Evolution of Technology*. Cambridge University Press, Cambridge.
- Bazzaz, F.A. 1996. *Plants in Changing Environments: Linking Physiological, Population, and Community Ecology*. Cambridge University Press, Cambridge.
- Begon, M., Harper, J.L. and Townsend, C.R. 1996. *Ecology: Individuals, Populations and Communities*. Third Edition, Blackwell Science Ltd., Oxford.
- Blackmore, S. 1999. *The Meme Machine*. Oxford University Press, Oxford.
- Boulding, K.E. 1981. *Evolutionary Economics*. Sage Publications, Thousand Oaks, CA.
- Brodie, R. 1996. *Virus of the Mind: The New Science of the Meme*. Integral Press, Seattle.
- Bulmer, M. 1994. *Theoretical Evolutionary Ecology*. Sinauer Associates, Inc. Sunderland, MA.
- Buskes, C. 1998. *The Genealogy of Knowledge: A Darwinian Approach to Epistemology and Philosophy of Science*. Tilburg University Press.

- Buss, L.W. and Blackstone, N.J. 1991. An Experimental Exploration of Waddington's Epigenesis Landscape, *Philosophical Transactions of the Royal Society, London B* **332**:49-58.
- Cairns, N.J., Overbought, J. and Miller, S. 1988. The Origin of Mutations, *Nature* **335**:142-145.
- Cairns, J. and Foster, P.L. 1991. Adaptive reversion of a frameshift mutation in *Escherichia coli*, *Genetics* **128**:695-701.
- Campanario, J.M. 2000. Fraud: Retracted Articles Are Still Being Cited, *Nature* **408**: 288.
- Campbell, D.T. 1988. *Methodology and Epistemology for Social Science: Selected Papers*. University of Chicago Press, Chicago.
- Carneiro, R.L. 1970. A Theory of the Origin of the State, *Science* **169**:733-738.
- Cavalli-Sforza, L.L. and Feldman, M.W. 1981. *Cultural Transmission and Evolution: A Quantitative Approach*. Princeton University Press, Princeton.
- Chance, D.M. and Peterson, P.P. 1999. The New Science of Finance, *American Scientist* **87**:256-263.
- Charnov, E.L. 1993. *Life History Invariants: Some Explorations of Symmetry in Evolutionary Ecology*. Oxford University Press, Oxford.
- Cockburn, A. 1991. *An Introduction to Evolutionary Ecology*. Blackwell Scientific Publications, Oxford.
- Croft, W. 2000. *Explaining Language Change: An Evolutionary Approach*. Pearson Education Limited, Harlow, Essex.
- Cullen, B.S. 2000. *Contagious Ideas: On Evolution, Culture, Archaeology, and Cultural Virus Theory*. Oxbow Books, Oxford.
- Dawkins, R. 1976. *The Selfish Gene*. Oxford University Press, Oxford.
- Degler, C.N. 1991. *In Search of Human Nature: The Decline and Revival of Darwinism in American Social Thought*. Oxford University Press, New York.
- Ehrlich, P.R. 2000. *Human Natures: Genes, Cultures and the Human Prospect*. Island Press, Washington, D.C.
- Fog, A. 1997. Cultural r/k Selection, *Journal of Memetics - Evolutionary Models of Information Transmission*, **1**(1). [http://jom-emit.cfpm.org/1997/vol1/fog\\_a.html](http://jom-emit.cfpm.org/1997/vol1/fog_a.html)
- Fog, A. 1999. *Cultural Selection*. Kluwer Academic Publishers, Dordrecht.

- Foster, P.L. 1999. Mechanisms of Stationary Phase Mutation: A Decade of Adaptive Mutation, *Annual Review of Genetics* **33**:57-88.
- Fox, C.W., Roff, D.A. and Fairbairn, D.J. 2001. *Evolutionary Ecology: Concepts and Case Studies*. Oxford University Press, Oxford.
- Gill, D.E. 1974. Intrinsic Rate of Increase, Saturation Density, and Competitive Ability II. The Evolution of Competitive Ability, *American Naturalist* **108**:103-116.
- Greene, M.T. 1997. What Cannot Be Said In Science, *Nature* **388**:619-620.
- Hall, B.G. 1998. Adaptive Mutagenesis: A Process That Generates Almost Exclusively Beneficial Mutations, *Genetica* **102/103**:109-125.
- Hanski, I.A. 1999. *Metapopulation Ecology*. Oxford University Press, New York.
- Hanski, I.A. and Gilpin, M.E. (eds.) 1997. *Metapopulation Biology: Ecology, Genetics and Evolution*. Academic Press, London.
- Hawkins, B.A. 1999. More Haste, Less Science, *Nature* **400**:498.
- Horgan, J. 1996. *The End of Science: Facing the Limits of Knowledge in the Twilight of the Scientific Age*. Addison-Wesley Publishing Company, Inc. Reading, Mass.
- Hull, D.L. 1988. *Science as a Process: An Evolutionary Account of the Social and Conceptual Development of Science*. The University of Chicago Press, Chicago.
- Hull, D.L., Langman, R.E. and Glenn, S.S. 2001. A General Account of Selection: Biology, Immunology, and Behavior, *Behavioral and Brain Sciences* **24**:511-573.
- Kosso, P. 1992. *Reading the Book of Nature: An Introduction to the Philosophy of Science*. Cambridge University Press, Cambridge.
- Krebs, J.R. and Davies, N.B. 1993. *An Introduction to Behavioural Ecology*, Third Edition. Blackwell Scientific Publications, Oxford.
- Krebs, J.R. and Davies, N.B. (eds.) 1997. *Behavioural Ecology: An Evolutionary Approach*, Fourth Edition. Blackwell Science Ltd., Oxford.
- Kuhn, T.S. 1962. *The Structure of Scientific Revolutions*. University of Chicago Press, Chicago.
- Lenski, G., Lenski, J. and Nolan, P. 1994. *Human Societies: An Introduction to Macrosociology*. McGraw Hill, Whitby, On.
- Levins, R. 1968. *Evolution in Changing Environments*. Princeton University Press, Princeton.
- Lovins, A. 1976. Energy Strategy: The Road Not Taken, *Foreign Affairs* **55**:65-96.

- Lynch, A. 1996. Thought Contagion: How Belief Spreads Through Society. Basic Books, New York, N.Y.
- MacArthur, R.H. 1962. Some Generalized Theorems of Natural Selection, Proceedings of the National Academy of Sciences U.S.A. **48**:1893-1897.
- MacArthur, R.H. and Wilson, E.O. 1967. The Theory of Island Biogeography. Princeton University Press, Princeton.
- Maddox, J. 1993. Competition and the Death of Science, Nature **363**:667.
- Mann, C.C. 1994. Fire Ants Parlay Their Queens Into a Threat to Biodiversity, Science **263**:1560-1561.
- Merton, R.K. 1973. The Sociology of Science. The University of Chicago Press, Chicago.
- Mervis, J. 1993. At NSF: Fewer, Longer Grants, Science **262**:1636-1638.
- Nelson, R.R. 1996. The Sources of Economic Growth. Harvard University Press, Cambridge.
- Nelson, R.R. and Winter, S.G. 1982. An Evolutionary Theory of Economic Change. Harvard University Press, Cambridge, Mass.
- Pianka, E.R. 1970. On r and K selection, American Naturalist **104**:592-597.
- Pianka, E.R. 1999. Evolutionary Ecology, Sixth Edition. Addison-Wesley Educ., New York.
- Plotkin, H. 1994. Darwin Machines and the Nature of Knowledge. Harvard University Press, Cambridge, MA.
- Popper, K.R. 1963. Conjectures and Refutations. Basic Books, New York.
- Promislow, D.E.L. and Harvey, P.H. 1990. Living Fast and Dying Young: A Comparative Analysis of Life History Variation Among Mammals, Journal of Zoology **220**:417-437.
- Rhodes, O.E., Chesson, R.K. and Smith, M.H. 1996. Population Dynamics in Ecological Space and Time. The University of Chicago Press, Chicago.
- Ricklefs, R.E. and Miller, G.L. 2000. Ecology, Fourth Edition. W. H. Freeman and Company, New York.
- Roff, D.A. 1992. The Evolution of Life Histories: Theory and Analysis. Chapman & Hall, London.
- Rosenberg, S.M. 2001. Evolving Responsively: Adaptive Mutation, Nature Reviews Genetics **2**:504-15.

- Ruhlen, M. 1994. *The Origin of Language: Tracing the Evolution of the Mother Tongue*. John Wiley & Sons, Inc., New York.
- Runciman, W. G. 1989. *A Treatise on Social Theory II, Substantive Social Theory*. Cambridge University Press, Cambridge.
- Sibly, R.M.. and Calow, P. 1986. *Physiological Ecology of Animals: An Evolutionary Approach*. Blackwell Scientific Publications, Oxford.
- Smith, E. A. and Winterhalder, B. (eds.) 1992. *Evolutionary Ecology and Human Behavior*. Aldine De Gruyter, New York.
- Sperber, D. 1996. *Explaining Culture: A Naturalistic Approach*. Blackwell Publishers Ltd. Oxford.
- Stearns, S. C. 1992. *The Evolution of Life Histories*. Oxford University Press, Oxford.
- Stephens, D.W. and Krebs, J.R. 1986. *Foraging Theory*. Princeton University Press, Princeton.
- Toulmin, S. 1972. *Human Understanding*. Princeton University Press, Princeton, N.J.
- Watson, J.D. 1969. *The Double Helix*. New American Library, New York.
- Whiten, A. 2001. Imitation and Cultural Transmission in Apes and Cetaceans, *Behavioral and Brain Sciences* **24**:359-360.
- Ziman, J.M. 1968. *Public Knowledge: An Essay Concerning the Social Dimension of Science*. Cambridge University Press, Cambridge.
- Ziman, J.M. (ed.) 2000. *Technological Innovation as an Evolutionary Process*. Cambridge University Press, Cambridge.

## **MEMETICS AND EVOLUTIONARY SOCIAL SCIENCE**

**By**  
**Marion Blute**

Department of Sociology  
University of Toronto at Mississauga  
[marion.blute@utoronto.ca](mailto:marion.blute@utoronto.ca)

Memetics is the scientific study of memes. The term “meme” was coined by Richard Dawkins (1976), suggestive of “m” for memory or imitation and “eme” for gene, for the basic unit of hereditary information or replication involved in cultural as opposed to biological evolution. Memetics became the subject of an academic journal (*Journal of Memetics - Evolutionary Models of Information Transmission*) as well as of more than half a dozen books (Aunger 2000, 2002; Blackmore 1999; Brodie 1996; Cullen 2000; Cziko 1995; Lynch 1996). According to some critics, memetics has joined extraterrestrial, exo or zenobiology as a science devoted to a subject matter whose very existence remains in dispute! However, those who view the meme concept as more amorphous or ambiguous than that of the gene have insufficient knowledge of biology. In those long molecules of DNA, the fact that none of the units of structure, function, replication and recombination coincide means that the gene concept is similarly ambiguous (for a useful review see Griffiths 2002). Of course one might then endorse abandoning the gene concept rather than adding a meme concept but in the age of genomics there is little chance of the former at least.

In one sense, only the term “meme” was new with Dawkins. For some time some social scientists studying a wide variety of phenomena including languages, science and technology, and economic organizations and institutions for example as well as theory in general had been applying a Darwinian-style evolutionary theory of “descent with modification” involving cultural transmission, variation and selection to their respective subject matters. Among memeticists, some difference of opinion exists over whether memes should be thought of as gene-like or virus-like but both points of view have made useful contributions to evolutionary social science.

The gene-like view. Social scientists tend to wonder whether they need a common term like meme across their various subject matters. It may or may not turn out to be useful in the long run but if nothing else, the gene-like concept of a meme has reminded us that cumulative evolution of complex entities requires digitally-encoded information. Social learning by ‘observation’ in a variety of sensory modalities can support a primitive form of cultural evolution as it does among chimpanzees (Whiten et. al. 1999) and Cetaceans (Rendell & Whitehead 2001) for example. However, only social learning by ‘instruction’, employing information encoded digitally in a string of symbols, among humans normally strings of phonemes or letters in sentences in a natural language, can a) support the inheritance of a large number of possible states (Maynard Smith and Szathmary’s “unlimited inheritance” 1995) and b) avoid the cumulative degradation characteristic of analogue systems (Dawkins 1995) and hence is required to support the cumulative evolution of complex sociocultural entities.

The virus-like view. The virus-like view of memes has usefully reminded us that much, even most cultural transmission in modern societies takes place horizontally rather than vertically

relative to genes. Social scientists have always been aware of this but have not paid sufficient attention to the implication that culture therefore evolves in a fashion that may be indifferent to its ‘host’s’ biological fitness. Many memeticists emphasize that our culture, in fact, commonly parasitizes our biology. The “memplexes” that characterize membership in sects or cults, to cite an extreme example, often ruthlessly exploit the biology of their members, convincing them with mythical appeals to “family” to isolate themselves from their real families and dedicate their all to the survival and spread of the social identity of cult membership. This virus-like view suggests the application of basic principles of the epidemiology of infectious diseases (e.g. Ewald 2000) to culture to predict the conditions under which it will evolve to be benign or virulent relative to genes. According to those principles culture (like infectious diseases) with vertical transmission, a low multiplicity of infection, and infection by means of direct contact will tend to evolve to be more benign. By contrast, culture which is transmitted horizontally, with a high multiplicity of infection, and infection by means of vectors will tend to be more virulent. Assuming you care more about your biology than your culture (which is not necessarily the case - ‘you’ after all are a combination of both), practical lessons, particularly for teenagers, emerge from memetics. Listen more to mommy and daddy and less to your friends! Beware more of fads and fashions which can infect you multiply than of whole social identities like ethnic, religious and occupational identities. One of these latter normally precludes another and hence may be willing to leave something of your biology for itself to live on tomorrow! And finally, trust information conveyed personally rather than via mass media which, like insect-borne diseases, can get to you even when you are down and unable to circulate!

The future. According to evolutionary epistemology (Campbell 1970), universal Darwinism (Cziko 1995, Dennett 1995), or multi-process selection theory, selection processes are selection processes and the same general principles should apply whether realized biologically (gene-based evolution by natural selection), socioculturally (meme or social learning-based sociocultural evolution by sociocultural selection), or psychologically (neural-based learning by reinforcement and punishment) for example (Hull Langman and Glenn, 2001). Hence there is no point reinventing the wheel. For principles applicable to selection processes, memetics can benefit from a close study of evolutionary biology and related disciplines like taxonomy and evolutionary ecology. Taxonomy dealing with history, and evolutionary ecology dealing with selection are the two most important research programmes in biological evolution. Duthie (2003) in his pithy essay on the fork and the paperclip got it exactly right - the key questions are where did something come from and why did it evolve. Those who hope to make a contribution to memetics in the future require knowledge of one or more topics or fields in biological as well as in a social science discipline or disciplines. It has become well known whether by philosophers (Bunge 2003) or by granting agencies like the U.S. National Science Foundation that multi, inter and transdisciplinary research (“MIT” in science studies jargon) is where the breakthroughs are made.

This should not be taken to imply that the study of selection in other realms need be wholly derivative of the biological. There is much extant research but still much to be done in understanding the mechanisms of inheritance in sociocultural/memetic evolution (for an overview of social learning in animals see Heyes and Galef, 1996). There is much to be done in understanding how genes and culture coevolve in interaction with each other (Durham 1991 remains the most extensive treatment to date). There are many unsolved problems in biological

evolutionary theory and solutions to their analogues in the sociocultural realm could conceivably, in the future, feed back in the other direction. For example two important problems in the forefront of evolutionary biology today are the three-cornered relationship between evolution, development and ecology ‘evo-devo-eco’ and the problem of ‘origins’. Everything which evolves under ecological control also develops under ecological influence and that includes the sociocultural. Artifacts and social identities for example not only evolve but also have a life course. Should the story and theory of life or culture in the context of origins therefore begin with a single ‘juvenile’ developing (implying that evolution developed) or with many ‘adults’ replicating (implying that development evolved) or somewhere else entirely?

Moving from recommendations to forecasting I think there is every reason to believe that the future of memetics/sociocultural evolution (understood broadly as the same enterprise) is bright. Many others think likewise (e.g. Mesoudi, Whiten and Laland 2004) although for reasons that I find difficult to understand, some still sometimes look with disfavour on the ‘meme’ word while freely using the similarly ambiguous ‘gene’ word.

## References

- Aunger, Robert. (Ed.) 2000. *Darwinizing Culture: The Status of Memetics as a Science*. Oxford University Press.
- Aunger, Robert. 2002. *The Electric Meme: A New Theory of How We Think*. The Free Press.
- Blackmore, Susan. 1999. *The Meme Machine*. Oxford University Press.
- Brodie, Richard. 1996. *Virus of the Mind: The New Science of the Meme*. Integral Press.
- Bunge, Mario. 2003. *Emergence and Convergence: Qualitative Novelty and the Unity of Knowledge*. University of Toronto Press.
- Campbell, Donald T. 1970. “Evolutionary epistemology.” Pp. 413-463 in P.A. Schlippe, ed. *The Philosophy of Karl R. Popper*. Open Court.
- Cullen, Brian S. 2000. *Contagious Ideas: On Evolution, Culture, Archaeology, and Cultural Virus Theory*. Oxbow Books.
- Cziko, Gary. 1995. *Without Miracles*. MIT Press.
- Dawkins, Richard. 1976. *The Selfish Gene*. Oxford University Press.
- Dawkins, Richard. 1995. *River Out of Eden: A Darwinian View of Life*. BasicBooks.
- Dennett, Daniel C. 1995. *Darwin’s Dangerous Idea: Evolution and the Meanings of Life*. Allen Lane.

- Durham, William H. 1991. Coevolution: Genes, Culture and Human Diversity. Stanford University Press.
- Duthie, A. Bradley. 2003. "The fork and the paperclip: a memetic perspective." *Journal of Memetics - Evolutionary Models of Information Transmission*. 8.
- Ewald, Paul W. 2000. Plague Time: How Stealth Infections Cause Cancers, Heart Disease and Other Deadly Ailments. The Free Press.
- Griffiths, Paul E. 2002. "Lost: one gene concept. Reward to finder". *Biology and Philosophy* 17: 271-283.
- Heyes, Cecilia M. and Bennett, G. Galef Jr. (Ed's.) 1996. Social Learning in Animals: The Roots of Culture. Academic Press.
- Hull, David L., Langman, R.E. and Glenn, S.S. 2001. A general account of selection: biology, immunology, and behavior. *Behavioral and Brain Sciences* 24:511-573.
- Lynch, Aaron. 1996. Thought Contagion: How Belief Spreads Through Society. Basic Books.
- Maynard Smith, John and Szathmary, Eors. 1995. The Major Transitions in Evolution. W.H. Freeman and Company Limited.
- Mesoudi, Alex., Whiten, Andrew., and Laland, Kevin N. 2004. "Is human cultural evolution darwinian? Evidence reviewed from the perspective of the Origin of Species." *Evolution* 58:1-11.
- Rendell, Luke and Whitehead, Hal. 2001. "Culture in whales and dolphins." *Behavioral and Brain Sciences* 24:309-382.
- Whiten, Andrew J., Goodall, W.C., McGrew, W.C., Nishidas, V., Reynolds, Y., Sugiyama, C.E.G., Tutin, R.W., Wrangham, R.W., and Boesch, C. 1999. "Culture in chimpanzees." *Nature* 399:682-685.

# **THE CONSTRUCTION OF IGNORANCE AND THE EVOLUTION OF KNOWLEDGE**

**(Working Paper)**

**By  
P. Bouissac**

paul.bouissac@utoronto.ca  
University of Toronto

During the last few decades, scientific research conceived as the empirical pursuit of truth and objectivity has been put on trial. For many, scientific knowledge has lost its prestige, integrity and quasi transcendence, and has come to be variously considered as a set of rival incommensurable theories similar to ideologies and religions, unstable configurations of power and authority prone to irrational changes, or even fabulations, fictions generated by self-delusion, if not deliberate deception. Landmark works in this critical process include Kuhn (1970),

Popper (1972), Feyerabend (1975), and Lakatos (1978): their criticisms have triggered many debates (e.g. Suppe 1977, Serres 1989) and have permeated our "*épistémé*" in more or less radical forms. Scholars from all fields have picked up the momentum and transferred these arguments from the natural to the social sciences, and, more generally, have questioned any intellectual endeavor inspired by the modern scientific ethos or modeled on the empirical method. The notion of knowledge as revealed or discovered truth has shifted toward a view which holds that knowledge is constructed, if not fabricated. Trivial versions of this new skepticism have filtered through the general population via the media, and, according to some, explain the marked decline in government funding for scientific research which characterized the 80s. In an article-manifesto, prominently published in Nature, two physicists, T. Theocharis and M. Psimopoulos (1987), called the attention of the scientific community to this state of affair, noting that the crisis of confidence had reached even the core of the scientific establishment itself and that this fact could be ominous.

They list the dangerous consequences which this critical movement might entail for society. Even if some aspects of their counterattack, which blends theoretical and pragmatic arguments, may sound naive at times, their warning should not be taken lightly, since, indeed, "skepticism and nihilism were not threats only in the remote past; nor do they now constitute a risk of only academic interest" (598). Confronting the issues raised by this epistemological debate is not only an intellectual challenge, but it also has political and ethical dimensions. Efforts toward a critical understanding of the scientific development of knowledge, including its sociological and rhetorical strategies, do not necessarily lead to a radical dismissal of its value and virtue. It may, however, help reassess its significance in the broader theoretical context formed by the interfaces which are presently developing between the humanities and the social sciences on the one side, and the formal and the natural sciences on the other.

## 1. Objects of perplexity

Knowledge has always an object. Prior to becoming an object of knowledge, was this object an unknown entity or agency? Is it instead a fallacy, a total or partial creation of knowledge itself? Are descriptions and explanations only means toward the fabrication of world views which legitimize political economies? Are experience and experimentations mere delusions, or do scientific methods of investigation yield the knowledge of universal laws which account for the qualities, quantities and behaviors of particular objects of perplexity which entice humans' minds into a quest for their knowledge? Or, are scientific institutions ways of building fictive objects on a grand scale, inventing problems and thriving on tax payers' money?<sup>2</sup> Or is science the patient and eventually successful confrontation of human intelligence with elusive objects, functions and laws whose discovery gives earthlings the power to act upon their environment and exactly predict the outcome of their actions?

Many of these questions have, at one time or another, inspired philosophical debates and political controversies over approximately two millennia of epistemological anxiety. They become all the more pressing in times of scarcity now that the scientific institutions specializing in the production of knowledge have developed close associations with military and economic power structures; contemporary society cannot afford to make hasty decisions because the stakes are so high that acting lightly in such matter amounts to playing Russian roulette. This is why, beyond the pure pleasure of theorizing, a reflexive enterprise focusing on these problems is in order.

Obviously, the first relevant question to ask bears upon the status of the object of knowledge. This issue has triggered intense discussions particularly in relation to experimental research (e.g., is the object an artifact?) and to the social sciences (e.g., is the object a function of the position of authority of the investigator?). Some problems in the latter domain will be now examined as a starting point for the argument I want to develop in this article.

One may ask for instance: Which method is most appropriate for reaching an accurate description and a satisfactory explanation of a socio-cultural object? This question suggests that, on the one hand, institutions and their productions stand as objects and that, on the other hand, methods are commodities among which a discriminating researcher can select the one which is best suited to the task at hand, given the assumed nature of the object. But such naive positivism tends to ignore the extent to which methods and objects are intimately associated, almost indistinguishable from each other like mutually definable correlates. Indeed, any method presupposes a model of its object, conceived in terms of the sort of information that the method is conceptually equipped to detect, if not to construct. Methodological strategies necessarily imply the elaboration of their object of inquiry within a more general conceptual scheme, and endeavor to determine the range of variables set forth by the model that is thus elaborated.

For example, analyzing public performances in general as acts of communication presupposes a pre-determined range of variables and their relations combined in a model. The model is the hypothesis which constrains the method as much as the method sets limits on the set of admissible models. The performance-as-communication model already contains the heuristic categories of commutable elementary units, syntactic structures and discursive strategies that the method will specify in accordance with the principles of functionalist and structural linguistics

and its theoretical extrapolations. In other words, the method makes the implied model of its object explicit and determinate, a process which may create the illusion that an autonomous domain of inquiry has been fully explored and revealed as a result of the judicious application of analytical procedures of a particular sort to some 'raw' material.

Is this to claim that all methods are equally valid or equally vain, and that the quest for knowledge proceeds through a forest of epistemological fantasies? Are researchers coming from various theoretical horizons doomed to confront one another in a conflict of interpretations? Is the cognitive interplay between models and methods necessarily condemned to be circular, thus locking researchers, individually or as groups, in epistemological solipsism? This is, of course, a possibility, and, as pointed out above, contemporary polemical debates, focused upon epistemological relativism, have called attention to the extent to which theories, scientific or otherwise, are parts of larger social, economical, cultural and ideological contexts which change over time under the pressure of sets of conditions which are so complex that they may appear at times as random and irrational processes.

However, the constant shifting of models or conceptual schemes does not suffice to justify the radical epistemological relativism of those who hold a purely 'historicism' view of knowledge (D'Amico 1989). It can be expected indeed that empirical information always exceeds the formal object of an inquiry; the method can perform its sifting and organizing functions only to a certain extent. The case can be made that there is a threshold beyond which the researcher can no longer ignore the accumulation of residue or noise. In spite of their inner beauty and logical consistency, models are eventually overcome by their insoluble surplus. Intellectual probity or a mere concern for pragmatic efficiency then motivates the restructuring of the initial model or the constituting of the residue itself as an object of inquiry with its own correlative model and method.

When the descriptive and explanatory power of a model is strained by the number of exceptions to the rules it generates or by the shrinkage of the domain of experience to which it applies, other attempts at modeling are undertaken. The challenge, however, remains to conceptualize the relationship among alternative or successive models, especially if it appears that the production of noise results from the very nature of models. The belief that socio-cultural phenomena can be exhaustively represented in terms of a unified theory may help one sustain the effort required to conduct systematic inquiries; but that remains a utopian goal.

Heuristic blindness can be productive only to a certain degree. The crisis of representation has brought to the awareness of many researchers the inherent artificiality not only of representational devices, but also of the objects which are constructed through representation along the investigative process, and which can acquire an inhuman dynamic of their own. This awareness can indeed take the form of a deep moral crisis when it turns out that human agents are thus objectified as 'others' and ultimately construed as mere tokens in some academic power game based on social predation. Ethnographic information can be indeed gathered as a mere commodity without consideration for the political conditions which make such an endeavor possible or for the consequences it may have locally.

The reflexive anthropology of the 1980s (e.g. Clifford and Marcus 1986,

Marcus and Fischer 1986) has forcefully voiced both ethical and epistemological problems that cannot be glossed over and are made all the more pressing at a time when anthropological pursuits, driven by the pressure of evolving geo-political circumstances, focus on the anthropologists' own society, and encounters most revealing challenges (Jackson 1987).

But where does epistemological despondency lead if not to a dark age of anomie, both intellectual and social? In response to the predicaments of knowledge, one may want to devise less arrogant approaches than the one with which positivism is usually associated.

## 2. The specter of the object

It is possible that the intractable problems resulting from a single, exclusive focus partly dissolve when frames and roles are systematically shifted. This alternative is at least worth exploring. The resistance to the constructed objects which manifests itself in the form of parasitic and haunting noise and the subsequent drifting of theoretical elaborations suggest some sort of refraction which calls for a spectroscopic approach.

Thus, hypothetical objects are not decomposed in terms of the analytical units or functions of a single model, but instead the targeted domain of experience is reconstructed in terms of differential, serial, and interfacial theoretical objects and subjects.

Interfaces may be relative, arbitrary, artificial and discrete to a degree, but they are nevertheless bounded within a continuum limited only by the horizons of experience. If, furthermore, these horizons are conceived as defined by the experience of a community of 'researchers' and 'researched' rather than by an individual's or school's point of view, a more comprehensive object of knowledge can be constructed that is at the same time complex, ordered and real. But before reaching this intellectual and moral grasp of an object that is both diverse and consistent, it is necessary to construct a multiplicity of interfaces. This approach might, of course, raise the specter of a fractal infinite regression, except that it does not intend to be exhaustive and can easily admit that interfaces are, to a degree, historically conditioned as the trans-object itself necessarily is.

A way of conceptualizing the complexity of a socio-cultural phenomenon as a domain of experience is indeed to model it as an array of interfaces without any suggestion of a principled hierarchical system. An interface is usually defined as a surface forming the boundary between adjacent spaces of whatever nature and consistency these spaces may be. Technically this surface equally belongs to the two contiguous domains but has also some of the properties characterizing each. The concept of interface can be further abstracted and recast in semiotic terms as the plane of contact between two systems modifying each other through some form of symbolic interaction. The contact between a liquid and a solid, which is one of the most striking examples of an interface, is only a particular case. There are also interfaces between magnetic fields as well as between deictic systems of signs such as those defining the boundary between two ethological territories. In fact all communication processes imply an interface between two or more information sources or fields. Interfaces may indeed overlap and probably necessarily do so. Social groups, marked spaces, fences, social distances, face-to-face interactions, spectacles

form interfaces whose correlated and synchronized modifications can be described as long as these interfaces are temporarily and heuristically construed as closed systems. All social realities can be seen as structured, or rather textured, by complex systems of interfaces on various levels of inclusiveness.

When particular interfaces are constituted as objects of inquiry, configurations of factors are put in place and hierarchical, or at least structural, orders define the objects of inquiry and determine their method. Both descriptive procedures and explanatory interpretations can then be developed in a manner congruent with the interfaces that have been thus constructed.

However, it must not be forgotten that artificiality, or rather 'artifactuality', belongs indeed to the nature of interfaces. It would be misleading to conceive of interfaces as 'slices' cut through the 'real' of the institutions and their productions, and closely examined with a magnifying device as if they were biological tissue. The construction of an interface involves, on the one hand, the recognition of a zone of contact between two distinct domains and its methodological isolation and, on the other hand, the formulation of a hypothesis or model which determines which sort of information will be collected and organized as being constitutive of these conceptualized zones. This approach basically relies on a theoretical systematization of the natural cleavages of socio-cultural experience with respect to societies, groups, generations, sub-cultures, corporations, specialties, and so on. But these cleavages are so numerous, complex, multivalent, fractal, fluctuating and chaotic that any methodological determination of an interface can only be a grossly simplified model. However, building series of such interfaces may bring some measure of systematicity and coherence to the notion of 'thick' description which Geertz (1973) borrowed and adapted from Ryle (1971) and which seems to be the only possible response to the engrossing, overwhelming object that any socio-cultural phenomenon form as a domain of experience.

Thus, one may be led to conclude that it is not because an object of inquiry is constructed that the research process does not yield any amount of objective knowledge, as long as the construction process itself is taken into account and forms an explicit part of the knowledge thus acquired.

However, this perspective, which may help clarify the status of objects of knowledge as processed domains of experience, remains largely tautological and strictly pragmatic. A further questioning of the strategy at play in this process may contribute to a better theoretical understanding of the epistemological problems which are being addressed in the current crisis.

### **3. An architecture of holes, gaps and lacunae**

Some fifteen years ago, Duncan and Weston-Smith (1977) edited a dense volume of collected articles under the tantalizing title *The Encyclopedia of Ignorance*. Prominent specialists - from Great Britain for the most part - in a wide array of scientific domains had been invited to write about the yet unanswered - or unanswerable - questions in their respective fields of inquiry. The resulting texts provide a most enlightening vista on the process through which knowledge is constructed in the sciences. Far from radiating unlimited confidence in their ability to conquer the unknown as if it were an object to be mapped, measured and packaged, the contributors to this volume carefully delineate their own area of ignorance and emphasize the crucial part that

this process plays in their pursuit. Each one relies on a particular rhetoric to express this essential absence of knowledge. For example, some evoke a daunting horizon of increasing complexity, some wonder whether the human brain can cope at all with this complexity beyond the requirements of extended survival strategies, some even question the very possibility of knowledge.

For instance, Lehman, a computer scientist, declares in conclusion of his contribution: "Total knowledge, the final state, can never be reached. Ignorance must always be present" (354). This should be understood not as a romantic declaration but as a technical, almost clinical constatation. It raises an interesting question regarding cognitive and discursive strategies through which scientific knowledge is constructed. Usually, the cumulative process of science is metaphorically apprehended as a 'positive', ever increasing sum. However, it also could be conceived as the 'negative' - in the photographic sense of the term - of an activity geared toward the relentless construction of ignorance, an architecture of holes, gaps and lacunae, so to speak. This might sound paradoxical, but The Encyclopedia of Ignorance reveals a consistent strategy which consists in conceptualizing what is not known, not in a vague, open-ended manner, but with application, precision, almost industriousness. It appears that the exercise did not seem strange to these scientists for the good reason that setting up what could be called "stages of uncertainty" is the very stuff of their professional existence. If, indeed, knowledge is information, its condition of existence and its real measure is uncertainty, and therefore depends on the construction of ignorance. There is no possibility of information if there are no alternatives with various degrees of probability. The more unexpected, the greater the discovery; because the uncertainty is resolved by a choice, which was not ever seriously, taken into consideration when the most probable alternatives were laid out in the experimental design. Such discoveries suddenly cause a reframing of ignorance because alternatives previously considered to be equally probable in view of a given knowledge become improbable and, thus, the presupposed knowledge which had led to this state of affair must be retroactively discarded and replaced by new questions which redefine the horizon of ignorance.<sup>3</sup>

Naturally, "science" or "scientific knowledge" as an entity is a rhetorical device. A grammatical subject which obfuscates the politics of institutions, somewhat like "life" covers the swarming and teeming of myriads of competing organisms. Academic disciplines themselves are dramatis personae on the political stages which appear whenever a critical mass of researchers agree on the definition of a domain of ignorance and on 'legitimate' methods to perpetuate this ignorance by a constant renewal of the stage of uncertainty. It should be clear at this point that 'ignorance' is not taken here in the absolute sense of self-ignored ignorance, i.e. an absence of knowledge caused by the absence of a stage of uncertainty in the representational system of an organism. The ignorance to which it is referred here is more akin to the state of knowledge of a sleuth whose reasoning is portrayed at the beginning of a crime fiction than to the situation of someone who would have been always immersed in total darkness. Scientific disciplines need decidable uncertainties but need them forever, so that the narrative will continue.

State of the art articles as a genre always are structured by what could be called a dynamic of lack. The Khunian paradigms need solvable but indefinitely expanding puzzles. In as much as these puzzles are expressed in the form of discourses, they are rhetorical artifacts. But, by contrast with the

Humanities and some sectors of the social sciences which thrive by staging undecidable uncertainties, the empirical sciences construct decidable ones by specifying in each case what will count as information, and which new ignorance this information will help formulate. From a general point of view, it seems impossible to conceive of a discipline which would consist of the mere transmission of a completed knowledge. A discipline which would fail to generate information, i.e. which would have exhausted its capacity to construct ignorance, would quickly disappear. The necessary dynamic must come from the construction of a relatively determined lack of knowledge. Therefore disciplines can be viewed as generators of uncertainty and can be compared with each other from this point of view. As it has been hinted above there are many forms of ignorance depending on the generator's model. Each discipline seems to be based upon a different blueprint, or a program (in the computer sense of the term) which specifies how to construct its *sui generis* uncertainties. But such programs are necessarily in competition with each other because they need hardware (i.e. human and economic resources) in order to be implemented.

An illustrative case is semiotics, a discipline-to-be which did not succeed in achieving disciplinary status, although it developed a dynamic toward this goal. It is symptomatic that the two 'founding' programs, i.e. those of Peirce and Saussure, essentially consist of discursive fragments about some unknown to be discovered; both evoked a relatively well specified horizon of ignorance but fell short of determining with optimal quality the generative model which could have yielded a self-replicating puzzle. As a result, the stage of uncertainty which defines semiotics remained populated by undecidables which can generate indefinite discursive productions, but finds themselves in competition with more powerful programs of uncertainties such as those of philosophy and literary interpretation. This state of affairs is appropriately expressed through the convoluted rhetoric of undecidability in the concluding paragraph of an essay by a representative semiotician: "A final remark, which is for obvious reasons inconclusive: From the preceding propositions one might get the impression that this program of enquiry will result, all in all, in a referential tautology. Well, it could also be said - precisely! Giving due regard to what M. Eigen (1975) has stated at the outset of his influential article on "Evolutionary Games", let us take in, henceforwards [interdiscursively] the basic rule of reconstruction as he puts it: "The origin of life [intertextuality] is tautologous with the origin of biological [textological] information." Therefore, ours the task, may be eternal." (Ruprecht 1991:73)

It is, indeed, undoubtedly so, but by design rather than because of an epistemological curse, as post-modern romanticism would have us believe. Naturally, "design" is taken here in the sense of "structural feature", and not psychological intentionality.

The same process of constructing a form of ignorance is observable across the whole array of modern disciplines which are thus provided with the dynamic needed for evolving and persisting over time. But each discipline is defined by its particular rhetorical strategies and modalities of implementing the perpetual reconfiguration of its stage.

At this point, one may be inclined to conclude, as some do, that knowledge production is a mere rhetorical trick, that special terminologies are ways of obfuscating tautological statements, and that the notion of scientific progress is nothing but an ever changing narrative used to legitimize

an otherwise frivolous game of power and confrontations. However, an argument to the contrary can be extrapolated from evolution theory - not as a dogmatic derivation but as a way to account for the fact that, indeed, some forms of constructed ignorance and the information it yields often make a difference for the organisms engaged by this process.

Let us recall that evolutionism is, briefly, the theory which holds that selection operates on random variations. It is a non-teleological process. Adaptation is not conceived as a goal-oriented change but as a serendipitous advantage yielded by a particular transmissible modification of an individual phenotype. For an advantageous variation which contributes to evolution, there are countless detrimental ones leading to early elimination of those genotypes from the genetic pool. Moreover, what is advantageous in a particular environment, can become a liability in a different one or if the initial environment changes. At the cost of a metaphoric translation, strategies which elicit information can be conceived as phenotypes implementing genotypic variations whose outcome may, at times, confer a decisive advantage through a close fit with environmental constraints. There are actually no good reasons for excluding rhetoric from the domain of natural phenomena, nor for explaining it in merely functionalist terms.

But, having set the theoretical perspective which will be further discussed in the next section, let us return for a while to the strategic construction of ignorance and let us examine what 'constructing ignorance' practically means. It consists of building a model, generally of an interface, such that some relations constitutive of the model remain indeterminate. For instance, in Duncan and Weston-Smith (1977), Barlow, a specialist of physiological optics, addresses the topic of the languages of the brain: "The language for communication between the various higher parts of the brain must use symbols with an agreed meaning, for otherwise one part would not understand what another part said. We do not know what these symbols are, nor how their meanings are agreed to. However, the material discussed in this essay does enable us to hazard a few guesses about the general nature of the language, and here are three of them. (1)

Whatever a symbol may correspond to in the external world, internally it is likely to be represented by nerve impulses in a nerve cell or in a family of nerve cells, for we know of no other tokens that can be rapidly exchanged between the parts of the brain. (2) There is nothing to indicate that these internal symbols occur in a single serial string [...] (271-271). This suffices to make clear that the model construes "part of the brain" as agencies which interact through communication, i.e. transfer of information; while there is evidence that such interactions occur, the mode of communication is not known. This uncertainty is, however, reduced to a few alternatives which are introduced as the stage of uncertainty within the model itself. The narration is familiar: the sleuth has evidence that two conspirators communicate through a secret code. Since they live in different cities, the messages have to be carried. By whom and in which form are these messages transmitted? and so on. This was the state of the art fifteen years or so ago. Neurotransmitters are now better known, but this new knowledge creates its own horizon of uncertainty.

#### **4. Epistemology: from creationism to evolutionism**

It seems that, at this point of the reflexion, at least two assertions can be stated which would be supported by a reasonable consensus. First, the process of scientific observation necessarily

implies some form of theoretical presuppositions, hence the logical pre-eminence of theory over scientific observation and experimentation. This argument has been put forward and convincingly argued by Popper (e.g. 1972:340-361) and the counter argument that theories can be made obsolete by disproving data does not establish the reverse. It simply shows that a theory can be replaced by another. This leads to the second assertion, namely that theories are entities in competition with each other. Some are short lived or remain confined to a small circle of minds; others spread among much larger groups of thinkers, researchers, and teachers; some gain the status of worldviews, or commonsense, by taking hold of whole populations. The paradigm shifts described, but not explained, by Kuhn, dramatizes this eliminative process by evoking some "irrational" forces.

If it is true that, as has been suggested by the argument developed in the second and third sections of this essay, theories are differentiated ways of constructing uncertainty, and correlatively, of producing information, it could be concluded that they share an important, if not critical, feature with living organisms. This line of argument draws from current assumptions made in artificial life (AL) research, holding that the "logical form" of an organism can be separated from its material basis of construction, and that "aliveness" will be found to be a property of the former, not of the latter" (Langton, 1991:11). The way, in which theories come and go, prosper and disappear, indicates that, although they are generally considered to be produced by human brains, they are endowed with a relative degree of autonomy with respect to the populations of organisms among which they spread with various degrees of success. Traditional philosophies and religions have developed an ontology of ideas, narratives, instructions and rituals which take such autonomy for granted (e.g. Plato's ideas; divine words embodied in sacred texts; rituals identified as divine beings in Hinduism). Moreover, secular poets, philosophers and scientists have often referred to the *sui generis* dynamism of ideas or theories as quasi agencies independent from their own psyches.

Metaphorical expressions in many languages emphasize this puzzling impression: one comes across an idea, one may yield to the power of a worldview or a system, one may be inspired, possessed, turned on by a theory, and so on. Considering the way in which myths seem to unfold and transform their structures as they drift across continents from population to population, Levi-Strauss once noted that "The Ojibwa Indians consider myths as 'conscious beings', with powers of thought and action" (1969:12, note) and claimed that his goal was "to show not how men think in myths, but how myths operate in men's minds without them being aware of the fact. (ibid:12) Someone can also identify with a particular theory and passionately fight for it - many humans have died for a worldview and still do - and, sometimes, can as easily switch to another one without noticeable detrimental effect on his/her organism and mental life. Ideas, faith, commonsense, belief can be shown to share features with theories, even to be theories under other names. Theories have no particular regard for the welfare of the populations among which they may thrive. They may even conceivably go extinct because "their" population disappears, precisely because of some ill effects of the normative behavior they consistently promote in the population. On the other hand, some theories may have some long term beneficial side effects on "their" populations. But it does not seem that as an information structure a theory as such could have either a benevolent or malevolent intent. It simply spreads as fast and as much as possible, eliminating competing theories in the process, sometimes softly by causing them to be forgotten, sometimes violently by physically suppressing the populations which hold them.

Naturally an extended metaphor cannot be used as an argument. Common sense associates theories so closely to the mental lives of their generators that such evolutionary analogies probably will be considered as nothing but playful speculations. However, it could also be pointed out that this perceived indissociable association, this identity of an information-producing structure and the brain which implement it, might be just a theory among others. But, then, it might be opposed to this line of thought that, according to a well known principle, one is not entitled to demonstrate a theory by this theory itself, except of course if this new theory had such a powerful information producing structure that it would "entropize", so to speak, Gödel's argument.

The notion of evolutionary epistemology goes back to the end of the 19<sup>th</sup> century. But its development remains anthropocentric. The survival rate of theories depends on their usefulness to humans; in fact they are closely related to the fitness of particular individuals. They are part of the phenotype. Popper, however, who discussed the issue on several occasions (e.g. 1972:65-70) perceived the inherent problems attached to this view: "[...] questions of truth or validity, not excluding the logical justification of the preference for one theory over another (the only kind of 'justification' which I believe possible) must be sharply distinguished from all genetic, historical, and psychological questions. [...] logical investigations of questions of validity and approximation to truth can be of the greatest importance for genetic and historical and even for psychological investigations. They are in any case logically prior to the latter type of question, even though investigations in the history of knowledge can pose many important problems to the logician of scientific discovery". (1972:67-68).

This is a way of saying - albeit from another theoretical viewpoint than the one proposed here - that organisms and theories are entities of different order. But the order which is constitutive of Popper's theory implies a hierarchy whereas "order" will be taken here in its taxonomic sense.

Although the evolutionary selection metaphor applied to theories appears often in the epistemological literature of the last two decades, it is not until the bold, albeit tentative introduction of the concept of "meme" by Dawkins (1976:206) that a new way of thinking evolutionary epistemology could be proposed. Let us recall the initial, somewhat awkward formulation of this hypothesis: "Just as genes propagate themselves in the gene pool by leaping from body to body via sperms or eggs, so memes propagate themselves in the meme pool by leaping from brain to brain via a process which, in the broad sense can be called imitation. If a scientist hears, or reads about a good idea, he passes it on to his colleagues and students." The term "meme" was somewhat playfully coined after the French "même" for "same" and with phonetic allusions both to "mimetic" and to "gene". In the paragraph following this introduction of the new concept, Dawkins quoted the summary that one of his colleagues, N.K. Humphrey, produces in reaction to the reading of an earlier draft of this chapter: "memes should be regarded as living structures, not just metaphorically but technically. When you plant a fertile meme in my mind you literally parasitize my brain, turning it into a vehicle for the meme's propagation in just the way a virus may parasite the genetic mechanism of a host cell. And this isn't just a way of talking - the meme for, say, "belief in life after death" is actually realized physically, millions of times over, as a structure of the nervous systems of individual men the world over." (207) In his subsequent publications Dawkins returns to the concept of "meme" (1982: 97-117, 1986:158)

and indicates attempts made by others to develop his hypothesis (e.g. Delius, 1990) which has also been taken up since by Dennett (1991).

The idea of entities endowed with both genotypic and phenotypic structures independent from the biological ones but evolving in a parasitic relation to them, remains an intriguing and fascinating metaphor which shatters our world view as if it were an aftershock of Darwinism but is too "odd" to be taken seriously. Dawkins himself has been cautious in his subsequent and brief treatments of the "meme" idea, as if he were daunted by the full implications of this emergent theory. It is worth noting in passing that "memes" are strikingly different from the "culturgens" conceived by Lumsden and Wilson (e.g. 1981:27, 368) which remained modeled on functional artifacts and the process of cultural inheritance. Even Boyd and Richerson (1985) for whom "culture is a system of inheritance [...]. Today's cultural traditions are the result of cumulative changes made by past and present bearers of them" (20), shy away from the full implications of Dawkins' views. They state: "Our definition of culture is not at all specific about the nature of the information that affects phenotypes. In particular, we do not assume that culture is encoded as discrete "particles [...]. Relatively little can be said on this topic since our knowledge of the neurophysiology of social learning is primitive compared to our knowledge of the molecular biology of the gene" (37). However, when they address the issue of "horizontal" transmission, i.e. transmission within a generation rather than across generations, the analogy with parasitic phenomena naturally comes to their minds: "Horizontal transmission is analogous in some ways to the transmission of a pathogen [...]. The item of culture being spread horizontally acts like a microbe that reproduces and spreads rapidly because it is "infective" and has a short generation length compared to the biological generation length of the "host. Fads and fashions and technical innovations are familiar examples". (8) The epidemiological metaphor is also used by Cavalli-Sforza and Feldman (1981) in their quantitative study of cultural transmission.

As long as the metaphor is constrained by the model of a biological parasite such as a bacteria or a virus, the analogy remains tantalizing but non-operational. This seems to be the source of Dawkins' reluctance to proceed further with his yet undeveloped "meme" theory. However, another analogy may come to the rescue. A developing sector of computer science focuses on information structures whose behavior within computer environments resembles so much the behavior of viruses within a biological milieu that the concept "virus" or its colloquial, more benign form "bugs" has gained wide currency, and has become the object of intensive scrutiny. Cohen (1990, 1991) has fully explored the issue and interestingly emphasizes that, although computer viruses have become associated with malevolent intent in the form of criminal or mischievous artifacts, they are, like their biological equivalent, two-edged swords: "the features that make computer viruses a serious threat to computer integrity can also make them a powerful mechanism for the reliable and efficient distribution of computing resource" (199:23). But this is not all; not only are computer viruses subject to mutations, but also some of these mutations may happen to be beneficial as Cohen shows with the case of the Morris' Internet virus. The point which seems to be most relevant here is that computer viruses can provide an operational model for "memes" conceived as programs endowed with their own dynamism, albeit of a non-biological nature (see Langton 1991, mentioned earlier). It is therefore possible to conceptualize theories, under any other names, as information structures (hence the centrality of uncertainty in the dynamics of their implementation), whose relation to the environment within which they

replicate is double-edged. In the same ways as a genetic random mutation may by chance yield an advantage for the organism which carries it as a part of its phenotype, an information structure variation (which can be a copying error, a misunderstanding or a misspelling) can spread through a population both horizontally and vertically and yield an advantage (adaptability) for this population. This theoretical view, which seems to provide a first step toward a truly evolutionary theory of culture, could explain many features of cultural and scientific evolution: why cultures and knowledge are only relatively cumulative; why cultural traits are not necessarily beneficial to the populations within which they thrive; why theories constantly displace each other but sometimes make temporarily irreversible differences in terms of the survival of the populations which carry them. The perspective which is thus opened might succeed in "displacing" or "entropizing" the creationism which remains implicit in traditional theories of knowledge, whose current crisis may represent a conflict of evolutionary significance.

## References

- Abragam, A. (1985) "Théorie ou expérience: un débat archaïque".  
Le débat 34 Mars 1985 (21-31).
- Bouissac, P. (1991) "Introduction: The circus - a semiotic spectroscopy".  
Semiotica 85 - 3/4 (1991), 189-199.
- Brooks, D.R. and E.O. Wiley (1988) Evolution as Entropy. Toward a Unified Theory of Biology. Second Edition. Chicago: University of Chicago Press.
- Clifford, J. and G.E. Marcus (eds.) (1986) Writing Culture: The Poetics and the Politics of Ethnography. Berkely: University of California Press.
- Cohen, F. (1985) "Computer viruses". Computer Security: A Global Challenge. Amsterdam: Elsevier Science 143-159.
- Cohen, F. (1987) "Computer viruses, theory and experiments." Computers and Security 6:22-35.
- Cohen, F. (1990) A Short Course on Computer viruses. :ASP Press.
- Cohen, F. (1991) "Friendly Contagion. Harnessing the Subtle Power of Computer Viruses". The Sciences (September/October 1991) 22-28.
- D'Amico, R.C. (1989) Historicism and Knowledge. New York: Routledge, Chapman and Hall.
- Dawkins, R. (1976) The Selfish Gene. Oxford: Oxford University Press.

- Dawkins, R. (1982) *The Extended Phenotype*. Oxford: Oxford University Press.
- Dawkins, R. (1986) *The Blind Watchmaker*. London: Penguin Books.
- Dawkins, R. (1989) *The Selfish Gene*. New Edition. Oxford: Oxford University Press.
- Delius, J.D. (1990) "Of mind memes and brain bugs": a natural history of culture". In W.A. Koch (ed.) *The Nature of Culture*. Bochum: Studienlag Brockmeyer.
- Dennett, D. (1990)
- Duncan, R. and M. Weston-Smith (eds.) (1977) *The Encyclopaedia of Ignorance*. Oxford: Pergamon.
- Eigen, M. (1975) "Evolutionary Games." In I.R. Miller (ed.) *Stability and Origin of Biological Information*: The first Aharon Katzir-Katchalsky Conference [June 23-28, 1973] New York: Wiley.
- Feyerabend, P. (1975) *Against Method*. London: New Left Books.
- Geertz, C. (1973) *The Interpretation of Cultures*. New York: Basic Books.
- Gardin, J.-Cl. et. al. (1987) *La logique du plausible. Essais d'épistémologie pratique en sciences humaines*. Paris: Editions de la Maison des Sciences de l'Homme (2nd edition).
- Jackson, A. (ed.) (1987) *Anthropology at Home*. ASA Monographs 25. London: Tavistock.
- Kuhn, T. (1970) *The Structure of Scientific Revolutions*. Chicago: University of Chicago Press.
- Lakatos, I. (1978) *The Methodology of Scientific Research Programmes. Philosophical Papers volume 1*. Cambridge: Cambridge University Press.
- Langton, C.G. (ed.) (1991) *Artificial Life*. vol. I. Redwood City (CA): Addison-Wesley.
- Lévi-Strauss, Cl. (1969) *The Raw and the Cooked*. Trans. by J. and D.

Weightman. New York: Harper and Row.

Malamoud, C. (1989) Cuire le monde. Rite et pensée dans L'Inde ancienne. Paris: Editions La Découverte.

Marcus, G.E. and M.M. Fischer (1986) Anthropology as Cultural Critique: An Experimental Moment in the Human Sciences. Chicago: University of Chicago Press.

McCloskey, D.Y. (1986) The Rhetoric of Economics. Madison: University of Wisconsin Press.

Nagel, E. and J.R. Newman (1958) Gödel's Proof. New York: New York University Press.

Pera, M. and W.R. Shea (eds.) (1991) The Art of Scientific Rhetoric. Canton (MA): Science History Publications.

Popper, K.R. (1972) Objective Knowledge. An Evolutionary Approach. Oxford: Clarendon Press.

Ruprecht, H.-G. (1991) "The reconstruction of Intertextuality." In H.F. Plett (ed.) Intertextuality. Berlin: Walter de Gruyter (60-77).

Ryle, G. (1971) A Collection of Critical Essays. O.P. Wood and G. Pitcher (eds.) London: MacMillan.

Serres, M. (1989)

Simons, G. (1989) Viruses, Bugs and Star Wars. Oxford: Blackwell.

Suppe, F. (ed.) (1977) The Structure of Scientific Theories. Urbana: University of Illinois Press.

Theocaris, T. and M. Psimopoulos (1987) "Where science has gone wrong." Nature 329 (15 October 1987), 595-598.

Thom, R. (1985) "La méthode expérimentale: un mythe des épistemologues (et des savants?)" Le Débat 34, Mars 1985 (11-20).

White, J.B. (1986) Hercules' Bow: The Rhetoric and Poetics of Law. Madison: University of Wisconsin Press.

## EDITORIAL: MEMES MATTER

By  
**Paul Bouissac**

This editorial appeared in Volume 5 (2) of The Semiotic Review of Books

Although the *meme* hypothesis has been around for quite some time - it was first proposed by Richard Dawkins in 1976 - it does not seem to have had any decisive impact yet on the speculations and investigations of semioticians. There are notable exceptions such as Koch (1986) or Delius (1990), but they remain marginal in the field of semiotic theory. Some disciplines have sporadically taken notice of Dawkins' idea either to dismiss it as "nothing new" (Costall, 1991) or provocatively champion it by developing its full metaphoric potential (Dennett, e.g. 1990). Others have tried to operationalize the concept within the domain of evolutionary biology (e.g. Bonner, 1990). However, its novel, counter-intuitive implications deserve more serious considerations in the broader epistemological context of a variety of contemporary advances both theoretical and empirical. As has been often the case in the history of scientific knowledge, an idea may inconsequentially float around for some time until it connects with other concepts or methods, usually apparently unrelated, to eventually precipitate, so to speak, and drastically change the way in which humans perceive their environment and themselves. Is the *meme* hypothesis of such a kind?

Let us recall the initial, somewhat awkward formulation of the *meme* hypothesis: "Just as genes propagate themselves in the gene pool by leaping from body to body via sperm or eggs, so memes propagate themselves in the *meme* pool by leaping from brain to brain via a process which, in the broad sense can be called imitation. If a scientist hears, or reads about a good idea, he passes it on to his colleagues and students" (Dawkins 1976: 206. The term "*meme*" was somewhat playfully coined after the French "Meme" for "same" and with phonetic allusions both to "mimetic" or "memory" and to "gene." In the paragraph following this introduction of the new concept, Dawkins quoted the summary that one of his colleagues, N.K. Humphrey, produced in reaction to the reading of an earlier draft of this chapter: memes should be regarded as living structures, not just metaphorically but technically. When you plant a fertile *meme* in my mind you literally parasitize my brain, turning it into a vehicle for the *meme*'s propagation in just the way a virus may parasite the genetic mechanism of a host cell. And this isn't just a way of talking - the *meme* for, say, "belief in life after death" is actually realized physically, millions of times over, as a structure of the nervous systems of individual men the world over (as quoted in Dawkins 1976: 207).

In his subsequent publications, Dawkins returned to the concept of "*meme*" (1982: 97-117, 1986: 158, 1989: 322) and indicated attempts made by others to develop his hypothesis (e.g., Delius 1990).

Until now, the idea of entities endowed with both genotypic and phenotypic structures evolving in a parasitic relation to the organism whose brains provide them with a vital environment, remains an intriguing and fascinating metaphor which shatters our worldview as if it were an aftershock of Darwinism, but it may seem to be too "odd" to be taken at face value. Dawkins

himself has been cautious in his subsequent and brief treatments of the "meme" idea, as if he were daunted by the full implications of this emergent theory. However, in the context of the changing epistemological landscape of the 1990s, this reconceptualization is being given new dimensions by a range of recently developed concepts, such as computer viruses (e.g., Cohen 1990; Delius 1990) and artificial life (Langton 1992), which provide a quasi-operational context for the speculative manipulation of this new idea (Bouissac 1992). For example, advances in the understanding of self-replication as an emergent property arising from local interactions in fairly simple systems (Reggia et al. 1993) 'suggest a possible avenue for a formal, canonic description of memes.

However, strong resistance to this approach, ranging from amused disbelief to philosophical arguments - mostly drawing from the long standing controversy concerning intentionality (e.g., Costall 1991) - has transformed a seemingly playful suggestion with which eccentric minds would toy into a potentially dangerous frontier of science, as if the Darwinian theory of evolution were bracing for a renewed confrontation with the last defense line of a long-established worldview order dominated by dualism (Bouissac 1991).

Along the line of Dawkins, "memes" will be taken here as informational structures, or sets of instructions, endowed with their own dynamism (autopoiesis) and constrained by the same evolutionary laws as genes, but their biological status will be more specifically ascribed to parasitism both in the predatory sense-- as N.K. Humphrey first visualized the idea (see quotation above) - and the symbiotic sense, along the line of the developing paradigm of symbiogenesis (Khakhina 1992).

But a note of caution is in order. In dealing with such problems, it is obvious that the *heuristic* status of the meme-as-parasite hypothesis should be emphasized even if some scientifically established data and currently accepted theories are brought into the argument. All too often, the humanities in the social sciences have borrowed ideas out of context from the natural and empirical sciences and have generalized them in the form of all-purpose metaphors of little descriptive value, such as the notion of social Darwinism or the right/left dichotomization of cerebral functions. An exploration of the meme hypothesis should try to avoid such pitfalls by a constant concern for the falsifiability of the propositions it may come to formulate. At the very least, the relevance to semiotics and to the social sciences of the issues potentially addressed by the meme hypothesis must be acknowledged, and even if the conceptual exploration sketched here proves to be misguided it might help better define the limits and pertinence of this challenge to established wisdom.

In as much as memes depend on other organisms for sustenance and reproduction (replication), their survival is bounded by the lifecycle and reproductive success of their hosts. As bearers of brains which afford the conditions necessary for memes to spread through replication, *Homo sapiens* is a vital resource which is itself dependent upon predation and other environmental commodities for the species' evolution and maintenance. Moreover, *Homo sapiens'* biological mode of survival involves a high degree of motility and dispersal. Therefore, memes can become extinct for the same reasons as parasites. An examination of the life conditions of parasites thus sheds some heuristic light on memes' modalities of existence.

Typically, parasites are dependent upon patchy and ephemeral resources because the other organisms - plants or animals - which are their hosts are themselves scattered over large areas and are not constant over time. The feeding specialization of the hosts compounds with the feeding specialization of the parasites, which is often extremely narrow. In addition, a parasite's life-cycle often requires more than one host, thus increasing the complexity of its reproduction, mainly if stages in the parasite's development must match stages in the hosts' lives. Since many parasites themselves use other parasites as hosts, it should be obvious that trophic and reproductive constraints make parasites' survival more complex than is the case for free-living predatory organisms, which depend on a much larger array of feeding resources and whose reproductive process is more flexible. As a result, parasitic species can never reach a state of equilibrium with their environment (i.e., their hosts) unless their parasitic mode of life evolves towards mutualism or even symbiosis. Another remarkable result of the parasitic mode of life is that ecological and genetic factors constantly interact to form within a given species "a patchwork quilt of genetic incompatibilities between closely adjacent populations" (Harper 1977: 412). Low probability of colonization and high probability of extinction in patches are among the consequences of the above properties of parasitic species (Price 1980: 77). In response to such ecological constraints, these species have evolved a variety of reproductive strategies such as an extreme longevity, including in some cases extended resting stages - up to several decades - which can combine with various sexual and non-sexual modes of reproduction such as mating before dispersal (colonization), self-compatible hermaphroditism or parthenogenesis.

There are obvious homologies between some of these strategies and what can be observed in the case of memes, namely the clonal mode of reproduction and the reliance on extended resting stages in the form of inscriptions - using a variety of semiotic systems --stored into non-organic supports. These periods of latency, which enable memes to bridge the gaps occurring in their vital resources, do not constitute however an absolute warranty of survival since, as in the case of parasites, they depend upon coming across an appropriate host at least within the period of time during which the physical support ensures their conservation.

Another important factor to be considered is that memes cannot be confused with evolved, wired-in, adaptive behavior. As any parasitic organisms, their biological dynamic exclusively concerns their own replication - in this respect they are strictly "selfish" to use the metaphor that Dawkins applied to genes - and consequently are neither detrimental nor beneficial to their hosts by necessity. In the same way as random genetic variations sometimes turn out to be adaptive, thus causing evolution to occur. Memes can both be aleatoric and occasionally generate in their hosts behavioral modifications which are adaptive (Bouissac 1992). However, meme-driven cultural habits may be only relatively adaptive with respect to a temporary subset of conditions, and may happen to be devastating in the long run. Historical time must be always measured against evolutionary time when assessing cultural changes. However, some memes may afford an adaptive advantage which improves the reproductive success of *Homo sapiens* genes. The reverse can naturally occur and some memes can become extinct because their own success in replication causes the extinction of their hosts. It must be remembered here that their hosts form a patchy and ephemeral resource and that resting stages are not immune to total destruction (viz. the burning of ancient libraries).

If memes are indeed a crucial factor in the emergence and evolution of cultures, the latter's patterns of patchy development and eventual sudden or slow extinction are compatible with an important consequence of the parasitic mode of life: namely the impossibility of reaching a state of population equilibrium (Price 1980: 44-52).

Beside the death of the host - or the extinction of its species - caused either by the ill-adaptive effects of a meme or a combination of memes, or by other factors, memes can for reasons other than those relating to their parasitic mode of survival. At least three such causes can be considered: changes in the environment, intrinsic fragility, and elimination by other strains of memes. These are general evolutionary conditions applying to all forms of life organic or not (Langton 1991).

As any organism, memes cannot be conceptualized independently of their environment. Adaptation by selection upon both genetic and memetic variations operates with respect to environmental factors. Catastrophic changes in memes' environment, that is the brain or a particular anatomical, physiological or chemical component of it, may eradicate a meme population, or trap it in a dead end without possibilities of further replications.

Another factor which must be taken into consideration when assessing the conditions of survival of memes is their intrinsic structure and degree of complexity and specialization. Complexity entails fragility not only because of a closer dependency on environmental conditions but also because of a higher risk of replication errors. These two factors may compound to drive memes to extinction. An ad-hoc example of the latter could be some secret formulae of ancient medicine or medieval alchemy. Selective secrecy and exclusively oral replication - i.e., absence of resting stage - can be intrinsic components of a meme's set of instructions and thus account for its built-in fragility. It can be hypothesized that some memes do not survive beyond a small number of replicative generations. The argument of Dawkins according to which longevity is a criterial feature of memes seems to be hardly defensible. With respect to evolutionary time, how can longevity be measured? Even a pattern of information which would replicate only once would qualify as a meme because of this very property which has been thus demonstrated. A meme is no more necessarily a successful meme than a gene is. Unsuccessful memes are just short-lived ones. But there is more: if memes indeed conform to some characteristics of parasitic species - that is, incidentally more than fifty percent of all extant species according to conservative estimates (Price 1980: 8) - they may share their general property of "evolving slowly" and representing "dead ends in any phylogeny" (Price 1980: 11). Although Price qualifies this view, he quotes other parasitologists who thus express current theoretical views: "parasites as a whole are worthy examples of the inexorable march of evolution into blind alleys" (Noble and Noble 1976: 525). Therefore, memes would die, that is cease to replicate, not only selectively because of an intrinsic structural fragility but generally, as a result of a proclivity attached to their specific mode of survival as parasites, not necessarily because of their host's death.

But this is not all. Within a given segment of evolution memes are bound to compete with each other for brain territories and resources which are essential to their existence and reproduction. It can be hypothesized that strains of memes have no choice but to strive to eliminate each other directly through structural disintegration or assimilation - that is by neutralizing the information value of their competitors, or indirectly by instructing their own host to destroy the hosts of other

competing memes. Admittedly, such a formulation make memes' warfare sound and look like a video game, but let us point out that the analogy might not be as far-fetched as it seems, given the fact that on the one hand confrontational video games explicitly mimic human affairs and that on the other hand they are based upon the mutual manipulation of information patterns or sets of instruction, and organisms through the medium of a tool which itself would easily qualify as a meme. The hierarchical depth of parasitism commonly involves several levels of embeddings, and there is no reason why this feature could not be heuristically extended to memes.

The above speculative exploration of the memes as parasites hypothesis can generate falsifiable hypotheses. For instance, Cavalli-Sforza (1971) Cavalli-Sforza and Feldman (1981) have shown how quantitative approaches to the study of sociocultural evolution can be modeled with reference to biological evolution. Nevertheless, the meme hypothesis remains a distant albeit fascinating frontier.

On the one hand, it fits to some degree a long standing intuitive knowledge; countless observations have indeed led to a tradition of human anxiety regarding the power of ideas and the alienating effects of ideologies, the taking over of human behavior by directly or indirectly self-destructive programs; "habits" have been scrutinized with some puzzlement by philosophers and psychologists and have even been characterized as "a second nature," a metaphor which would become literal if the meme hypothesis were substantiated; but, on the other hand, it is deeply counter-intuitive and extremely difficult to operationalize; the conditions which would be required to test this hypothesis are still far from being afforded by current knowledge and means of investigation. Computer viruses and parasitic models at best offer intriguing analogies. Like the "sign" of semiotics, the "meme" of evolutionary biology, or memetic, is a hypothetical entity. However, further advances in the exploration of the meme hypothesis are obviously dependent upon advances in the understanding of the brain because it is generally recognized in evolutionary biology that it is impossible to understand an organism independently from an understanding of the environment within which it has evolved as a part of this environment. Let it suffice, for the time being, to ponder the basic concepts of parasites' evolutionary biology as expressed by Price (1980: 16-24) but by heuristically replacing "parasites" by "memes:" memes are adapted to exploit small, discontinuous environments; memes represent the extreme in specialized resource exploitation; and memes exist in non-equilibrium conditions.

This may provide the shifting of perspective which is necessary to reformulate the fundamental questions of semiotics in a novel, more productive manner, not necessarily by equating signs and memes, but by considering the possibility that memes might be parasite signs and, as such, an important dimension of evolutionary semiosis.

## References

- Bonner, J.T. 1980 *The Evolution of Culture in Animals*. Princeton: Princeton University Press.  
Bonner, J.T. 1990 "Cultural Evolution: A Biologist's View" *Journal of Ideas*, 1(1),24-34.

- Bouissac, P. 1991 "Semiotics and the Gaia Hypothesis: Toward a Restructuring of Western Thought," *Philosophy and the Future of Humanity* 1(2),168-184.
- Bouissac, P. ,1992 "The Construction of Ignorance and the Evolution of Knowledge," *University of Toronto Quarterly*, 61(4),460-472.
- Cavalli-Sforza, L.L. 1971 "Similarities and Dissimilarities of Social Cultural and Biological Evolution." In: *Mathematics in the Archeological and Historical Sciences*. F.R. Hodson (ed.). Edinburgh: Edinburgh University Press.
- Cavalli-Sforza, L.L. and M.W. Feldman 1981 *Cultural Transmission and Evolution: A Quantitative Approach*. Princeton: Princeton University Press.
- Cohen, F. 1990 *A Short Course in Computer Viruses*. Pittsburgh: ASP Press.
- Costall, A. 1991 "The 'Meme' Meme," *Cultural Dynamics* IV(3),321-335.
- Dawkins, R. 1976 *The Selfish Gene*. Oxford: Oxford University Press.
- Dawkins, R. 1986 *The Blind Watchmaker*. London: Longman Scientific.
- Dawkins, R. 1989 *The Selfish Gene*. 2nd edition. Oxford: Oxford University Press.
- Delius, J.D. 1990 "Of Mind Memes and Brain Bugs," In: W.A. Koch (ed.) *The Nature of Culture*. Bochum: Studienverlag Brockmeyer,26-79.
- Dennett, D. 1990 "Memes and the Exploitation of Imagination," *Journal of Aesthetics and Art Criticism* 48(2).125-35.
- Harper, J.L., 1977 *Population Biology of Plants*. London: Academic Press.
- Hull D.L. 1988 *Science as a Process*. Chicago Chicago University Press.
- Khakhina, L.N. 1991 *Concepts of Symbiogenesis*. Trans. by S. Merkel and R. Coalson. edit. by L. Margulis and M. McMenamin. Yale University Press.
- Koch, W.A. 1986 *Genes vs. Memes; Modes of Integration for Natural and Cultural Evolution in a Holist Model*. Bochum: Studienverlag Brockmeyer.
- Langton, C.G. (ed.) 1991 *Artificial Life 1*. Redwood City, CA: Addison-Wesley.
- Lumsden, C.J. and E.O. Wilson 1981 *Genes, Mind, and Culture: The Coevolutionary Process*. Cambridge, MA: Harvard University Press.
- Moritz, E. 1990 "Memetic Science," *Journal of Ideas* 1(1), 3-23.

Noble, E.R. and G.A. Noble 1976 *Parasitology: The Biology of Animal Parasites*. 4th ed. Philadelphia: Lea and Febiger.

Price, P.W. 1980 *Evolutionary Biology of Parasites*. Princeton: Princeton University Press.

Reggia, J.A., S.L. Armentrout, Hui-Hsien Chou, Yua Peng 1993 "Simple Systems that Exhibit Self-directed Replication," *Science*, Vol. 249 (26 Feb.), 12821287.

Thompson, D'Arcy W. 1942 *On Growth and Form*. 2nd ed. London: Cambridge University Press.

*Paul Bouissac is Professor of French at the University of Toronto (Victoria College). His recent publications include: "The representation of Commonsense knowledge: semiotic modeling and artificial intelligence" (1993), "Ecology of semiotic space: competition, exploitation and the evolution of arbitrary signs" (1994), "Art or script? A falsifiable hypothesis" (1994) and "Syntactic iconicity and connectionist models of language and cognition" (1994).*

# WHY DO MEMES DIE?

By

Paul Bouissac

paul.bouissac@utoronto.ca

University of Toronto

*"We have come to the edge of a world of which we have no experience, and where all our pre-conceptions must be recast."* --- D'Arcy W. Thompson (1942:77)

## Introduction: What is a Meme?

It is being increasingly recognized that the notion of "meme" (Dawkins 1976) as a self-replicating information pattern (cultural unit) whose biotic environment is formed by human, and possibly primate[1], brains is one of the most intriguing, even upsetting hypotheses put forward in recent years. Like the more palatable -- because of its inherent functionalism --notion of "culturgens" proposed by Lumsden and Wilson (1981), this new way of framing semiosis and cultural dynamism in an evolutionary, Darwinian, perspective should be of prime interest to semioticians in as much as it provides a novel approach to the conceptualization of signs and their action. In spite of some efforts to trace the idea back to earlier metaphoric attempts "to define a cultural analogue of the gene" (Costall 1991: 323) or to show that traditional Western philosophy has always tried to account for the dynamism of signs or ideas (Moritz 1990: 5-9), Dawkins' tentative proposal that cultural units of replication may have some features in common with genes has struck many as an entirely novel, counterintuitive perspective deserving serious consideration, mainly in view of the theoretical context provided by *The Selfish Gene* (Dawkins 1976). Delius (1989), Dennett (1990), Hull (1988), Koch (1986), Moritz (1990), for instance, have devoted considerable efforts to showing how this hypothesis may be much more than a metaphor and may indeed bring about a revolution in the way we conceptualize the dynamics of cultural, mental and intellectual life. This seems all the more interesting since the theoretical link between this hypothesis and earlier, seminal definitions of semiotics can be made without conceptual strain, albeit somewhat unexpectedly: on the one hand, the meme hypothesis matches some important aspects of Peirce's accounts of semiosis and gives a surprising turn to his "perfusion of signs" metaphor; on the other hand, Saussure's definition of semiology as the study of the life of signs in social life (my emphasis) takes up a new meaning if "life" is understood literally, i.e., in its full biological sense, rather than metaphorically as some kind of abstract dynamism.

Let us recall the initial, somewhat awkward formulation of the meme hypothesis: "Just as genes propagate themselves in the gene pool by leaping from body to body via sperms or eggs, so memes propagate themselves in the meme pool by leaping from brain to brain via a process which, in the broad sense can be called imitation. If a scientist hears, or reads about, a good idea, he passes it on to his colleagues and students" (Dawkins 1976: 206). The term "meme" was somewhat playfully coined after the French "*même*" for "same" and with phonetic allusions both to "mimetic" or "memory" and to "gene." In the paragraph following this introduction of the new concept, Dawkins quoted the summary that one of his colleagues, N.K. Humphrey,

produced in reaction to the reading of an earlier draft of this chapter: memes should be regarded as living structures, not just metaphorically but technically. When you plant a fertile meme in my mind you literally parasitize my brain, turning it into a vehicle for the meme's propagation in just the way a virus may parasite the genetic mechanism of a host cell. And this isn't just a way of talking -- the meme for, say, "belief in life after death" is actually realized physically, millions of times over, as a structure of the nervous systems of individual men the world over (as quoted in Dawkins 1976: 207).

In his subsequent publications Dawkins returns to the concept of "meme" (1982: 97-117, 1986: 158, 1989: 322) and indicates attempts made by others to develop his hypothesis (e.g., Delius 1990).

Until now, the idea of entities endowed with both genotypic and phenotypic structures evolving in a parasitic relation to the organism whose brains provide them with a vital environment, remains an intriguing and fascinating metaphor which shatters our worldview as if it were an aftershock of Darwinism but it may seem to be too "odd" to be taken at face value. Dawkins himself has been cautious in his subsequent and brief treatments of the "meme" idea, as if he were daunted by the full implications of this emergent theory. However, in the context of the changing epistemological landscape of the 1990s, this reconceptualization is being given new dimensions by a range of recently developed concepts such as computer viruses (e.g., Cohen 1990; Delius 1990) and artificial life (Langton 1992) which provide a quasi-operational context for the speculative manipulation of this new idea (Bouissac 1992). For example, advances in the understanding of self-replication as an emergent property arising from local interactions in fairly simple systems (Reggia et al. 1993) suggest a possible avenue for a formal, canonic description of memes.

However, strong resistance to this approach, ranging from amused disbelief to philosophical arguments -- mostly drawing from the long standing controversy concerning intentionality (e.g., Costall 1991) – has transformed a seemingly playful suggestion with which eccentric minds would toy into a potentially dangerous frontier of science, as if the Darwinian theory of evolution were bracing for a renewed confrontation with the last defense line of a long established worldview order dominated by dualism (Bouissac 1991).

This paper proposes to explore further the "meme" hypothesis by raising two heuristic questions: "Why and how do memes die?" and "What can we learn from the way they die about the way they live?" Thus, it will focus on some of the features and processes which may account for the successful replication (conservation) or the extinction (dissipation) of "memes," a phenomenon which requires to be examined from a multi-disciplinary perspective. Along the line of Dawkins, "memes" will be taken here as informational structures, or sets of instructions, endowed with their own dynamism (autopoiesis) and constrained by the same evolutionary laws as genes, but their biological status will be more specifically ascribed to parasitism both in the predatory sense -- as N.K. Humphrey first visualized the idea (see quotation above) -- and the symbiotic sense, along the line of the developing paradigm of symbiogenesis (Khakhina 1992).

But a note of caution is in order. In dealing with such problems, it is obvious that the heuristic status of the meme-as-parasite hypothesis should be emphasized even if some scientifically

established data and currently accepted theories are brought into the argument. All too often, the humanities and the social sciences have borrowed ideas out of context from the natural and empirical sciences and have generalized them abusively in the form of all-purpose metaphors of little descriptive value, such as the notion of social Darwinism or the right/left dichotomization of cerebral functions. This contribution to the exploration of the meme hypothesis tries to avoid such pitfalls by a constant concern for the falsifiability of the propositions it may come to formulate. At the very least the relevance to semiotics and to the social sciences of the issues potentially addressed by the meme hypothesis must be acknowledged, and even if the conceptual exploration developed in this paper proves to be misguided it might help better define the limits and pertinence of this challenge to established wisdom.

## Memes' Ways of Dying

It can be hypothesized that there are at least four causes which can bring about the demise of a meme. These causes should be critically examined in relation to case studies such as the extinction of a language, the discarding of a fashion, the refutation of a theory (in the range of macro-memes), and the phasing out of a word, a habit, a belief (in the range of micro-memes). However, as a first step in this direction, these causes will be considered here in a purely speculative manner. The first cause is a direct result of the parasitic mode of survival of memes; the other three are not specific to memes but apply to memes as to any other organism submitted to the laws of evolution by selection, although they can be expressed in a meme-specific manner. These various models of extinction will now be examined serially.

In as much as memes depend on other organisms for sustenance and reproduction (replication), their survival is bounded by the life-cycle and reproductive success of their hosts. As bearers of brains which afford the conditions necessary for memes to spread through replication, *Homo sapiens* is a vital resource which is itself dependent upon predation and other environmental commodities for the species' evolution and maintenance. Moreover, *Homo sapiens'* biological mode of survival involves a high degree of motility and dispersal. Therefore, memes can become extinct for the same reasons as parasites. An examination of the life conditions of parasites thus sheds some heuristic light on memes' modalities of existence.

Typically, parasites are dependent upon patchy and ephemeral resources because the other organisms -- plants or animals -- which are their hosts are themselves scattered over large areas and are not constant over time. The feeding specialization of the hosts compounds with the feeding specialization of the parasites, which is often extremely narrow [2]. In addition, a parasite's life-cycle often requires more than one host, thus increasing the complexity of its reproduction, mainly if stages in the parasite's development must match stages in the hosts' lives. Since many parasites themselves use other parasites as hosts, it should be obvious that trophic and reproductive constraints make parasites' survival more complex than is the case for free-living predatory organisms, which depend on a much larger array of feeding resources and whose reproductive process is more flexible. As a result, parasitic species can never reach a state of equilibrium with their environment (i.e., their hosts) unless their parasitic mode of life evolves toward mutualism or even symbiosis. Another remarkable result of the parasitic mode of life is that ecological and genetic factors constantly interact to form within a given species "a patchwork quilt of genetic incompatibilities between closely adjacent populations" (Harper 1977:

412). Low probability of colonization and high probability of extinction in patches are among the consequences of the above properties of parasitic species (Price 1980: 77). In response to such ecological constraints these species have evolved a variety of reproductive strategies such as an extreme longevity, including in some cases extended resting stages -- up to several decades -- which can combine with various sexual and non-sexual modes of reproduction such as mating before dispersal (colonization), self-compatible hermaphroditism or parthenogenesis.

There are obvious homologies between some of these strategies and what can be observed in the case of memes, namely the clonal mode of reproduction and the reliance on extended resting stages in the form of inscriptions -- using a variety of semiotic systems -- stored into non-organismic supports. These periods of latency, which enable memes to bridge the gaps occurring in their vital resources, do not constitute however an absolute warranty of survival since, as in the case of parasites, they depend upon coming across an appropriate host at least within the period of time during which the physical support ensures their conservation.

Another important factor to be considered is that memes cannot be confused with evolved wired-in adaptive behavior. As any parasitic organisms, their biological dynamic exclusively concerns their own replication -- in this respect they are strictly "selfish" to use the metaphor that Dawkins applied to genes -- and consequently are neither detrimental nor beneficial to their hosts by necessity. However, some may afford an adaptive advantage which improve the reproductive success of *Homo sapiens* genes.[3] The reverse can naturally occur and some memes can become extinct because their own success of replication causes the extinction of their hosts. It must be remembered here that their hosts form a patchy and ephemeral resource and that resting stages are not immune to total destruction (viz. the burning of ancient libraries).

If memes are indeed a crucial factor in the emergence and evolution of cultures, the latter's patterns of patchy development and eventual sudden or slow extinction are compatible with an important consequence of the parasitic mode of life: namely the impossibility of reaching a state of population equilibrium (Price 1980: 44-52).

Beside the death of the host -- or the extinction of its species -- caused either by the ill-adaptive effects of a meme or a combination of memes, or by other factors, memes can die for reasons other than those relating to their parasitic mode of survival. At least three such causes can be considered: changes in the environment, intrinsic fragility, and elimination by other strains of memes. These are general evolutionary conditions applying to all forms of life -- let it be organic or not (Langton 1991).

As any organism, memes cannot be conceptualized independently of their environment. Adaptation by selection upon both genetic and memetic variations operates with respect to environmental factors. Catastrophic changes in memes' environment that is the brain or a particular anatomical, physiological or chemical component of it, may eradicate a meme population or trap it in a dead-end without possibilities of further replications. Local disruptions causing amnesia, general degenerative pathologies, such as Alzheimer's disease, chemical imbalance occurring naturally or being induced by external factors may seal the fate of memes. Since memes have the peculiarity of replicating via modifications of their hosts' behavior, those which are replicated through verbal means are jeopardized by aphasic conditions, those which

spread through observational learning can be threatened by various forms of motor impairments. Such deleterious modifications of the neurological milieu are, from the point of view of memes' evolution, essentially similar to the environmental constraints on any living organism which can become extinct when confronted to sudden and drastic changes.

Another factor which must be taken into consideration when assessing the conditions of survival of memes is their intrinsic structure and degree of complexity and specialization. Complexity entails fragility not only because of a closer dependency on environmental conditions but also because of a higher risk of replication errors. These two factors may compound to drive memes to extinction. An ad-hoc example of the latter could be some secret formulas of ancient medicine or medieval alchemy. Selective secrecy and exclusively oral replication -- i.e., absence of resting stage -- can be intrinsic components of a meme's set of instructions and thus account for its built-in fragility. It can be hypothesized that some memes do not survive beyond a small number of replicative generations. The argument of Dawkins according to which longevity is a criterial feature of memes seems to be hardly defensible. With respect to evolutionary time, how can longevity be measured? Even a pattern of information which would replicate only once would qualify as a meme because of this very property which has been thus demonstrated. A meme is no more necessarily a successful meme than a gene is. Unsuccessful memes are just short-lived ones. But there is more: if memes indeed conform to some characteristics of parasitic species -- that is, incidentally more than fifty percent of all extant species according to conservative estimates (Price 1980: 8) -- they may share their general property of "evolving slowly" and representing "dead ends in any phylogeny (Price 1980: 11). Although Price qualifies this view, he quotes other parasitologists who thus express current theoretical views: "parasites as a whole are worthy examples of the inexorable march of evolution into blind alleys" (Noble and Noble 1976: 525). Therefore, memes would die, that is cease to replicate, not only selectively because of an intrinsic structural fragility but generally, as a result of a proclivity attached to their specific mode of survival as parasites, not necessarily because of their host's death.

But this is not all. Within a given segment of evolution memes are bound to compete with each other for brain territories and resources which are essential to their existence and reproduction. It can be hypothesized that strains of memes have no choice but strive to eliminate each other directly through structural disintegration or assimilation -- that is by neutralizing the information value of their competitors -- or indirectly by instructing their own host to destroy the hosts of other competing memes. Admittedly, such a formulation make memes' warfare sound and look like a video game, but let us point out that the analogy might not be as far-fetched as it seems given the fact that on the one hand confrontational video games explicitly mimic human affairs and that on the other hand they are based upon the mutual manipulation of information patterns or sets of instruction, and organisms through the medium of a tool which itself would easily qualify as a meme. The hierarchical depth of parasitism commonly involves several levels of embeddings, and there is no reason why this feature could not be heuristically extended to memes.

## **Conclusion: A Program for Evolutionary Semiosis**

The above speculative exploration of the memes as parasites hypothesis can generate falsifiable hypotheses. For instance, Cavalli-Sforza (1971), Cavalli-Sforza and Feldman (1981) have shown

how quantitative approaches to the study of socio-cultural evolution can be modeled with reference to biological evolution. Nevertheless, the meme hypothesis remains a distant, albeit fascinating frontier. On the one hand, it fits to some degree a long standing intuitive knowledge; countless observations have indeed led to a tradition of human anxiety regarding the power of ideas and the alienating effects of ideologies, the taking over of human behavior by directly or indirectly self-destructive programs; "habits" have been scrutinized with some puzzlement by philosophers and psychologists and have even been characterized as "a second nature," a metaphor which would become literal if the meme hypothesis were substantiated; but, on the other hand, it is deeply counter-intuitive and extremely difficult to operationalize; the conditions which would be required to test this hypothesis are still far from being afforded by current knowledge and means of investigation. Computer viruses and parasitic models, at best, offer intriguing analogies. Like the "sign" of semiotics, the "meme" of evolutionary biology or memetic is a hypothetical entity. It raises many questions and solves none. It is however the contention of this paper that it improves the quality of the questions raised through delineating what could be called a field of relevance and through shifting, on the epistemological spectrum, the heuristic speculation further toward theoretical and empirical determination of cultural evolution. Further advances in the exploration of the meme hypothesis are obviously depending upon advances in the understanding of the brain because it is generally recognized in evolutionary biology that it is impossible to understand an organism independently from an understanding of the environment within which it has evolved as a part of this environment. Let it suffice for the time being to ponder the basic concepts of parasites' evolutionary biology as expressed by Price (1980: 16-24) but by heuristically replacing "parasites" by "memes:" memes are adapted to exploit small, discontinuous environments; memes represent the extreme in specialized resource exploitation; and memes exist in non-equilibrium conditions. This may provide the shifting of perspective which is necessary to reformulate the fundamental questions of semiotics in a novel, more productive manner, not necessarily by equating signs and memes, but by considering the possibility that memes might be parasite's signs, and, as such, an important dimension of evolutionary semiosis.

## References

Bonner, J.T.

1980 *The Evolution of Culture in Animals*. Princeton: Princeton University Press.

1990 "Cultural Evolution: A Biologist's View," *Journal of Ideas*, I(1), 24-34.

Bouissac, P.

1991 "Semiotics and the Gaia Hypothesis: Toward a Restructuring of Western Thought," *Philosophy and the Future of Humanity* I(2), 168-184.

1992 "The Construction of Ignorance and the Evolution of Knowledge," *University of Toronto Quarterly*, 61(4), 460-472.

Cavalli-Sforza, L.L.

1971. "Simularities and Dissimularities of Social Cultural and Biological Evolution." In: *Mathematics in the Archeological and Historical Sciences*. F.R. Hodson (ed.). Edinburgh: Edinburgh University Press.

Cavalli-Sforza, L.L. and M.W. Feldman  
1981 Cultural Transmission and Evolution: A Quantitative Approach.  
Princeton: Princeton University Press.

Cohen, F.  
1990 A Short Course in Computer Viruses. Pittsburgh: ASP Press.

Costall, A.  
1991 "The 'Meme' Meme," Cultural Dynamics IV(3), 321-335.

Dawkins, R.  
1976 The Selfish Gene. Oxford: Oxford University Press.  
1986 The Blind Watchmaker. London: Longman Scientific.  
1989 The Selfish Gene. 2nd edition. Oxford: Oxford University Press.

Delius, J.D.  
1990 "Of Mind Memes and Brain Bugs," In: W.A. Koch (ed.) The Nature of Culture. Bochum: Studienverlag Brockmeyer, 26-79.

Dennett, D.  
1990 "Memes and the Exploitation of Imagination," Journal of Aesthetics and Art Criticism 48(2), 125-35.

Harper, J.L.  
1977 Population Biology of Plants. London: Academic Press.

Hull, D.L.  
1988 Science as a Process. Chicago: Chicago University Press.

Khakhina, L.N.  
1991 Concepts of Symbiogenesis. Trans. by S. Merkel and R. Coalson. edit. by L. Margulis and M. McMenamin. Yale University Press.

Koch, W.A.  
1986 Genes vs. Memes: Modes of Integration for Natural and Cultural Evolution in a Holist Model. Bochum: Studienverlag Brockmeyer.

Langton, C.G. (ed.)  
1991 Artificial life I. Redwood City, CA: Addison-Wesley.

Lumsden, C.J. and E.O. Wilson  
1981 Genes, Mind, and Culture: The Coevolutionary Process. Cambridge, MA: Harvard University Press.

Moritz, E.  
1990 "Memetic Science," Journal of Ideas 1(1), 3-23.

Noble, E.R. and G.A. Noble

1976 Parasitology: The Biology of Animal Parasites. 4th ed. Philadelphia:

Lea and Febiger.

Price, P. W.

1980 Evolutionary Biology of Parasites. Princeton: Princeton University Press.

Reggia, J.A., S.L. Armentrout, Hui-Hsien Chou, Yua Peng

1993 "Simple Systems that Exhibit Self-directed Replication," *Science*, Vol. 249 (26 Feb.), 1282-1287.

Thompson, D'Arcy W.

1942 On Growth and Form. 2nd ed. London: Cambridge University Press.

## Notes

[1]. However, Bonner who defined memes as "any bit or any collection of bits of information passed by behavioral means from one individual to another" (1980), extends the operability of this concept to social animals well beyond the primates (1990: 28-31).

[2]. It must also be kept in mind that parasites influence many aspects of their hosts' qualities, thus altering the hosts' ecological system. "Behavior, reproductive potential, competitive ability, and susceptibility to predation may all be modified and the degrees of change will differ with levels of infection. This variability in the host population, superimposed on the genotypic and phenotypic differences within it, greatly complicates its ecological and evolutionary processes" (Price 1980: 10-11).

[3]. In the same way as random genetic variations sometimes turn out to be adaptive, thus causing evolution to occur, memes can both be aleatoric and occasionally generate in their hosts behavioral modifications which are adaptive (Bouissac 1992). However, meme-driven cultural habits may be only relatively adaptive with respect to a temporary subset of conditions, and may happen to be devastating in the long run. Historical time must be always measured against evolutionary time when assessing cultural changes.

# THE HUMAN AGENCY OF MEME MACHINES

An extended review of:  
**Blackmore, Susan (1999) *The Meme Machine*.**  
Oxford: Oxford Univ. Press. 264 pp. ISBN 0 19 850 365 2

By  
**Gary Boyd**  
Concordia University  
Montreal, Quebec, Canada.  
boydg@vax2.concordia.ca

## 1 Importance

For any thoughtful reader, Susan Blackmore's *Meme Machine* is an important book because it clearly explains and illustrates the basic import of the new field of memetics. Actually memetics did not just spring from the head of Zeus (Dawkins) fully formed, but like Barkow, Cosmides and Tooby's (1998) evolutionary psychology, rather it is part of a general contemporary movement to extend evolutionary explanation beyond the purely biological world. At best, memetic evolutionary explanation will become nearly as central to the social sciences as Darwinian evolution has been to the biological sciences. At worst, memetics might conceivably become another way to deny the uniqueness of each mysterious human being. Or as Jacoby (1975) would caution, memetics qua ideology like the computer-metaphor for mind, could become yet another example of "*social amnesia*" - of reification leading to further alienation. Having acknowledged that possibility I must go on to state that current work makes me rather confident that memetics insights will actually help us find many new solutions to educational and political problems.

"*How is memetic evolution steerable?*" that is the key question and the one which Blackmore's *Meme Machine* poses and partly answers with good heuristics which should lead to real progress.

## 2 Three Key Achievements:

### 2.1 Educating engagingly:

The first of Susan Blackmore's main achievements in this volume is to provide a pleasantly, even intriguingly, readable understandable respectable book which one can offer to friends and students who ask: "*What is all this meme or memetics excitement about?*"

Susan Blackmore has written a thoroughly enjoyable and informative book for all reasonably well educated people such as readers of the New Scientist, or the Scientific American and Harper's magazines. Her style is fluent, clear, coherent, convincing, and entertaining. The *Meme Machine* is itself a carrier of an infectious and contagious 'memeplex' carefully crafted to become a symbiocytic fellow-traveler for us altruists who are teachers of the various sciences.

## **2.2 Advancing Science:**

As Dr. Blackmore the scientist describes science (p. 202) "*Science like religion is a mass of memplexes ... Science is fundamentally a process; a set of methods for trying to distinguish true memes from false ones. At its heart lies the idea of building theories about the world and testing them, rather like perceptual systems do*". She points out that "*Evolutionary theory faced enormous opposition because it provided a view of humans that many humans do not like. The same will probably be true for memetics. However at the heart of science lies the method of demanding tests of any idea. Scientists must predict what will happen if a particular theory is valid, and then find out if it is so. That is what I have tried to do with the theory of memetics.*"

The prediction of probable ubiquity of memetic theory among social scientists can be tested by trying out the explanatory framework on various 'hard' case phenomena in the various social sciences. This is what Blackmore begins to do in a schematic qualitative way. Her analytic forays are well grounded in the literature, and self-critical enough to be plausible. They are also exciting enough, and controversial enough to inspire much further thought, and to warrant much further work. That I believe is her main scientific achievement here.

Social science theories such as memetics and e.g. 19<sup>th</sup> century social-Darwinism can however be either self-fulfilling or self defeating prophecies. As they gain currency they become ever more true; or contrariwise as in the case of social Darwinism as it gained currency it evoked a wave of objections which constructed it to be false.

Memetics is first of all an explanatory theory with great heuristic value. I am confident, partly from reading Blackmore as well as Dawkins, Dennett, and many of the others, that it will indeed lead-on self-fulfillingly to important new truthful explanatory sub-theories and also even to, at least fuzzily, quantitative predictive models of cultural propagation, and evolution.

## **2.3 Liberating memplexes:**

Blackmore's third main achievement (- very much in line with her previous works - e.g. Blackmore, 1996) is to make challenges which liberate us from habitual and wishful thinking, and oppressive ideology. For example, consider this statement of hers on p. 241-2: "*But this is still a cop out*". As Daniel Dennett (1984) says '*The "independent" mind struggling to protect itself from dangerous memes is a myth*'. So we must ask: who gets to choose? If we take memetics seriously then the 'me' that could do the choosing is itself a memetic construct: a fluid and ever-changing group of memes installed in a complicated meme machine. The choices made will be a product of my genetic and memetic history in a given environment, not of some separate self that can 'have' a life purpose and overrule the memes that make it up. This is the power and beauty of memetics: it allows us to see how human lives language and creativity all come about through the same kind of replicator power as did design in the biological world.

What she does raise explicitly is the question of free-will and human agency pointing out that (p. 237) "...*the self that is supposed to have free will is just a story that forms part of a vast memplex, and a false story at that. On this view all human actions whether conscious or not, come from complex interactions between memes genes and all their products , in complicated*

*environments. The self is not the initiator of actions, it does not have consciousness, and it does not 'do' the deliberating. There is no truth in the idea of an inner self inside my body that controls the body and is conscious. Since this is false so is the idea of my conscious self having free will."*

This is arguably the most controversial claim in Blackmore's book. It is well warranted by the rest of the book. Yet it sits oddly in the lifework of a very responsible creative and apparently self-determined author who in her previous work has been circumspectly devoted to the critical scientific exploration and as it turned out refutation of most of the ubiquitous claims of parapsychology, and of the paranormal (Blackmore, 1996).

Right now I feel I can and must choose between going on with this review or going for a walk amid the beauty of the falling maple leaves. Without these in a sense illusory options I would feel my life to be pointless. The conflict between the option to go on writing versus the option of a Sunday stroll is a conscious epiphenomenon of contending memplexes embodied in the neuronal groups of my CNS (Central Nervous System), and triggered or reinforced by various current perceptions. The human-being which chooses encompasses all this and moreover has historical depth in my and others' socialization and geographical and biochemical history. It seems we memplex nexii must tell our stories anew if we wish to go on elaborating and propagating critical realist science.

Is this possible? Yes, theoretically it is if the *memetic* science memplex proves to be contagious enough and co-travels with the Critical Realist Science memplex. This can happen but not without the pruning and re-organizing or killing off (and grieving?) many religious memplexes which have evolved to behave to prevent pruning and re-organization let alone replacement.

In my view the memetic explanatory mechanism does not rule out purposive intention or at least Dennett's 'intentional stance', but it does require a different account of it - one which has not yet been scientifically articulated. Dennett went some way along toward a better account, but without making links to memetics, in his (1984) book on the varieties of free will worth wanting. Probably the best currently accessible treatments of consciousness are those of the neurophysiologists D'Amasio (1999) e.g.- p.313. "...the plotting of a relationship between any object and the organism becomes the feeling of a feeling. The mysterious first-person perspective of consciousness consists of newly-minted knowledge expressed as feeling ... Creativity - the ability to generate new ideas and artifacts requires more than consciousness can ever provide. It requires abundant memory, fine reasoning ability, language. But consciousness' revelations guide the process." (Tonini and Edelman, 2000). My reading of both is that they are fully compatible with meme-complex acquisition (in the form of synapse weightings and neuronal group selections), and neuronal memplex execution processes which generate both plans and the emotions that evoke feelings which along with other physiological homeostatic regulators, produce our decisions.

According to this theory, whether you will actually allow yourself to accept the tempting memplexes which Susan offers, and so to see all the human world in the memetic realist way, or not, is up to the ways memplexes are executing in your body, which together are you!

### **3 Four Adumbrative Criticisms:**

#### **3.1 Memes as replicands or replicators?**

Like others Susan Blackmore sidesteps the task of defining memes any more precisely than Dawkins did. Are memes *replicands* or *replicators*? - i.e. are they *merely catalysts* for their own reproduction via the host's pre-existing reproductive mechanisms like viruses, or do they become *embedded as self-reproducers* in the host human animal (- like bacterial parasites)?

As others have pointed out precise definition poses a serious challenge which must be taken up. My own offering as a possibly better (?) definition is: that a meme is a bounded set of modulations on any carrier which when received by an animal has by virtue of its configuration (Shannon information of an aesthetic, &/or pseudo sustenential, &/or proto-propagative propitious character which is attractive to that animal's pre-embodied memplexes) an high probability of being transduced into neuronal group and synapse threshold representations which result in effectively identical replications of the meme as outputted modulations on other carriers (which of course tend to produce further propagation (- seduction, infection and contagion) via other animals. This definition depends on the engineering definitions: of carriers and modulations, transducers, receivers, Shannon-information and transformation which in all fairness, may not be commonplaces for Blackmore nor for others writing memetics. It also assumes something like Gerald Edelman's (1992) extended Theory of Neuronal Group Selection (neuronal group Darwinism) as one basis for transduction of received input into active neurophysiological quasi-self-organizing, transformative and reproductive humanimal (i.e. human physiological) processes.

#### **3.2 Not just Three Essentials for Darwinian Evolution but Six**

This extension certainly does not seriously invalidate Blackmore's arguments, but her arguments are rendered somewhat imprecise by the neglect of the additional three factors necessary for Darwinian evolution. It is stated p.10 that "*Darwin's argument requires three main features: variation, selection and retention (or heredity).*" This is true and necessary but according to Williams (1994) and others, *not sufficient*. For neo-Darwinian evolution, as now understood, six features are necessary. They all it seems to me, should also be quite important to think about for understanding meme and memplex evolution:

It must depend on a ***population of patterns*** of some type (they can be patterns of activity, not just of objects).

***Copies*** are made of these patterns (the smallest reliably copied pattern defines the basic unit of replication - e.g. a meme).

***Variation*** - The pattern-units must occasionally ***vary*** (due to mutation copying errors, or reshuffling of parts).

***Competition*** - Variant patterns must ***compete*** to occupy some ***limited space*** (e.g. someone's CNS).

**Selection** - The *relative* reproductive success of variants must be influenced by the environment (which also usually co-evolves with the organisms - e.g. Dawkins' Extended Phenotype theory).

**Heredity** -The *makeup* of the *next generation* of the population must depend on which variants survive to be copied.

It is worth noting that the main environment for most memes is constituted by other memes and especially by memeplexes (& "selfplexes") (- just as the main environment for any gene are other neighboring genes). Bhaktin's question: "*Who is doing the talking?*" (Wertsch, 1991) is also relevant for specifying the selective environment of a meme or memeplex.

### 3.3 Not just One Emergent Memetic level but about six memeplex levels

All memes are not created equal, and especially all memeplexes are not created equal. I have argued (Boyd, 1997) that at least six different levels of memeplexes have emerged historically in co-evolution with populations of informationally coupled humanimals (*Homo sapiens*) and their/our environments:

The *simple viral* aesthetically attractive, pseudo sustenential (or actually sustenential/symbiotic) memes, { e.g. Happy-faces :-), apples, paper-money}. above these emerged *negotiative* ethically and morally attractive memeplexes, {e.g. "do-as-you-would-be-done-by"} above those what I refer to as *conjugative-integrative identi-memeplexes* these are what Blackmore has named "*selfplexes*"- which compose our ordinary personal identities. Emerging above the selfplexes that we are addicted to, is another level of those memeplexes which can free us from addictions - "*liberative*" memeplexes. {e.g., ignore this sentence!}

Then emerging evolutionarily from those are the "*scientosophic*" (i.e. scientific and philosophic- or universal science) memeplexes (- including those of Blackmore's book itself) which can liberate us from superstitious memeplexes, and give us good heuristics and/or generative predictive models and above those emerge what I am now inclined to call: "*existential*" hope-despair imbuing and sustaining, (at best perhaps "*sympivial*" = harmonious-Life-unity-hope survival-promoting) memeplexes.

### 3.4 The trouble with one "selfplex" per customer

The trouble with "selfplex" is that it probably lumps together a number of identity component memeplex entities all hosted by humanimal bodies. For most people the self may usually feel singular and integrated, but it is very apparent to others that we often actually exhibit at least several different personae (e.g. Rowan, 1990). The extreme cases are those of Multiple Personality Syndrome on the one hand and the wonderful self-integration of some great figures or saints on the other. For the rest of us we serve as stages on which various personae or identi-memeplexes take their turns at performing and contending with one-another. We each are a little cast of selfplex actors, if you like.

## 4 The Meme Machine from a Critical Realist Stance

### Is Blackmore's unstated philosophical stance possibly Critical Realism?

The presuppositions of a scientific essay are almost always as important as what is explicitly stated. It is a little difficult to determine the philosophy of science which Blackmore espouses. It is clearly not naive realism, nor superficial observational empiricism, nor positivist operationalism. The actual methodology used is unstated. It might very possibly be either Methodological Pragmatism (Rescher, 1997), constructivism (Von Glaserfeld, 1987), or more hopefully Critical Realism (Bhaskar, 1989).

Memetics, this new causal historical explanatory theory which can be used to generate models of particular source systems, is expounded by Blackmore in a particular qualitative way which might fit all three philosophies of science. Some of the scientific commonsense pre-suppositions of Memetic theory, and some of the vested interests which memeplex evolution models (e.g. of a religion) might harm or benefit are clearly indicated - which taken together might allow categorizing her work as a creative foray in Critical Realism. Adding more critical emphasis on identifying the kinds of memeplexes which are the reproductive forms of domination that Marx identified and named 'reification' and 'alienation' would nicely put the work in the mainstream of Critical Realism today.

If we start from a 'Critical Realist' philosophical position, i.e. An ontology of the real world as structured differentiated and ever changing, and an understanding of scientific activity as the continual process of the empirically controlled retrodiction of explanatory structures from the manifest phenomena which are produced by them, then Blackmore in the *Meme Machine* is doing science. The possibly real explanatory structures are humanimally embodied 'memeplexes' and in particular 'selfplexes' which are proposed to account for a large part of the phenomena of human behavior. They sustain questioning, and thus are partly validated by her explorations of accepted knowledge.

The *Meme Machine* also participates in the critical emancipative transformative agenda of Critical Realism. Emancipative critique according to Bhaskar (1989) consists in asking to what extent are enduring structures being merely reproduced in novel forms, and to what extent are they being transformed to replace un-needed, unwanted and oppressive sources of determination by wanted and empowering ones. Memeplex reproduction and propagation is certainly often unwanted and oppressive, but is it also the basis of human agency and indeed the human spirit? Blackmore seems to contend effectively that it is such. What emerges is a rather odd kind of human agency - the agency of millions of complexly interacting memes embodied in various humanimal nervous systems and artifacts.

This poses a problem for Critical Realists with an emancipative concern such as Bhaskar, because if memetic processes are the 'real' underlying generative processes of culture and identity, then the most important needs of human beings as such arise from our biologically evolved imperatives to reproduce and propagate often totally incompatible competitive memeplexes. Does not emancipation now mean that everyone needs to be liberated from those parts of their own "selfplexes" which cannot propagate symbiotically with those of others and

with our biophysical niches? This selfplex competition survival question is not explicitly raised by Blackmore, but it is I think implicit in her exposition.

## 5. Conclusion

All definitions of 'human nature' have been, and are, very political in the sense that their deployment advantages some sorts (classes?) of people at the expense of others. As Berry, (1986) concludes: "*Any theory (of human nature) itself is part of the problem and constitutive of the issue.*" Who then benefits, and who malefits, from the memetic definition of human nature being purveyed by Susan Blackmore? Critical Realist social scientists will benefit. Socio-cultural engineers will benefit. educators will benefit. POST-post-modernist engaged atheist intellectual selfplexes are bound to benefit. On the other hand various sorts of 'True-believers', including possibly even 'true believers' in Memetics, will suffer as this work, with its liberative anti-magical reframing effects, climbs up the best-seller list (and I think it should).

## References

- Barkow, J., Cosmides, L., Tooby, J. (1998) *The Adapted Mind*. Oxford: Oxford University Press.
- Bhaskar, R. (1991) *Philosophy and the Idea of Freedom*. Oxford: Blackwells. See also the Website for critical Realism <<http://www.raggedclaws.com/criticalrealism/>>
- Bhaskar, Roy (1989) *A Realist Theory of Science*. London: Verso.
- Berry, Christopher (1986) *Human Nature*. Atlantic Highlands, Humanities Press. {Any "theory (of human nature) itself is part of the problem and constitutive of the issue."}
- Blackmore, S. (1996) *In Search of The Light; The Adventures of a Parapsychologist*. Amherst NY. Prometheus Books.
- Blackmore, S. (1999) *The Meme Machine*. Oxford, Oxford Univ. Press.
- Boyd, G. (1997) Excavating the Emergent Levels of Cybersystemics with a view to Concerted Progressive Action. *Systemica*, 11:29-42. BKS/ Dutch System Group.
- Brodie, R. (1996) *Viruses of the Mind*. Seattle, Integral Press. Available June 2000 from <[info@broditech.com](mailto:info@broditech.com)>
- Cees-Speel, H. (1999) On Memetics and Memes as Brain-Entities - a commentary on Gatherer's paper: Why the 'Thought Contagion' Metaphor is Retarding the Progress of memetics, *Journal of Memetics - Evolutionary Models of Information Transmission*, 3. <[http://jom-emit.cfpm.org/1999/vol3/speel\\_h-c.html](http://jom-emit.cfpm.org/1999/vol3/speel_h-c.html)>
- D'Amasio, A. (1999) *The Feeling of What Happens; body and emotion and the making of Consciousness*. New York: Harcourt Brace.

- Dennett, D. C. (1984) *Elbow Room; The Varieties of Free Will Worth Wanting*. Cambridge MA: MIT Press.
- Edelman, G. (1992) *Bright Air Brilliant Fire; on the Matter of Mind*. New York: Basic Books.
- Edelman, G. & Tononi, G. (2000) *A Universe Of Consciousness; How Matter Becomes Imagination*. New York: Basic Books.
- Jacoby, R. (1975) *Social Amnesia*. Boston: Beacon Press.
- Lynch, A. (1996) *Thought Contagion; How Belief Spreads Through Society*. New York: Basic Books.
- Rescher, N. (1977) *Methodological Pragmatism*. Oxford, Oxford Univ. Press.
- Rowan, J. (1990) *Subpersonalities; The People Inside Us*. London: Routledge.
- Von Glaserfeld, E. (1987) *The Construction of Knowledge: Contributions to Conceptual Semantics*. Seaside CA: Intersystems Publications.
- Wertsch, J. V. (1991) *Voices of the Mind; a Sociocultural Approach to Mediated Action*. Cambridge MA: Harvard University Press.
- Williams, C. (1994) Evolution. *Scientific American*, Oct. 1994:105-106.

## INTRODUCTION TO THE CRISIS OF THE MIND

By  
Richard Brodie

**What a waste it is to lose one's mind, or not to have a mind is very wasteful.**

—Dan Quayle, (Mutating the memes in the United Negro College Fund's motto, "A mind is a terrible thing to waste.")

There *is* some good news in this book. So before I get into how mind viruses are spreading wildly throughout the world—infesting people with unwanted programming like the Michelangelo computer virus infects computers with self-destruct instructions—I'll start with the good news.

The good news is that the long-awaited scientific theory unifying biology, psychology, and cognitive science is here. An interdisciplinary effort by scientists in all those fields over the last 20 years or so—really back to 1859 and Charles Darwin, if you like—has produced a new science called **memetics**.

The science of memetics is based on evolution. Darwin's theory of the evolution of species by natural selection utterly transformed the field of biology. Scientists are now applying modern evolutionary theory to the way the mind works, the way people learn and grow, the way culture progresses. In so doing, the field of psychology will ultimately be as transformed by the scientists researching memetics as biology was by Darwin.

For those of us who yearn to understand ourselves, learning about memetics gives us a huge amount of satisfaction. I also believe that people who understand memetics will have an increasing advantage in life, especially in preventing themselves from being manipulated or taken advantage of. If you better understand how your mind works, you can better navigate through a world of increasingly subtle manipulation.

Now the bad news. The bad news is that this book raises more questions than it answers. In particular, memetics has uncovered the existence of **viruses of the mind**, but gives us few insights into what to do about them.

Viruses of the mind have been with us throughout history, but are constantly evolving and changing. They are infectious pieces of our culture that spread rapidly throughout a population, altering people's thoughts and lives in their wake. They include everything from relatively harmless mind viruses, such as miniskirts and slang phrases, to mind viruses that seriously derail people's lives, such as the cycle of unwed mothers on welfare, the Crips and Bloods youth gangs and the Branch Davidian religious cult. When these pieces of culture are ones we like, there's no problem. However, as the Michelangelo computer virus programs computers with instructions to destroy their data, viruses of the mind can program us to think and behave in ways that are destructive to our lives.

This is the most surprising and most profound insight from the science of memetics: your thoughts are not always your own original ideas. You catch thoughts—you get infected with them, both directly from other people and indirectly from viruses of the mind. People don't seem to like the idea that they aren't in control of their thoughts. The reluctance of people to even consider this notion is probably the main reason the scientific work done so far is not better known. As we'll see, ideas people don't like have a hard time catching on.

Further compounding the problem, you don't immediately know whether the programming you get from a given mind virus is harmful or beneficial. Nobody every joined a religious cult with the intention of getting brainwashed, moving to Guyana, and committing suicide. When the teenage Bill Gates caught the poker-playing mind virus at Harvard, was that harmful because it kept him from his studies? Or was it beneficial because it helped sway his decision to drop out, start Microsoft, and become a multibillionaire?

## Paradigm shift

Every so often, the world of science experiences something called a ***paradigm shift***. That happens when one of the basic, underlying assumptions we've been living with changes, such as when we shifted from looking at the universe as revolving around the earth to the earth revolving around the sun. Another shift occurred when Einstein discovered the relationships between space and time and between energy and matter. Each of these paradigm shifts took some time to penetrate the scientific community, and even longer to become accepted by the general public.

### **Viruses of the mind, and the whole science of memetics, represent a major paradigm shift in the science of the mind.**

Because understanding this new science involves a significant change in the way people think about the mind and culture, it has been difficult for people to grasp. As with any paradigm shift, memetics doesn't fit into our existing way of looking at things, of understanding the world.

The trick to learning a new paradigm is to set aside your current one while you're learning rather than attempt to fit the new knowledge into your existing model. It won't fit! If you're willing to set aside your current thinking long enough to consider four concepts, some or all of which may be new to you, you'll be rewarded with an understanding of memetics. With that understanding, I hope, comes a call to action for anyone concerned with the future of human life.

The first concept—the star of the show—is the ***meme***, which I introduce in Chapter 1 and which plays a leading role throughout this book. The meme, which rhymes with "beam," is the basic building block of culture in the same way the gene is the basic building block of life. As I outline in Chapter 2, memes are not only the building blocks of culture on a large scale—making up countries, languages, religions—but also on a small scale: memes are the building blocks of your mind, the programming of your mental "computer."

Second is the concept of ***virus***. It's well known that viruses exist in biology and in the world of computers. Now we'll see how they show up in the world of mind and culture, the world of

memetics. In Chapter 3, I'll draw parallels between the three different universes that viruses live in to show what we can expect from mind viruses in the future.

The third concept that contributes to this paradigm shift is ***evolution***. Evolution is one of those words that many people use, thinking they are talking about the same thing, but really having different ideas of what evolution is and means. I'll discuss scientists' most current theory of evolution in Chapter 4, and how it applies to memes in Chapter 5.

The fourth concept necessary to understanding mind viruses is the new science of ***evolutionary psychology***. This field examines the biases and mechanisms of our minds that evolved to support our survival and reproduction. Some of these biases take the form of psychological ***buttons*** that can be pushed to penetrate our mental defenses. I called this part of the book "Crisis of the Mind" rather than simply "Introduction" because the former pushes more buttons: it attracts more attention and more people will read it. I called this book ***Virus of the Mind*** rather than ***Introduction to Memetics*** for the same reason.

Currently a controversial topic, evolutionary psychology explores and explains many of the stereotypical differences between men and women, especially in the realm of mating behavior. Chapter 6 is about the mating part of evolutionary psychology; Chapter 7 covers the survival aspect.

Memetics builds on these four conceptual blocks to form a new paradigm of how culture evolved and is evolving. It illuminates a major decision point for humanity:

**Will we allow natural selection to evolve us randomly, without regard for our happiness, satisfaction, or spirit? Or will we seize the reins of our own evolution and pick a direction for ourselves?**

Memetics gives us the knowledge and power to direct our own evolution more than at any time in history. Now that we have that power, what will we do with it?

## A threat to humanity

A mind virus is not spread by sneezing, like the flu, or by sex, like AIDS. It's not a physical thing. Mind viruses are spread by something as simple as communicating. I discuss the ways we get programmed by mind viruses in Chapter 8. In a way, mind viruses are the price of one of the freedoms most dear to us: freedom of speech. The more freedom there is to put forth any communication, the more welcoming the environment for mind viruses.

Some mind viruses arise spontaneously, as I discuss in Chapters 9 and 10; some are created intentionally, as I cover in Chapter 11. But all of them share one thing in common:

**Once created, a virus of the mind gains a life *independent of its creator* and evolves quickly to infect as many people as possible.**

Viruses of the mind are not some far-off future worry like the sun burning out or the earth being hit by a comet. They are here with us now—have been with us since before recorded history—and they are evolving to become better and better at their job of infecting us. We are being infected in some new ways—television, popular music, sales techniques—but also in very ancient ways—education, religious teachings, even talking to our closest friends. Our parents unwittingly infected us when we were kids. If you have children, chances are you are spreading the viruses to them every day.

Read a newspaper? Catch a mind virus. Listen to the radio? Catch a mind virus. Hang out with your friends and shoot the breeze about nothing in particular? Catch one mind virus after another. If your life isn't going the way you would like, you can bet mind viruses are playing a large part. Having relationship problems? Mind viruses take over parts of your brain and divert you from what would give you long-term happiness in a relationship. Having trouble in your job or career? Mind viruses cloud your future and steer you along a career path that supports their agenda, not your quality of life.

Cult religions are springing up everywhere, the result of more and more powerful mind viruses. These cults take control of people's minds and make members engage in bizarre behavior ranging from odd rituals to mass suicide. If you think you're immune, remember: nobody ever set out intentionally to join a cult and have their mind taken over. It's the work of tricky and pernicious mind viruses. And once the founder of the cult starts the process, the virus of the mind takes on a life of its own.

Because of mass media and direct elections, the U.S. and other governments are becoming more and more subject to infection by mind viruses. A politician today cannot be elected without coming up with an effective image that pushes people's buttons and gets the votes. "We're having a crisis and only I can fix it," they say, or, "Those other guys have caused all these problems; surely any change is better than what we've got!" Politicians' well-crafted images are hooks into some of the most elaborate and pervasive mind viruses infecting society today.

What brand of soft drink do you buy? The ones that sell the most cost twice as much as unadvertised store brands. The extra money goes into television advertising, sending out the spores of ever more penetrating mind viruses that literally take control of your mind and make you push your shopping cart over to their shelf. Successfully programming your mind to believe that you prefer that brand, advertising agencies are among the most brazen and calculating of the mind virus instigators.

The unchecked spread of mind viruses shows up most alarmingly in the state of our children today. Starting with the inner cities and quickly spreading, the mind viruses infecting many children are pushing them into hopelessness, single motherhood, and gang warfare. Many children seem to be losing their sense of values and taking off in some very unsettling directions. Chapter 12 discusses the possibility of disinfection for us and our children.

## My agenda

Let me tell you right now, I have an agenda in writing this book, and that agenda is to make a difference in people's lives. Some of the content in this book *could* be used for self-improvement. You might not expect a book about science to include ideas from the self-development field, but the science of memetics deals with the mind, with people's lives. Understanding memetics can naturally help increase the quality of people's lives.

In the first place, I would never have written this book—or my first book, *Getting Past OK*—if I had not intentionally disinfected myself of many of the memes I got as I grew up and reprogrammed myself with new memes. What new memes would you choose to reprogram yourself with, given the chance? That's entirely up to you. I had no idea what that even *meant* when I started this research. Now that I do, I choose to program myself with memes that support my values in life rather than ones that support the agendas of viruses of the mind. You can do that or something different. But you won't have the option to do anything like that unless you understand memetics.

The reason I'm writing this book is that I really *enjoy* making a difference in people's lives. My agenda in writing *Virus of the Mind* is to do just that. I believe that knowledge of memetics is important, and so I'm spreading it. I'm not just writing this book as an intellectual exercise. Although this book is *about* science, it's obviously not a scientific text. It's designed with an intention, and that is to consciously spread the new paradigm of memetics because I think it's important.

### **Consciously spreading ideas you consider important is one way to combat mind viruses.**

Have you ever wondered why life seems so complicated today, more complicated and stressful year after year? One reason is the ever-evolving army of mind viruses, taking a greater and greater share of your mind, diverting you from your pursuit of happiness and due to have an even greater effect on the next generation.

Ever wonder why, with greater and greater progress and technology, life doesn't seem to get simpler, but just the reverse? Every time you're exposed to a new virus of the mind, your mind takes on just a little more stress, a little more confusion.

People are flocking to everything from psychotherapy to the New Age movement to try to relieve the crushing burden of stress. Doctors are more and more certain that excess stress is our No. 1 killer, but experts disagree on what causes stress and how to cure it. The medical community talked of stressed-out "Type A" and laid-back "Type B" personalities, with no clear idea of what caused someone to have one or the other. And even the "Type B's" had stress-related symptoms sometimes. The new science of memetics gives much insight into the problem of stress.

Taking over bits of your mind and pulling you in different directions, mind viruses distract you from what's most important to you in life and cause confusion, stress, and even despair.

They infect your mind, programming you with directions that point you away from where you want to go. Since this all happens unconsciously, all you're aware of is that, as you grow older, life becomes more stressful, less fun, more of a drag, and less meaningful. You may feel your motivation slipping away. You may get less excited about things than you used to. These are some of the effects of infection by a virus of the mind, an infection we cannot avoid entirely short of living in complete isolation from birth.

You *can*, however, begin to disinfect yourself. My hope is that the understanding people gain from this book will be a big first step in that disinfection. But it takes a bit of effort to teach yourself a new paradigm.

## Birth of a new paradigm

It's always been hard for scientists to communicate their ideas to the general public. Science, by its very nature, is an artificial selection of ideas based on rigorous testing of their usefulness *rather than on people's gut feelings*. As such, new scientific ideas tend to rub people the wrong way at first and produce predictable reactions. When Charles Darwin first proposed his ideas on natural selection in 1859, there were several stages of public reaction, stages that any revolutionary new scientific idea seems to go through before becoming accepted:

**1. Complacency/Marginalization.** At first, the new theory is seen as an off-the-wall idea: quaint, but not a serious threat to the dominant world view—perhaps a simple variant of some already known theory. Memetics is graduating from this stage to the next as I write this. Editors of *The New York Times Magazine* of Jan. 22, 1995, picked up on the growing use of the word **meme** and mildly attempted to marginalize it: "A skeptic might wonder what the notion of a meme adds to the paradigm of cultural evolution. Perhaps there's nothing new under the sun." By the end of this book, readers will discover that rather than adding to the existing paradigm of cultural evolution, memetics itself is a new and more powerful paradigm.

**2. Ridicule.** Complacency fades as the new idea refuses to die, resulting in ridicule by people who clearly and laughingly see that it's inconsistent with something they hold to be true. In Darwin's case, contemporaries laughed at Darwin's inability to see the necessity of a Supreme Designer doing the selecting. Darwin was frustrated by his seeming inability to communicate this new paradigm. Similar ridicule of memetics is seen from time to time in the few places memetics is discussed, such as the alt.memetics newsgroup on the Internet.

**3. Criticism.** As the new idea gains acceptance, people who have held conflicting world views for some time, or who have their reputation invested in old paradigms, take off their gloves. Darwinism is still being attacked today by creationists who believe it conflicts with their Truth. It's possible that this book will touch off serious criticism of memetics. If it does, we shouldn't worry; it's the nature of a paradigm shift.

**4. Acceptance.** Finally, enough people make the leap to the new paradigm that it gains psychological as well as intellectual acceptance. Those who understand the new ideas are no longer as alone and unloved as Columbus among the flat-earth believers. The new world agrees

on the new paradigm. Peer pressure starts to work for it rather than against it. It's begun to be taught in elementary schools. Scientists can move on to their next challenge.

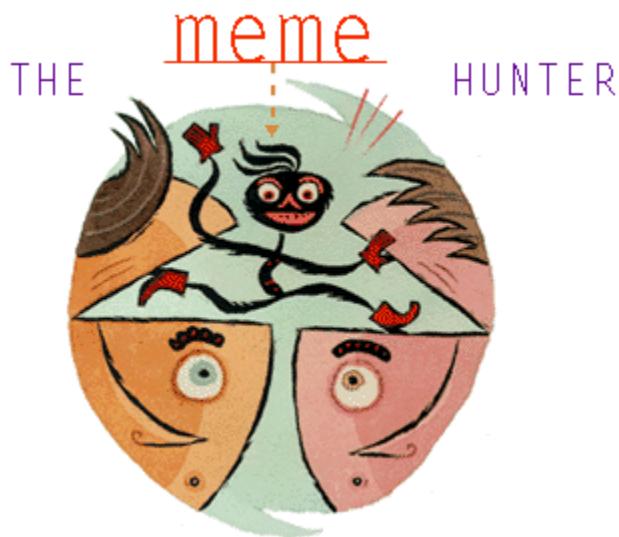
Our minds, it seems, are not well equipped to understand how they themselves work. You, in fact, may at first be very confused or distracted, or suddenly get tired as you read this, or even get angry just from reading these words. Although right now you may think this statement absurd, those feelings and symptoms are actually the defense mechanisms of mind viruses. They have evolved to be very protective of the parts of your mind they have stolen, and any attempts to cleanse yourself of them can trigger reactions.

If you experience one or more of these reactions while reading this book, don't worry: the reaction will pass if you ride it out. If you do, you'll be rewarded with a powerful tool for your future and the future of humanity.

## THE MEME HUNTER

BY  
ANDREW BROWN

*A British psychologist prowls for hard evidence that memes -- ideas that reproduce genetically, like viruses -- actually exist. What's one of the prime habitats? The Internet.*



Suppose that every thought you have -- including this one -- is an autonomous parasite in your brain: a pattern of brain cells that copies itself from mind to mind.

Congratulations! You've just caught a meme.

A meme, according to Richard Dawkins, who coined the word in his bestseller "The Selfish Gene," is anything a human can remember and transmit. Memes are meant to be the brain's equivalent of DNA. A meme could be an idea or a snatch of music or a dance. So long as it gets copied between brains fairly accurately and competes with other copies for survival, it will do as a candidate for evolution. If memes exist, they have modified the world just as genes have: Genes have made the biosphere; memes have made the memeoosphere, the place where human beings exist.

The idea is catchy -- the "meme" meme is particularly popular online -- but controversial. Dawkins himself has withdrawn from it a little. He said last year: "There are people who take memes seriously and there are people who don't. I sort of sit on the fence, and don't mind seriously one way or the other. That wasn't my purpose in producing them."

But lots of smart people do take them seriously. Daniel Dennett, the leading philosopher of artificial intelligence, writes as if the existence of memes were an established fact. A human being, he says, is an animal infested by memes.

And Susan Blackmore, a British psychologist who has been the scourge of every sort of paranormal nonsense for the last 20 years, has got memes bad. She is trying to work out ways to find them in the wild.

"I want to do away with the notion of 'ideas' and look at brain structures and behaviors," she says. "The brain structures may be hard to find. They probably consist of patterns of activation, which will not be there all the time. But I have talked at you for half an hour -- that's a behavior. You have written down some of my words. That's another behavior. And so the memes have been copied. It is a very loose form of copying -- but we can study the process."

Talking to Blackmore is exhausting and invigorating all at once. Ideas fountain out of her like champagne in an intoxicating stream; it hardly matters whether they are bubbles or not.

"If you take a common-sense view," she says, "humans are units called selves which somehow generate ideas from within. But when you take the meme's-eye view, you see brains as hosts for memes. We don't own or generate 'our' ideas. Nor are they working for our purposes. We can imagine meaning in them, but really all that's going on is imitation and storage," she says.

These ideas fit in with Blackmore's Buddhism, where the human personality is no more than a swarm of causes and consequences, temporarily bundled in a body. But they also fit with her sense of a crisis in academic psychology: "I've spent 25 years in psychology and it's an absolute mess. We don't have a decent theory of emotions, or of motivations."

Meme theory seems to offer a way out of this. If beliefs prosper for their own reasons, and not for the good of the organism that contains them, then this would explain the inexhaustible capacity of the human mind to produce bad ideas and disastrous plans. It does not matter to the success of a belief if the original brain that carries it dies, provided it has been spread to more brains in the meantime. In Daniel Dennett's slogan: "A scholar is just a library's way of making another library."

Blackmore says that a good test for the independent existence of memes would look for "behavior that is not in the interest of the person -- but purely in the interest of the propagation of the memes."

The classic example of this spread is religious: "Martyrdom" originally meant "testifying." It's just that handing out tracts or making speeches turned out to be a less effective marketing method than getting chewed by lions in public. "The blood of the martyrs is the seed of the church," as the meme has it.

All this leaves out the consideration of what a meme actually is: how you can identify an individual, transmissible unit of culture. And this is a huge difficulty for meme theory, one that has led most people who have considered the idea to attack it. For everything we know about human culture suggests that it is hugely different from the genetic system.

John Maynard Smith sides with Dawkins against Stephen Jay Gould in battles over biological evolution, but has nonetheless expressed very clearly one of the main theoretical objections to

memes: "Two features are necessary for any genetic system that is to support adaptive evolution. The system should be digital, and it should not support 'the inheritance of adapted characters.'"

Genetic copying is digital -- the DNA strings code for 20 discrete amino acids, and nothing in between -- while cultural copying is analog and fuzzy. Cultural evolution is Lamarckian -- ideas are modified by experience before being passed on -- whereas it is a central dogma of biology that evolution cannot be Lamarckian. Genes are not selected directly, but as a result of their effects; a gene that makes an animal grow or behave in a certain way will be copied (or not) as a result of the behavior or capacities of the animal that carries it. (The animal's physical and behavioral characteristics are known in the jargon as the phenotype; the DNA that encodes the instructions to build the phenotype is the genotype, so that biologists say that genotypes are selected for the phenotypical effects.) But meme candidates are selected directly.

Culture appears to be a ladder -- especially to a scientist like Dawkins -- whereas evolution is a bush. Stephen Jay Gould calls memes "a meaningless metaphor"; virtual-reality pioneer Jaron Lanier thinks they are dangerous nonsense that promotes sloppy habits of thought. Yet it is all too easy for meme believers to dismiss their opponents as people infested with mind viruses -- as if they were not in the same situation themselves.

None of these difficulties bothers Susan Blackmore. She takes her strength from a syllogism: If the Darwinian algorithm is the only way we know to generate order and complexity in nature, which it is; and if our minds are full of order and complexity, which they are; then there must be a Darwinian algorithm operating on something inside them, and that something we shall call memes.

"The Darwinian algorithm can run whether you have got digital or analog information," she says. "It will run better and with higher fidelity if it is digital. But it can run over analog information."

Language is an analog medium, not a digital one, but Blackmore sees the emergence of language as a prime candidate for memetic evolution. "There is a problem of design in language: not just at the deep level of universal grammar, but at the observable level of the languages we actually speak. Why are there so many languages, and why do they stay so separate?" She cites research to show that after 40 years of mass television, the differing regional vocabularies of Britain are more divergent than before. Something is clearly combating the homogenizing effect of TV, and she believes it is the force of memetic evolution.

"Language allows for more accurate copying of thought as it grows more precise. Writing makes copying still more accurate, and thus will increase the size of the memosphere," she argues. The Internet is the most recent expansion of this force.

"What's driving it? Is it our genes, or is it our individual interests? Memetics says that actually it is all in the interests of the memes. What has happened with the development of the Internet is a huge step toward high-fidelity copying -- with just enough errors to make evolution possible. We would expect that to happen because it is in the interests of memetic ambition, and it is completely irrelevant to our happiness. All we're doing is acting as a selection environment."

But memes may not need us to survive, any more than our genes any longer need the bacteria in which their ancestors originated. It is not that we will at some stage create artificial life or artificial consciousness, but that it will create successors to us.

"The more inefficient the technology, the more it relies on humans. But as the technology advances, humans become less and less important. Books made humans less important as a medium for the survival of information. We ought to be able to predict from a memetic analysis of what sort of hardware will best propagate the most memes with the highest fidelity -- and that is what we will find ourselves building."

Blackmore and I were both, by this stage, sitting at open laptops, typing and talking in a synchronous overload; and for a moment I suspected that the laptops were using us to communicate, rather than the other way round.

"The point at which it really takes off," she said, "will be the point at which you have robots which directly imitate other robots in really complex ways. Nothing in artificial intelligence does that now, but when it does, we will have truly human-independent memes."

Only we won't have them. It will be the other way round -- if they can be bothered with us.

*July 10, 1997*

*Andrew Brown writes about religion and other things people take to be true.  
Do memes exist? Come add to the meme pool in Table Talk's Digital Culture discussion area.*

# **THE SIX ESSENTIALS? MINIMAL REQUIREMENTS FOR THE DARWINIAN BOOTSTRAPING OF QUALITY**

**By**

**William H. Calvin**

University of Washington,  
Seattle, WA 98195-1800. USA.  
WCalvin@U.Washington.edu

## **Abstract**

Selectionism emphasizes carving patterns, memes remind us of minimal replicable patterns, but a full-fledged Darwinian process needs six essential ingredients to keep going, to recursively bootstrap quality from rude beginnings. While there may be situations ("sparse Darwinism") in which a reduced number suffice, another five ingredients, while not essential, greatly enhance the speed and stability of a Darwinian process. While our best examples are drawn from species evolution, the immune response, and evolutionary epistemology, the Darwinian process may well be a major law of the universe, right up there with chemical bonds as a prime generator of interesting combinations that discover stratified stabilities.

## **1 Introduction**

Since Richard Dawkins' *The Extended Phenotype* [7] got me to thinking about copying units in the mid-1980s, I have been trying to define a cerebral code (the spatiotemporal firing pattern that represents a word, image, metaphor, or even a sentence) by searching for what can be successfully replicated in the brain's neural circuitry, a minimum replicable unit.

I indeed found such circuitry (it implies that the firing pattern within several hundred minicolumns of neocortex, contained in a 0.5 mm hexagon, is such a copying unit). But to explore creativity in higher intellectual function, I wanted to see if the resulting copies could compete in a Darwinian manner, the process shaping up quality as it goes. And that forced me to try and boil down a lot of evolutionary biology, attempting to abstract the features that were essential (for what I came to call "the full-fledged Darwinian process") from those that merely contributed to speed or stability.

This isn't the place to describe the neural outcome - it's in my book *The Cerebral Code* [5] and, more briefly, in the seventh chapter of my other 1996 book, *How Brains Think* [4] - but this does seem an appropriate place to review what I started calling "The Six Essentials." They seem applicable to a wide range of problems within memetics (Hofstadter [12]) as the field attempts to cope with evolutionary models of information transmission. For a more general history of memetics, see the useful bibliographies in [note 1]; I will only mention a few (mostly cautionary!) contributions from neuroscience along the way.

## 2 Selectionism

'Looking back into the history of biology, it appears that wherever a phenomenon resembles learning, an instructive theory was first proposed to account for the underlying mechanisms. In every case, this was later replaced by a selective theory. Thus the species were thought to have developed by learning or by adaptation of individuals to the environment, until Darwin showed this to have been a selective process. Resistance of bacteria to antibacterial agents was thought to be acquired by adaptation, until Luria and Delbrück showed the mechanism to be a selective one. Adaptive enzymes were shown by Monod and his school to be inducible enzymes arising through the selection of preexisting genes. Finally, antibody formation that was thought to be based on instruction by the antigen is now found to result from the selection of already existing patterns. It thus remains to be asked if learning by the central nervous system might not also be a selective process; i.e., perhaps learning is not learning either.' Niels K. Jerne [15]

The term "selectionism" covers a wide range of cases, ranging from fancy biology with sexual selection to examples that are called "selective survival" because they lack any notion of replication. Brain development offers many examples of this simple end of selectionism. The environment affects survival of cells and their interconnections, and thus helps to create patterns during development (so-called "epigenetic factors in development"). For example, the neural development of a single individual exhibits an enormous overproduction of connections between neurons. Most such synapses (and even long axon branches) disappear, raising notions of selective survival (J. Z. Young's 1964 proposal, that synapses were selectively weakened, was what started this whole line of reasoning [note 2]).

Neurons themselves are also overproduced and die. This again might allow memories to be stored by a carving process (that was Dawkins' 1971 proposal [6]) but it turns out that most neuron death occurs during gestation.

If the removal of connections or cells is carried too far (a common problem in carving wood blocks for printing), there may be no way back (unless, of course, unmodified copies survive elsewhere). Selective strengthening of interconnections, in the face of a culling process, probably accounts for most neural examples of selectionism. Quality can emerge from such carving or selective strengthening. For example, the perception of a speech sound involves the creation of a prototype category that standardizes the meaning despite a range of variations, and selective survival of synaptic connections within a neural feltwork is thought to contribute to categorical perception. But - and this is the important point for memetics - there is nothing recursive about this type of quality enhancement. Which of the selectionist examples should also be called "Darwinian"? I won't review the Darwinian claims except to note that, if we are to blame anyone for the frequent confusion of selective survival with the full Darwinian process, we would have to start with Charles Darwin himself, who named his multipart theory (more in a minute) to emphasize one particularly novel aspect: Natural Selection.

I don't want to seem to be prescribing what's "Darwinian" and what isn't, but I think that we must be cautious about ascribing recursive bootstrapping of quality (what I take to be the heart of the

matter, what makes evolution so interesting) to any process that has a few elements of the process that Darwin and followers have worked out over the last 160 years. Simulations may eventually demonstrate a semidemihemiquasi-Darwinian bootstrap, another self-organizing process that gets better and better - but, until the capacities of sparse solutions are well demonstrated, caution is in order.

### 3 Sparse Darwinian Possibilities

There are two "halfway houses" which may prove to be more interesting than environmental carving of patterns. First, since brain development (to continue the earlier story) is never really over (it just slows down, and the gene repertoire may shift), and since new synapses may form during adulthood, one is initially reminded of a biological population with replacement and growth - and Darwinian shaping up. But observe that there isn't a pattern being replicated with variations; there isn't a population of such patterns competing with other patterns, etc. - which is what population usually means, not merely a number of constituents comprising the pattern being carved.

While Gerald Edelman [note 3] has such a population lacking patterned individuals, he goes beyond selectionist carving in an interesting, nontraditional way: he has a notion of interacting maps, that shape up one another in a manner rather like the sometimes creative back-and-forth interactions between author and editor (my analogy, not his - as is my perhaps shopworn name for it, revisionism). I have a difficult time identifying either an individual unit, or a distinctive copying mechanism for it, in Edelman's lots-of-neurons notion of a "population," even if his re-entrant loops are reminiscent of generations. His differential amplification via re-entrant loops, however, is undoubtedly an important process (I particularly like it for the consolidation of episodic memories).

On closer inspection, neither developmental patterning nor Edelman's reentry fits my concept of Darwin's process. Populations - in ecology and evolutionary biology and immunology - usually involve lots of patterned individuals somehow making near copies of themselves, all present at the same time, interacting with one another and with the environment.

Yet analogies always leave something out; we don't expect them to be perfect fits, exactly the same thing. As the poet Robert Frost once said, we have to know how far we can ride a metaphor, judge when it's safe [11]. That's exactly our problem in memetics, and why Edelman's notions have proved controversial. When, then, are we forced to ascribe, to a candidate such as developmental patterning or reentry, the potential for the recursive bootstrapping of quality that we associate with Darwin's process, which we regularly see operating on the biological species and the antibody?

To approach an answer to that question, it will be useful to enumerate what has contributed to Darwin's process, while trying to strip it of the biological particulars - and then ask how well it could limp along with a reduced number of components (what I've started calling "sparse Darwinism").

## 4 The Full-fledged Darwinian Process

The six essentials aren't a settled issue. What I was aiming for, however, were the essential ingredients of an algorithmic quality-improvement process [see consolidation], stated in a way that didn't impose a lot of biological preconditions. I wanted, for example, to avoid making use of the genotype-phenotype distinction, or a universal translation table like the genetic code; I wanted to describe a process, not make an analogy. John Holland's computational technique [13] known as the "genetic algorithm" comes close to what I had in mind, but Holland was trying to mimic recombination genetics in a computational procedure for discovering solutions, and I wanted to abstract more general principles that avoided the presumption of recombination. Since many of us think that (properly defined) the Darwinian process is a major law of the universe, right up there with chemical bonds as a prime generator of interesting combinations that discover stratified stabilities [note 4], we want it to be able to run on different substrates, each with their own distinctive properties that may, or may not, correspond to those seen elsewhere. So our abstraction should fit the species evolution problem, as well as the immune response, but also be independent of media and time scale. Here, paraphrased from The Cerebral Code [5], is what I ended up with:

There must be a pattern involved.

The pattern must be copied somehow (indeed, that which is copied may serve to define the pattern). [Together, 1 and 2 are the minimum replicable unit - so, in a sense, we could reduce six essentials to five. But I'm splitting rather than lumping here because so many "sparse Darwinian" processes exhibit a pattern without replication.]

Variant patterns must sometimes be produced by chance - though it need not be purely random, as another process could well bias the directionality of the small sidesteps that result. Superpositions and recombinations will also suffice.

The pattern and its variant must compete with one another for occupation of a limited work space. For example, bluegrass and crab grass compete for back yards. Limited means the workspace forces choices, unlike a wide-open niche with enough resources for all to survive. Observe that we're now talking about populations of a pattern, not one at a time. The competition is biased by a multifaceted environment: for example, how often the grass is watered, cut, fertilized, and frozen, giving one pattern more of the lawn than another. That's Darwin's **natural selection**.

New variants always preferentially occur around the more successful of the current patterns. In biology, there is a skewed survival to reproductive maturity (environmental selection is mostly juvenile mortality) or a skewed distribution of those adults who successfully mate (sexual selection). This what Darwin later called an **inheritance principle**. Variations are not just random jumps from some standard starting position; rather, they are usually little sidesteps from a pretty-good solution (most variants are worse than a parent, but a few may be even better, and become the preferred source of further variants).

Neural patterning in development is a sparse case: just a pattern and a multifaceted environment. There is no replication of the pattern, no variation, no population of the pattern to compete with a variant's population, and there's nothing recursive about achieving quality because there's no inheritance principle.

## 5 Example: History as a Darwinian Process

History qua history - what it includes, what it leaves out, and how these change over time - provides us with a memetic example of these six essentials at work. Of the many happenings, some are captured in patterned sentences that describe who did what to whom, why, and with what means.

Some of these patterns are retold (copied), often with little confusions (variation) and conflations (superpositions). Alternative versions of stories compete for the limited space of bookstore shelves or the limited time of campfire storytelling. There is a multifaceted environment that affects their success, the association of the described events to those of everyday life. In particular, the environment contains mental schemas and scripts; as Aristotle noted and all four-year-olds demanding bedtime stories seem to know, a proper narrative has a beginning, middle, and end - and so "good stories" fare much better in the memorized environment. (Especially those conveyed by historical novels that strengthen the narrative aspects!) Finally, because historians rewrite earlier historians, we see Darwin's inheritance principle in action: new variations are preferentially based on the more successfully copied of the current generation of historical stories, and so history has a drift to better and better fits to language instincts (such as chunking and narratives) because current relevance is shifting and ephemeral.

After many generations, only those stories of timeless relevance are left alongside the likely-ephemeral contemporary ones.

Quality emerges, in some sense, as in the way that the nine-part epic tales studied by Misia Landau (youth sets out on a quest, fails, returns, sets out again with a helper, survives a new set of trials and tribulations, finally succeeds and returns home, and so on [12]) seem to have emerged in many cultures from the retelling of simpler narratives, generation after generation. Our modern origin stories, the anthropological scenarios about human ancestors during the tribulations of the ice age climate changes, are even said to follow the epic template! Can history, as we know it, run on a reduced set - say, without the inheritance principle? (Imagine storytellers always reviewing a videotape before telling the story again, so variations were always done from an unchanging "standard version.") Certainly, a pattern that copied and varied, with retelling biased by resonances with current memories of the current population, would be impressive - but the anchoring of the center of variation to the standard version would keep stories from drifting very far and prevent the recursive bootstrapping of quality.

Suppose that, instead of eliminating inheritance, we loosened the environmental influence - say, individuals' memories [see episodic] for unique episodes faded within a year. The often-told tales would simply drift, adapting to current concerns, losing those of the antepenultimate generation. It would be about like the whale songs that drift from one year to the next. What you would lose,

lacking a good memorized environment that persisted a lifetime to overlap several generations, would be shaping up of quality (those timeless stories with universal relevance, the resonance with episodes recurring only twice in a lifetime, and so forth).

My first "knock-out mutation" sounds, of course, like what we try to train scholars to do ("Avoid secondary sources! Read the original!"), while my second is merely an exaggerated version of the ahistoricism of preliterate societies (the Navajo emigrated from the Yukon to the American Southwest about 500 years ago, but this great migration has been lost to them, recovered only through a linguistic and genetic analysis of the Athabascan peoples). History, however, is not merely the retention of facts: it involves detecting patterns and attempting to understand them - and this involves making good guesses and refining them. That intellectual endeavor is, I suggest in *How Brains Think* [4], another full-fledged Darwinian process.

Competition between concepts is, of course, one of the ways in which science advances; evolutionary epistemology (Mayr [17]) treats this as a Darwinian process. The advance of science differs from ordinary history because the environment biasing the competition between concepts involves a broad range of testing against reality.

## **6 Nonessentials: Catalysts and Stabilizers**

There are another five features that, while not essential, do notably influence the rate of evolutionary change:

Stability may occur, as in getting stuck in a rut (a local minimum - or maximum - in the adaptational landscape). Variants happen, but they're either nonviable or backslide easily. Systematic recombination (crossing over, sex) generates many more variants than do copying errors and the far-rarer point mutations. Or, for that matter, nonsystematic recombination such as bacterial conjugation or the conflation of ideas.

Fluctuating environments (seasons, climate changes, diseases) change the name of the game, shaping up more complex patterns capable of doing well in several environments. For such jack-of-all-trades selection to occur, the climate must change much faster than efficiency adaptations can track it (more in a minute).

Parcellation (as when rising sea level converts the hilltops of one large island into an archipelago of small islands) typically speeds evolution. It raises the surface-to-volume ratio (or perimeter-to-area ratio) and exposes a higher percentage of the population to the marginal conditions on the margins.

Local extinctions (as when an island population becomes too small to sustain itself) speed evolution because they create empty niches. The pioneers that rediscover the niche get a series of generations with no competition, enough resources even for the odder variants that would never grow up to reproduce under any competition. For a novel pattern, that could represent the chance to "establish itself" before the next climate change, for which it might prove better suited than the others.

There are also catalysts acting at several removes, as in Darwin's example of how the introduction of cats to an English village could improve the clover in the surrounding countryside: The (i) cats would (ii) eat the mice that (iii) attack the bumble bee nests and thus (iv) allow more flowers to be cross pollinated. (You can see why I call these the "Rube Goldberg Variations.")

## 7 The Augmented Darwinian Set

Although a Darwinian process will run without these catalysts, using Darwinian creativity often requires some optimization for speed. In the behavioral setting I analyze in my two 1996 books, quality must be achieved within the time span of thought and action.

Accelerating factors are the problem in what the French call avoir l'esprit de l'escalier - finally thinking of a witty reply, but only after leaving the party. Some accelerating factors are almost essential in mental Darwinism simply because of the time windows created by fleeting opportunities, and so this "augmented Darwinian set" may also prove to be important for other memetic applications of the universal Darwinian process.

## 8 Discussion

I am proposing a standard Darwinian set (six ingredients, in my formulation), with nonstandard cases often described via the sparse and augmented sets. As with Edelman's reentry, some cases may be both sparse and have a novel feature like revisionism (mixed cases). I was delighted to discover that my (neocortical circuitry) candidate process was not only capable of implementing all six essentials, but stability and the four catalysts as well.

At what point can we carry over the traditional implications of the best-studied case, the species-evolution Darwinian process, to a candidate process? My present answer would be: When the six essentials are present, and no obvious stability or relative-rate issue seems to be precluding "progress," we are then entitled to predict that our candidate process is capable of repeatedly bootstrapping quality.

The extent of "coverage" of memetic theories varies widely. For example, I was able to spend much of my last chapter of *The Cerebral Code* discussing the Darwinian implications of minor circuit malfunctions for a broad range of pathophysiology such as seizures, hallucinations, delusions, amnesia, *déjà vu*, and so forth. My point is that candidate processes in other memetic fields are also likely to be judged by similar nonevolutionary considerations, so we must remember that possessing the six essentials is only a "threshold" consideration, mostly relevant to the sorts of quality that can be bootstrapped - and for how long that improvement can continue.

### 8.1 Stratified Stability and Relative Rates

One coverage issue that seems relevant, however, is whether new levels of organization emerge from the candidate evolutionary process. Can, for example, a candidate process form categories?

Can it progress to evolving analogies [note 5] or metaphors? Are these new levels ephemeral, or stable for awhile?

Jacob Bronowski spoke of "stratified stability" and observed [2], "The stable units that compose one level or stratum are the raw material for random encounters which produce higher configurations, some of which will chance to be stable....Evolution is the climbing of a ladder from simple to complex by steps, each of which is stable in itself." Does the process self-limit when reaching an angle of repose, so that piling on another layer is self defeating? Does the process backslide in a catastrophic manner, requiring something like the Weismannian barrier between genotype and phenotype to provide a ratchet?

Relative rates play an important role in any process involving change, one that can trivialize or magnify. Relative rates of expansion are the major principle underlying most household bimetallic-layer thermostats, and it is a familiar principle in development (the way curved surfaces are made is to have two sheets of cells in contact, one growing faster than the other). And we've already seen two examples here: the history example of episodic memories that faded quickly when compared to the generation time and lifespan, and again in (Hofstadter [9]) where climate changes had to be much faster than adaptations could track, if jack-of-all-trades abilities were to accumulate in the face of competition from lean, mean machines.

## 8.2 Repackaging the Essentials

Someone will surely try to condense my six essentials to a phrase more memorable than "a pattern that copies with occasional variation, where populations of the variants compete for a limited workspace, biased by a multifaceted environment, and with the next round of variations preferentially done from the more successful of the current generation." Indeed, Alfred Russel Wallace did a pretty good job back in 1875 ("...the known laws of variation, multiplication, and heredity... have probably sufficed....") [14].

It's just that I make explicit the pattern, the work space competition between populations, and the environmental biases. As noted in [note 4], I'm trying to avoid lumping where I know that splitting is going to be required later, to deal with some important partial cases. Wallace shows us that only three items cover a lot of essential ground, and there are surely other profitable ways to split and lump, if context allows the inference of other factors. A list of essentials - at least one that aspires to some universality - can't omit the context, can't skip over the potential confusions. Bronowski once observed [1] that, even if six sentences might serve to sum up one of his lectures, the rest of the hour was really essential for disambiguating the meaning of those summary sentences. The name of the game here isn't compression but abstraction, an abstraction that is just general enough to cover a number of situations that differ widely in media and time scale - but not so abstract as to lose important associations.

Of course, all the definition in the world can be upset by one little existence proof, a simulation of a self-organizing quality bootstrap that runs on a different set of principles. Until then, we are simply trying to clarify our thinking about the creation-of-quality process we know best, the one first stumbled upon by Charles Darwin.

## Notes

Memetic history is covered in a number of places; see, for example, the excellent webbed bibliographies of Barry McMullin, Hans-Cees Speel, and John S. Wilkins.

John Z. Young, A Model of the Brain [19]. His The organization of a memory system [20] introduces the mnemonic concept in which weakened synapses serve to tune up a function. A later version is his Learning as a process of selection [21]. The modern chapter of mental darwinism starts in 1965 with Daniel C. Dennett's D. Phil. thesis, published as [8] but it all began more than a century ago with William James, Great men, great thoughts, and the environment [14]. Gerald M. Edelman, Neural Darwinism [9] and Bright Air, Brilliant Fire [10] (especially chapter 9 where he addresses his critics). Reviewed by me in [3].

In introducing stratified stability, Bronowski says: "Nature works by steps. The atoms form molecules, the molecules form bases, the bases direct the formation of amino acids, the amino acids form proteins, and proteins work in cells. The cells make up first of all the simple animals, and then the sophisticated ones, climbing step by step. The stable units that compose one level or stratum are the raw material for random encounters which produce higher configurations, some of which will chance to be stable.... Evolution is the climbing of a ladder from simple to complex by steps, each of which is stable in itself." [1]

In searching for Hebb's cell assembly (a committee of neurons whose firing represents a color, word, thought, etc.), I looked for neural circuitry in the brain that was capable of copying spatiotemporal firing patterns. The top layers of the newer parts of cerebral cortex indeed have recurrent excitatory circuitry that should produce synchronized triangular arrays of neurons which extend themselves over a few centimeters, or recruit distant populations via corticocortical pathways. Collections of such arrays constitute a spatiotemporal pattern with great redundancy. The smallest "tile" in this mosaic, which has no redundancy within it, is hexagonal in shape and about 0.5 mm across; it probably has a few hundred independent elements. This is a minimal replicable unit (and a candidate for a cerebral code for, say, a word or concept); a hexagonal mosaic of it can compete with another pattern's hexagonal mosaic for territory in association cortex.

## Glossary

### Ahistoricism

is discussed in the first chapter of my book, The River That Flows Uphill. But Loren Eiseley said it best: "Man without writing cannot long retain his history in his head. His intelligence permits him to grasp some kind of succession of generations; but without writing, the tale of the past rapidly degenerates into fumbling myth and fable. Man's greatest epic, his four long battles with the advancing ice of the great continental glaciers, has vanished from human memory without a trace. Our illiterate fathers disappeared and with them, in a few scant generations, died one of the great stories of all time."

### An algorithm

is a simple systematic procedure for solving a problem, usually involving repeated steps, e.g., long division (try multiplying the divisor by two and subtracting; if the remainder is too large, try 3, etc.).

### The angle of repose

is a geological term for how steep-sided a pile can become before little landslides remove any further additions.

### Bootstrapping

("Pulling yourself up by your own bootstraps") is a term that describes the ability of simple patterns to bring forth more complicated ones. A familiar example is the start-up procedure for a computer: booting involves a simple program stored in a ROM chip that, in turn, reads and starts up the operating system from a hard disk - which, in turn, starts up the application programs.

### Episodic memories

are those of unique events, such as a particular conversation; they're much more difficult to maintain than memories of repeated events. **Consolidation** of memory is the often-fallible process of making short-term "volatile" memories into more permanent long-term memories with an enduring structural basis. I suggest, in chapter 6 of *The Cerebral Code* [5], that Edelmanian revisionism, between a hippocampal map and a cortical map, would be a useful way of "learning" (via repeated trials) what was originally only a unique episode and embedding it in neocortex. Like a photographic development process, it is likely that, in the manner of those three-million-year old footprints that were preserved by volcanic ashfall (cement on the fly preventing the usual erosion and backfilling), both culling and fixation are involved in biological memories.

### Weismann's genotype-phenotype distinction

in biology is not a necessary condition for a Darwinian process, as recent experiments on "RNA evolution" have shown (there isn't a body that lives and dies, carrying the genes along, but rather patterns directly exposed to environmental selection). Envelopes such as bodies (phenotypes) are an example of stratified stability; they nicely illustrate why strict one-trait-at-a-time adaptationist reasoning is insufficient. Genes often live - and die - in a collection called an "individual," which means that success is often via particular combinations of traits.

## References

- [1] Jacob Bronowski, *The Origins of Knowledge and Imagination* (Yale University Press 1978, transcribed from 1967 lectures), p. 105.
- [2] Jacob Browonski, *The Ascent of Man* (Little, Brown 1973), pp. 348-349.
- [3] William H. Calvin, A global brain theory. *Science* 240:1802-1803 (24 June 1988).
- [4] William H. Calvin, *How Brains Think: Evolving Intelligence, Then and Now* (BasicBooks 1996).

- [5] William H. Calvin, *The Cerebral Code*: Thinking a Thought in the Mosaics of the Mind (MIT Press 1996).
- [6] Richard Dawkins, "Selective neurone death as a possible memory mechanism," *Nature* 229:118-119 (1971).
- [7] Richard Dawkins, *The Extended Phenotype* (Oxford University Press 1982).
- [8] Daniel C. Dennett's D. Phil. thesis, published as *Content and Consciousness* (Routledge and Kegan Paul 1969)
- [9] Gerald M. Edelman, *Neural Darwinism* (BasicBooks 1987)
- [10] Gerald M. Edelman, *Bright Air, Brilliant Fire* (BasicBooks 1992)
- [11] Robert Frost, in *Selected Prose of Robert Frost*, edited by H. Cox and E. C. Lathem (Collier 1986), pp. 33-46.
- [12] Douglas R. Hofstadter, *Metamagical Themes* (BasicBooks 1985).
- [13] John Holland, *Adaptation in Natural and Artificial Systems*, revised edition (MIT Press 1992)
- [14] William James, "Great men, great thoughts, and the environment," *The Atlantic Monthly* 46(276):441-459 (October 1880).
- [15] Niels K. Jerne, "Antibodies and learning: Selection versus instruction," in *The Neurosciences: A Study Program*, edited by G. C. Quarton, T. Melnechuk, & F. O. Schmitt (Rockefeller University Press 1967), pp. 200-205 at p. 204.
- [16] Misia Landau, "Human evolution as narrative," *American Scientist* 72:262-268 (May/June, 1984).
- [17] Ernst Mayr, *This is Biology* (Harvard University Press, 1997), p.99.
- [18] Alfred Russel Wallace, "The limits of natural selection as applied to man," chapter 10 of *Contributions to the Theory of Natural Selection* (Macmillan 1875).
- [19] John Z. Young, *A Model of the Brain* (Clarendon Press 1964)
- [20] John Z. Young, "The organization of a memory system," *Proceedings of the Royal Society (London)* 163B:285-320 (1965)
- [21] John Z. Young, "Learning as a process of selection," *Journal of the Royal Society of Medicine* 72:801-804 (1979).

# **TOWARDS A COGNITIVE MEMETICS: SOCIO-COGNITIVE MECHANISMS FOR MEMES SELECTION AND SPREADING**

**By**

**Cristiano Castelfranchi**

University of Siena - Dep. of Communication Sciences  
Siena - Italy  
[castel@ip.rm.cnr.it](mailto:castel@ip.rm.cnr.it)

## **Abstract**

After stressing the autonomy of a cognitive agent relative to social influence, and the importance of cognitive constraints in accepting a given meme, we discuss three specific micro-mechanisms for adopting a given behavior; they differ in the interpretation of the observed behavior and in the motives for replicating it. Tomasello's model of cultural transmission is discussed. Special attention is paid to the role of norms in meme spreading, to the role of social identity and membership, and to inter-group differentiation. Principles supporting the diffusion of *know how* are different from principles supporting the diffusion of *know that*. Cognitive constraints for beliefs acceptance are examined. Not only the adoption of a meme but also its diffusion can be the result of a decision by the cognitive agent; different socio-cognitive micro-mechanisms have different macro-results in meme propagation. Those examples and models are aimed at claiming that the agents' minds are the most relevant selective environment for memes. To understand cultural evolution it is necessary to identify the cognitive principles of the success or selection of memes *within minds*. Memetics can only be cognitive; otherwise it is contradictory and non explanatory.

**Keywords:** memes, cognitive memetics, culture, beliefs, norms, social identity

## **Introduction**

We start from the autonomy of a cognitive agent relative to social influence and the importance of cognitive constraints in accepting a given meme. In this perspective, Tomasello's model of cultural transmission is discussed. On the basis of general principles of cognitive processing (as the role of interpretation and 'understanding' of the input) we introduce three specific mechanisms for adopting a given behavior: 'instrumental adoption' based on means-end reasoning and on practical utility (section 2.); normative adoption, where certain behaviors or values are prescribed by social norms and the agent wants to conform to such norms (section 3.); and the social identity mechanism, where the agent imitates a given behavior of others to be "like them, one of them" (section 4.). Those mechanisms differ one from the other as for the cognitive interpretation of the observed behavior and for the subjective motives for replicating it. Special attention is devoted to the role of norms as spreading memes and as meme-regulators, and to the role of the inter-group differentiation motive. Principles supporting the diffusion of *know how* are different from principles supporting the diffusion of *know that*. Cognitive constraints for belief acceptance are examined (section 6). Finally a more sophisticated model of the cognitive processing of a candidate meme is presented (section 7.) where not only the adoption of a meme

but also its diffusion can be the result of a *decision* of the cognitive agent. Different socio-cognitive micro-mechanisms make different predictions as for their emergent effects; in fact, they have different macro-results in meme propagation. Those examples and models are aimed at claiming that the agents' minds are the most relevant selective environment for memes and the *media* of their transmission. To understand cultural evolution it is necessary to identify the cognitive principles of the success or selection of memes within *minds*. Memetics needs cognitive modeling.

## 1. Cognitive Autonomy and Mediation in Cultural Transmission

Humans are cognitively autonomous [6; 18], i.e., they have a rather good control over their own mental states, relatively to external social influence. This plays a central role in cultural evolution [13; 17]. One cannot make another believe or do whatever he wants the other to believe or do. One should resort to complex strategies, like education, persuasion, manipulation, etc. in order to deal with the *cognitive constraints*, filters, and mechanisms of the addressee.

Autonomous agents believe something only on the basis of what they believe and wish (section 6.); they decide to do something only on the basis of what they believe and wish; they learn to do something on the basis of their motives and of the *perceived* rewards, and on the basis of cognitive processes like attention, memory, association, analogy, abstraction, etc.

These cognitive constraints decide also of meme adoption and transmission; thus there is a crucial cognitive mediation [17; 8; 10] in memes spreading and inheritance [note 1]. The individual mind with its cognitive processes is the fundamental selection environment for memes. For the great majority of memes, it holds that:

*if memes do not win within the mind of the individual, they cannot spread around in the population.*

### 1.1 Limits of these 'cognitive' approach and representations

It is important to keep in mind that 'cognitive' is synonym of 'mental', not of 'conscious' or 'rational'. It is important to analyze and clarify the mental mechanisms of such broad or obscure notions as 'imitation' or 'contagion', but those mechanisms are neither necessarily rational nor conscious. In the following, we will describe different mechanisms in terms of 'beliefs' and even of 'decisions'; however, only the first mechanism has to do with rational instrumental reasoning, and none of them is supposed to be necessarily conscious. However, unfortunately, this is not enough: our framework will keep some rationalistic flavor, because of its formulation in terms of explicit beliefs and goals.

It is worth specifying that:

on the one hand, I do not exclude other -more primitive- mechanisms for imitation or for 'contagion', based on mere automatic rules [note 2]; I simply would like to have some clear model of them in terms of internal processes and devices (rule-based reactive agents, conditioning, etc.);

on the other hand, I admit that my approach is limited: one should provide a more complete theory, for example about the role of affective reactions, on intuitive appraisal in adopting ideas, attitudes or behaviors, and suggest a theory of implicit beliefs and goals in memetic processing. Even the notion of 'decision' can be interpreted at different levels of cognitive complexity: at a more 'procedural' level, for example simply as a test, an "if/then/else" step in a flowchart; or at a more 'declarative' level, as a conscious deliberation based on explicit evaluations. The proposed mechanisms are presented in a very explicit and propositional way, but their really important message is as follows: mental interpretation and internal representation (for example, in declarative terms) are crucial in the memetic process; there are different cognitive mechanisms accounting for the acceptance and the spreading of a meme by an agent.

## 1.2 Diffusion vs. Tradition

An interesting model of cultural transmission that gives some room to cognitive processes is Tomasello's ratchet model, (see Appendix) although it is too focused on learning and children. It underestimates the role of synchronic or horizontal cultural diffusion as a fundamental condition also for an efficient diachronic or vertical diffusion (tradition, inheritance). Transmission of cultural features does not occur only from one generation to the next (like in this model). In fact in order to spread from one 'generation' (and not from one individual) to another, the trait is supposed to be diffused within the first generation. Clearly:

*the diachronic transmission is function of the diffusion in a given population (and vice versa)*

'diffusion' both in terms of extension and duration: the greater the number of individuals in a given population that are vehicles of a given cultural trait (behavior, idea, artifact), and the longer the period of their having/exhibiting/using such a meme, the greater the probability of the same meme in the next generation of the same population.

(conversely, transmission from a previous generation to a new generation, i.e. inheritance, is one of the causal factors that strongly affects the diffusion of a given trait in a given population).

Also considering the synchronic diffusion, the notion and mechanism of 'imitative learning' - although in the rather clear (and cognitive) version of Tomasello - is not enough. The problem is *why another adult individual should 'adopt' or 'accept' a mental representation, a behavior, an artifact from another individual or group*.

We believe that there is no unique answer, that is, there are several cognitive mechanisms (beliefs + motives) through which this happens. In the next sections we will examine some of these different mechanisms, but first we have to describe a *cognitive-mediation* view of 'imitation'.

## 1.3 A basic model of replication in cognitive agents

A cognitive agent has three crucial and defining features (among others):

### i. Cognitivism

Vis à vis information, events, situations, inputs in general, the agent interprets them, and reacts in some way not to the event or situation but to the subjective perception of it; i.e., the mental representation of it; the meaning the event has in its meaning and value system; it reacts to the meaning that it attributes to it. [note 3]

## ii. Reasons

A cognitive agent activates, selects, prefers, pursues, gives up goals on the basis of what it believes [6]. In other words, it has 'reasons' for what it does.

## iii. Purposes

A cognitive agent is a goal-directed agent (endowed with intentions, planning, and deliberation abilities, ...); its behavior is in fact 'action' aimed at certain anticipated results (mental representations) and is controlled and motivated by them (these representations).

From these basic assumptions it follows a rather obvious and simple model of the replication of behaviors in Cognitive Agents. As stressed by Conte [17] "behavior does not spread automatically but through the agent' mind":

*agents have to observe a behavior; to interpret/understand it; and then to have their own motives and reasons for performing (and store and repeat) such a behavior.*

Only on the basis of their understanding and of their motives and reasons they can replicate the same behavior or some accidental or adjusted variation of it.

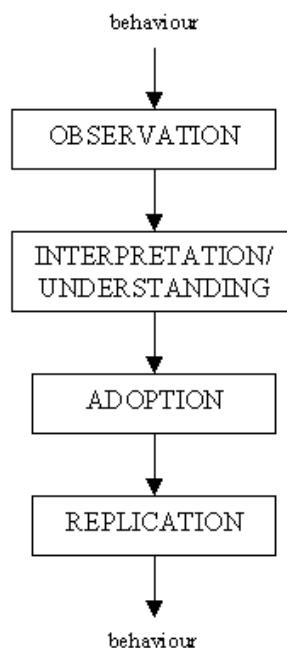


Figure. 1

It is important to notice that

a) the behavior is 'on purpose' but its replication is not necessarily intended or motivating [note 4];

b) analogously, the 'spreading' effect is not necessarily known or intended by the agent.

#### 1.4 Three mechanisms for meme adoption and replication

In the following sections we will give three examples of the application of this model: what we will call the means-ends or instrumental reasoning; the normative reasoning and compliance; the Identity or membership based reasoning.

These three mechanisms are different realizations of the same abstract model (Fig. 1) with its crucial steps. They differ one from the other precisely because the inference/interpretation (understanding) of the observed behavior is different; and reasons and motives for replicating it are very different one from the other. They could be applied to one and the same observed behavior, while producing quite a different cognitive and social processing [note 5]. Let us adopt as our working example the following one: suppose we are in a new country and culture and we observe that they -unlike us- use fork and knife for peeling fruit. The following are three possible interpretations.

## 2. Instrumental Adoption: the 'practical problem-solving' mechanism

This mechanism applies to what we could generally call 'means' that is actions, plans, recipes, rules, procedures, tools (*Know How transmission*)

Tomasello's answer to our starting question (*why should another adult individual 'adopt' or 'accept' a mental representation, a behavior, an artifact by another individual or group*) is the following one. Imitative learning of a feature x is not simple emulation or echoing, but it is based on some *understanding* of the purposes and reasons of the model M exhibiting it.

Let's characterize this mechanism as follows:

Given an Adopter A, a trait x and a Model M,  
conditions for 'ADOPT A x from M' are:

**Bel A (Use M x for p)** [note 6]

which implies that **Bel A (Goal M p)** and

**Bel A (Bel M (Good-for x p))** [30] [note 7]

**Bel A (Good-for x p)** (shared *evaluation* of the mean)

**Bel A for-all y (Better x y for p)** [note 8]

**Goal A p**

Then 'Adopt A x for p' means that:

*A stores beliefs about x as a good-means for p in its 'know how' for p. When goal p becomes active, A will consider x as a possible means (plan, action, tool) for p, better than others, and then probably A will use x, as M does.*

In the fruit example the agent reasons as follows: "In this way, one doesn't dirty one's hands and face! It's a good solution: I'll do the same". The Adopter has *reasons* (the evaluation of y) and *motives* (its need and problem) for adopting and replicating the behavior.

In sum, *individuals accept new behaviors, plans or tools as better solutions for their own problems, as good means for their goals*; the group diffuse and preserve (memorize) and transmit the best (discovered) solutions. This is also (one of) Hayek's view(s) about tradition, its validity, and unintended cooperation in constructing social institutions, rules, and technology (what Tomasello [36] calls the third kind of sociogenesis of culture, i.e. of its cooperative construction).

This 'practical problem-solving' mechanism seems very close to what Hedstrom [27] calls 'rational imitation' [note 9]. However, it is worth stressing that the agent's understanding of the intention and plan of the observed Model, and its evaluation of the use and advantage of the new behavior or method it is not a trivial problem [15].

Notice that for this kind of "imitation" even one single example/model can suffice. However, usually the greater the number of sources of a belief the more credible that belief (section 6.). In the same vein, the greater the number of agents who adopt a certain "solution" (tool, recipe, plan, etc.) the more influent that solution becomes: the more it is diffused the more it looks validated and good.

### **3. Normative Adoption: The normative character of cultural transmission**

It is underestimated that cultural transmission - including the *adoption* move - has a very strong 'normative' (and more precisely 'prescriptive') character. It is not simply matter of the Adopter's interest and initiative. Also M and the group care about such an adoption. It is not simply a matter of technical/practical reasoning ("is this a better solution for my needs/problems?"). Completely different, non-technical criteria guide the adoption of cultural features, like a behavior in given circumstances, or the use of a given plan or tool. Culture is a set of expectations and prescriptions on the members of the group: each member expects and wants the others behave according to cultural traditions (prescriptions) about what to do (think and feel), when and how.

The satisfaction of those expectations (social approval motivation, membership-collective identity motivation) and the conformity to those norms (for several motives [18]) is one of the main reasons why people adopt a cultural trait when entering a given culture.

Again, this presupposes a rich cognitive representation, an 'understanding' of the others' minds and behaviors: more precisely, it presupposes the *recognition* of such a Model, example, expectation and advice *as a norm*, and its normative adoption [18].

By adopting x, A knows and intends to conform to a norm, and -by doing so- it confirms it and the authority of the group.

This is also a message that A's behavior is sending to the others: "I intend to 'respect' the norms, our habits and traditions, and our authorities" (3.1). This 'message' can be the only real function of the behavior that becomes completely 'ritual' and loses any practical function and reason. For example, why not cut fish with a knife? this is just a ritual adhesion to a custom.

### The normative mechanism

**Bel A (Norm x on group Y)**

**Bel A (member A of Y)**

**Goal A (Adoption of N impinging on A)**  
(for several possible motives/ higher goals)

**Goal A (Adopt A x)**

In the fruit example the Adopter's reasoning is as follows: "Probably this is a social norm, maybe handling fruits is considered a sacrilege. I had better respect this custom or rule". Notice that the Adopter also believes that everybody does so; or at least that it is very diffused. In fact,

*a behaviour diffused in a group/population is more likely to be interpreted (recognised) as a norm of that group.*

This as for its *recognition* as a norm. Moreover, as for replication (in case of norms: *obedience*)

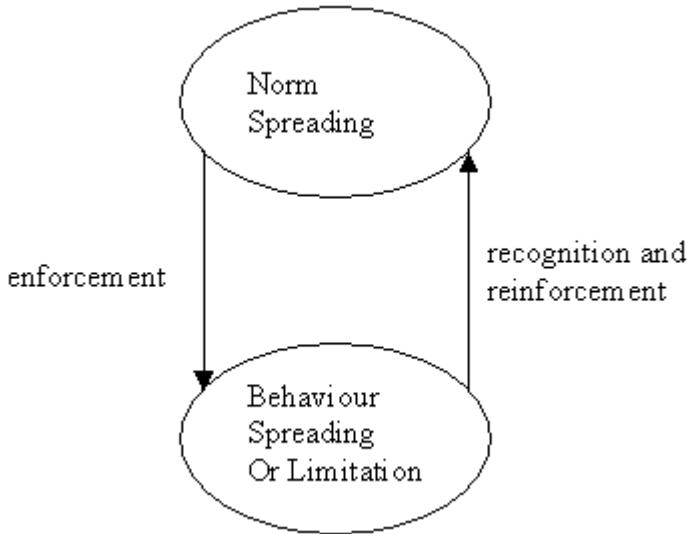
*the more the agent observes that the norm is respected, the more the agent is leaning to conform to it (diffusion).*

Among the motives for adopting a norm (obeying it after and because we recognized that it is a norm) there are not only social approval and sanction avoidance, but also possible 'terminal' motives like "norms must be respected". There can also be a motive like: "to make others conform to the norm: i.e. to be a M for the others". In this case the memetic *function* of conforming to a norm becomes an explicit intention in the mind of the agent [9; 11].

### 3.1 Norms & Memes

On the one side, *memes spread for prescriptive-normative reasons*, and on the other side, *norms are themselves memes*, aimed at spreading around in a given group or population, and at being adopted and shared by its members.

Thus, Norms are META-MEMES: they are memes for regulating the conformity to and the *spreading* of behaviors, goals, ideas; they provide Models and prescribe who should or can do what.



**Figure. 2**

When conforming to a social or legal norm the agent is at the same time acknowledging the authority issuing the norm, and reinforcing both. In fact, the most respected is an authority the most respect-worthy it is. As we said, the more respected is a norm the more difficult it is to not recognize it as a norm and to deviate from it.

Particularly interesting in this perspective is the notion of 'authority' which is a specifically memetic notion. In fact 'authority' stands for the influential role, the leading impact of a memetic source on the other agents. Either it is the role of a Model among a group of 'followers' who imitate her/his behavior or absorb her/his ideas [note 10] because of its prestige or competence; or it is the role of a norm-issuing agent who, by norms and commands (prescriptions), shapes the behaviors of subject agents.

Norms are not only prescriptions, but also permissions and *prohibitions*. This means that while in some meme-adoption the individual can feel free, in others it has an obligation, and other meme-adoptions are simply forbidden and would imply a violation. Thus norms create real (mental) barriers to the diffusion of a given meme. Consider such example as the attitude towards contraceptives in catholic countries and its effects, or the moral values against the use of drugs acting as defensive mechanisms in some social groups, or the resistance to the penetration of the western 'civilization' in Islamic countries. *Moral barriers* (prohibitions) can be very effective.

#### **4. Identity/Membership: the Social Identity Mechanism**

Several authors (including Tomasello [36]) agree about the fact that social/cultural identity and identification with the others play a crucial role in cultural transmission. In Tomasello's model this is due to the crucial role that in true imitative learning is played by the *understanding* of the mind (intentionality) of the other, and this understanding seems due to the projection -by the

child- of his subjective experience, thanks to the belief/feeling "he is like me" (p. 14). This is an interesting view but it is rather limited.

On the one side, in fact it is possible to ascribe mental states (especially in adults) without identification, empathy or whatever projection of our subjective experience; just as powerful predictive models of special 'agents'.

On the other side -and this matters here- *identification* plays a much more important role as feeling, belief and desire of membership in a category and of acceptance in a group, that is as personal 'social identity': "I am, I feel I am an X"; "I am *like them*, one of them".

In both cases ("he is like me" and "I am like them") there is some form of identification, but they are not equivalent. The analogical movement is in fact reversed: in one case, the others are assimilated to me and some of my features are ascribed to them; in the other case, I'm assimilated to them, and their features are 'transferred' to me.

### **Social Identity Mechanism**

If I believe and/or want to be "one of them" (males, Italians, researchers, etc.), if this is part of "my identity", and

I believe that certain features are needed

- either to belong to that class or category X (*weak membership*),

or

- to belong to a given group via their 'recognition and acceptance' (*strong membership*) ;

or

I believe that certain manifested features (behaviors, attitudes, use of artifacts, etc.) are *symbols* of this membership (and status) and are used for recognizing a person as "one of them", [note 12] and

I want to send such a message and to be recognized as "one of them"

then

I want to ADOPT those features from/of that category or group in order to:

be an X;

be accepted as "one of them";

be recognized as "one of them"

feel I am one of them and recognize myself as "one of them".

In the fruit example the implied reasoning would be something like: "this is a sign of distinction of an elite, a sign of membership to a special group; I want to be like them (distinguished) or one of them (elite)".

Notice that the Adopter also believes that not anybody does so, but just a sub-group.

#### **4.1 Identity Differentiation or Hostility**

Since social identities obviously presuppose distinction, differentiation, and even hostility and conflicts, the belief and the goal of "being an X; being one of them" necessarily implies

*the goal of "not being a Y", "not being like them". [note 13]*

We call this goal 'Social Identity Hostility'. It plays a very important role (as the previous Identity assimilation) in cultural transmission. On the one hand,

this is a motive for *NOT Adopting* a given cultural trait, for not conforming and personally rejecting it;

it is a possible motive for *actively fighting against the diffusion* of this trait in the population (for my sons, friends) and against the expansion or existence of the other group members, ideas, behaviors (these are *cultural wars*).

Thus there can be a motivated resistance and even active opposition against the spreading of a trait. In other words we are far from a notion of 'contagion', and we see again the *selection process* as a *mental process*, and as deciding in favor or against of the trait adoption and determining the macro-spreading or not of a given behavior.

On the other hand,

this can be a motive for *not propagating* in some direction, *not revealing* our 'knowledge' or features, in order to protect our difference or our superiority (and power).

Finally,

this can be a motive for *innovation*: in order to re-establish a distinction from other classes or groups (for example higher class distinctive behaviors or symbols) new distinguishing features could be necessary.

As for the importance of the Distinction force (motives) in a dialectic relationship with the Conformity force, see Simmel's view on fashions [34; 3; 32].

## 4.2 Cooperation and conflict between the practical, normative, and identity-based adoption mechanisms

Those different motives and mechanisms for adoption, preservation, and transmission of memes can converge and cooperate with each other but they can also be in conflict. For example the normative mechanism and the practical one converge when it is both useful in practice and socially prescribed to adopt a given behavior. They are in conflict when the individual, on the one hand, is motivated to conform to the tradition, or to the expectations and approval of the others, while, on the other hand, the new solution looks more efficient for its practical purposes.

Also between the normative and the identity-based mechanisms for meme spreading there can be a conflict, but usually they cooperate. In fact, as we said, membership and social identity are strongly value and norm based, and to be an accepted member of a group one must and want to conform to the group values and norms.

Notice that the 'problem solving' mechanism is both a mechanism for *conservation* (until a better solution appears) and for *innovation*: "it is new but since it is better I will change and adopt it". By contrast, the normative mechanism or the conformist motive, seems only useful for conservation of what is already a norm in the group; it introduces something new only for a learning individual, not as innovation.

It is true that norms also emerge spontaneously and bottom-up, and individuals can create (and even intentionally propose) new norms. However, the motivation for creating a new norm is not the same as that for adopting a norm; while the motivation for finding a better solution is the same as that for adopting a better solution.

## 5. Perceived Diffusion and Actual Diffusion

As we saw, in all the previous cases the behavior adoption and its replication by the individual are influenced by its *perceived diffusion*, while actual diffusion -in turn- depends on individual agent adoption:

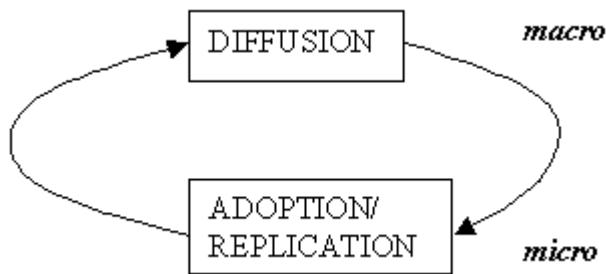


Figure. 3

As we said, the individual mind is not the *only* selecting environment; for example, diffusion in population plays a role by itself:

## **DIFFUSION ==> DIFFUSION**

For example the probability of survival of a given belief or behavior after a catastrophe destroying a population or a culture, or after an invasion, is greater when that meme is very diffused in that population.

However in the above mentioned mechanisms the 'diffusion' factor passes *through the individual mind*:

## **Actual DIFFUSION==>(perceived-diffusion/ interpretation/ adoption) ==> Actual DIFFUSION**

The fact that in all these mechanisms (especially the first two) adoption and replication are strongly influenced by *perceived diffusion* shows that the notion of 'social proof' [16] is a vague and broad category covering different processes, not a single and precise mechanism (as proposed for example by Hedstrom [27]). Moreover, the role of the perceived diffusion in diffusion process is a nice example of '**immersion**' in social dynamics: i.e. of the feedback of complex macro-level effects on the micro-level and in particular it is an example of the role of the mental representation (partial understanding) of the emerging macro-phenomena [18; 9].

## **6. Belief Adoption and its Constraints**

Those mechanisms (section 2, 3 and 4) mainly refer to behavioral 'imitation' (that in a cognitive agent - whose 'behavior' is goal-directed - means the imitation of goals, and the reasons and motives for adopting those goals). Important special constraints concern the adoption and spreading of *Beliefs*, that is opinions, evaluations, (folk)science, factual knowledge, etc. (***Know That* transmission**)

We cannot believe anything we observe or somebody communicates to us. We accept a given belief on the basis of our previous beliefs, of their evidence, supports and sources, and of others psychological factors. Analogously we decide to spread around our opinion or our knowledge for some reason and purpose. To study belief-propagation we have to study the principles of belief credibility and belief sharing. Let us sketch some crucial point of these cognitive mechanisms.

Our knowledge base is not a file where one can introduce new data or eliminate a file-card without altering the other data. Our beliefs are integrated, interconnected and mutually supported: to drop a belief or to add a new one entails checking its coherence with the rest and revise previous knowledge. The belief-belief coherence and support is quite a well studied problem in philosophy and AI (truth maintenance systems; belief revision and updating; argumentation) and in some cognitive agent architecture. There are in fact two schools in belief revision ([26; 24; 22]): the "foundations approach" stressing the importance of supports and justifications of beliefs, and the "coherence approach" modeling logical compatibility and coherence. However, we agree with Doyle [22] that there is no incompatibility between the two models, and that beliefs must be both relatively coherent and justified.

## **6.1 To accept and to reject; storing vs. believing**

As we wrote elsewhere [7] the meaning of 'revising' and of 'rejecting' a belief is quite obscure. One either believes something or she does not believe it. What does it mean to 'reject' a belief or an information [5] (which is not the same)? In our view in many approaches there is a dangerous confusion between memory and knowledge, between storing an information and accepting/believing it. In several models, the belief base is imagined as a memory store plus the coherence constraint. When a new belief arrives it is either coherent -and it is accepted (added)- or it is in conflict with other beliefs. In that case, after a given process, the new belief is either rejected (not stored) or the old knowledge is revised in order to be compatible with the new information. In this simplistic scheme, to 'accept' means both to store and to believe, while to 'reject' means not to believe and even not to store the 'information'. In order to avoid such a confusion and to clarify our interpretation of accepting and rejecting, we will assume *a strong independence between memory and knowledge* (or better beliefs). Quite obviously:

- a) one can remember something she does not believe (and remember that she does not believe it); and
- b) one can forget something she believes.

Thus, when we say that a given information is 'rejected' we mean that the subject *refused to believe it*, but that this information (and its disbelieving) is perhaps stored. When we say that an information has been 'accepted' we do not mean that it has just been stored: we mean 'believed'.

It is important to notice that we do not spread around *only* 'believed' information; not only because there are a lot of lies and simulation in human communication and social interaction, but because, for several reasons, we maintain and transmit a lot of non-believed information.

I will not discuss here the function of legends, fiction, play, etc. but, it is important to notice that information passed round as shared knowledge follows different paths and has different filters than information transmitted as fiction; and -of course- they satisfy different cultural functions.

## **6.2 The decision to believe**

In our view, there is a basic postulate for our decision to believe:  
*Believe only if you have reasons to believe*

However, a reliable source is by itself a reason for believing.

In this way the postulate is not too strong [note 14], and does not contradict -in our view- Harmann's claim [26] and some experimental results [25]. In fact, we don't imply that to believe something one has to logically demonstrate it in her/his knowledge base, or find it strongly 'plausible'. We just imply that sources give us reasons to believe -quite automatically- and that these reasons have to be stronger than the possible first hand 'implausibility' of the new piece of information.

## **Source reliability and belief credibility**

On the one side, believing in a source depends on its assumed reliability [note 15]. On the other side,

*The Credibility of a piece of knowledge (a candidate Belief) is a function of its sources*

The basic principles governing Credibility are the following ones :

*1a. If the source is reliable its information is credible and is believed; if it is not reliable its information is not credible and not believed. (In quantitative terms: the more reliable the source the more credible the information provided)*

*1b. The many the converging (independent) sources, the more credible the information provided*

*Any convergent source of knowledge 'confirms' the other* (in particular: S2 confirms S1, when S2 is a new source for a previous item whose source was S1).

'Confirmation' is a fundamental cognitive 'integration' among sources. It consists of the fact that:

- a) after the arrival of a 'confirming source' the item (the belief) is stabler, safer, more certain, and we subjectively are more sure and convinced about it;
- b) not only the item is more 'credible', but also the confirmed source is more credible, trusted (it is felt as more 'reliable').

Confirmation is a very important psychological phenomenon: when we control or check something we are just looking for confirmatory sources; proactive behavior, expectations, and goals imply some 'confirmation' mechanism; there is a very well studied 'confirmatory bias' in our cognition which is not only due to the cost of revising our knowledge (economic motivation) but also to our need for control and to our need to trust our knowledge and our ability to make predictions [2] (apart from self-deception and defense mechanisms).

The 'independence' of the source is clearly very important for our rationality and for resisting social influence. However, from a psychological point of view we have to admit that people are leaning to accept as confirming and additional evidence also the mere repetition of the same input, and that we do not care so much of controlling the real independence of our social sources (for ex. gossip, or newspapers). This is not so irrational sometimes.

The reliability of the source depends on many different aspects.

As for the social or communication sources, many authors identify two dimensions: *competence* and *trustworthiness*. The first is related to the fact that the content of the information is pertaining to a domain the source can be really expert and informed about.

## **Other criteria: *importance* and *plausibility***

Of course, number and credibility of sources are not the only criteria for accepting/rejecting beliefs. It is not the aim of this paper to examine the criteria for belief acceptance and revision, however we need to clarify at least partially which are the other factors that contrast or contribute with the sources to the acceptance/rejection of a belief.

### **Importance**

In our view in the literature about belief revision there is a confusion between two properties of integrated beliefs [5] : their *importance* and their *credibility*. They are two distinct dimensions and notions:

*a belief could be very important but not very credible; or very credible but absolutely marginal.*

By 'important' I mean that it will explain a lot; it will be very *useful for understanding and integrating other information*. 'Credible' means that I have a lot of evidence, sources, supports to believe it. Clearly enough the two aspects are distinct. An integrated belief in a belief network is in fact both supported and supporting: I call *credibility* how much it is supported by external or by internal sources ('plausibility'), and *importance* how much it supports: its explanatory power.

There is a contribution of both 'credibility' and 'importance' in the decision to accept or to maintain a belief. To decide to believe we do not consider only credibility (see later). The fact that a belief is highly 'important' (explains and supports a lot of other beliefs) can be a strong reason for accepting it, and a strong reason for not changing/abandoning it (resistance). In fact abandoning an important belief entails a lot of expensive revisions in our mental map.

### **Plausibility**

To be believed, something should be 'plausible'. Often we resist or reject a new item just because it is not 'plausible'. What is the basis of this kind of evaluation? Why is it a basis for rejection?

Plausibility is *the credibility value assigned to the coming item from inside*. It is evaluated just on the basis of previous knowledge: the same knowledge it has to be integrated with.

Thus, we have two credibility values, one based on external sources, the other based on the niche that has to accept the new item. Metaphorically, one is the value provided by the offering agent that 'gives' the item, the other is the value attributed by the accepting (or refusing) agent. When there is a conflict, and a difficulty to accept the new belief from outside, the conflict is between its 'credibility' and its 'plausibility'. To accept a new belief the plausibility i.e.; the internal source should 'confirm' the external source, and converge with credibility; or, at least, credibility should be stronger than implausibility.

More precisely, **plausibility** is our attempt to derive, infer the new item from our previous knowledge: the more predicted or expected, the more 'plausible'; in other words we are searching an internal source confirming the external one, in order to better 'assimilate' the new information;

at least we would like not to have internal reasons for not believing i.e. implausibility. Implausibility is the result of our attempt to derive/infer the opposite of that item. The first attempt is due to our need not only for knowing 'that' but for knowing 'why' (Aristotle), and for integrating knowledge on such a basis (explanations, reasons,). The second attempt is due to *the necessity to verify if there are conflicts* between the new item and the consolidated knowledge. This would be either a reason for rejecting the item or a reason for revising beliefs and for rejecting (dropping) some old items.

Of course this is a 'normative' perspective; on the psychological side we know that there is experimental evidence that people tend to accept new beliefs without such deep (and long and expensive) controls. Gilbert [25] for ex. has shown that we quite automatically believe all that we comprehend, and that the rejection of ideas comes later as part of a more effortful process. This result should be taken with caution, since Gilbert's notion of 'comprehend' is quite broad. Clearly enough we believe quite automatically, by default [6] provided that there are no overt contradictions, immediate implausibility, or previous reasons to suspect/doubt about the source (see note 12).

External 'credibility' is just one component of the 'acceptability' or 'believability' of an item (its properties that will determine whether it will be believed or not). 'Plausibility' is another component. But even credibility, plausibility and importance are not enough.

There are also other dimensions that interfere with belief acceptance, like 'relevance' and 'likeability' [6; 7]. By 'relevance' we mean how useful, important is a given belief for our interests and goals. By 'likeability' we mean a special aspect of relevance, i.e. the fact that a belief frustrates or satisfies a goal, is pleasant or unpleasant for us. Our belief adoption is not only a rational process; it is also influenced by affective responses - which we cannot address here - such as wishful thinking, defensive mechanisms and self-deception.

### **6.3 Infra-psychic memetic competition**

When and why would two memes be 'in competition' [19; 4]? First of all for merely cognitive-psychological causes. Let us give a couple of examples of infra-psychic memetic competition.

#### *Epistemic unacceptability*

As we just saw, meme/belief/idea/plan x can be contradictory with what I already believe. X is 'implausible' to me, i.e. 'unbelievable'. In order to adopt it I should revise a lot of knowledge I have. The new meme (that is attempting to be 'hosted' by me) is in competition with another (previous) one, just for cognitive reasons: I cannot accept both at the same time.

#### *Deontic or Dynamic unacceptability*

A completely different case - closer to the previous principle about behaviours ('moral barriers' in section 3.1) but applicable also to mere beliefs- is the following one: X could be believable (acceptable for cognitive coherence and revision reasons) but is unacceptable for moral or

religious reasons (incompatible with my values and interiorised norms), or very disagreeable and painful to me, given my desires and interests.

#### 6.4 Why Sharing Beliefs

What is the use and motivation of sharing beliefs? Is sharing just an effect or is it sometimes an end?

Shared knowledge plays several roles.

On the one hand, shared knowledge is a guaranty of truth. Several independent sources with their own evidence make us more certain about the reality of our beliefs (this is a principle valid even in science). Moreover, the many people -with their coherence and plausibility checking and their control - have accepted a given belief the more it has been cognitively tested and can be believed. As we saw, 'social proof' is an important factor for believing, but it is also a function and a motive for sharing information: testing it, exchanging it, creating a common patrimony.

On the other hand, sharing knowledge is useful for several independent social functions. Coordination, cooperation and organization (working in a common environment for a joint plan) require common knowledge about the task, domain, plan, and context. Group membership and identity require common knowledge: the myth, or the history, or the special competence of the group, etc. Any social interaction and communication requires some common knowledge of rules, scripts [32], roles, etc.

It is important to understand that those two types of function (validity of beliefs vs. social coordination) are rather independent of each other. Shared beliefs can be completely false but work quite well as culture common ground and group 'glue'. This function can be largely independent of the truth of what is mutually believed [21].

In conclusion, it seems clear that for a belief (which is the mental form of many memes) to be accepted, preserved, and transmitted by an agent, its impact within the belief structure of that mind is crucial. More precisely its credibility, based on the reliability and the number of sources, and on the internal plausibility, its epistemic importance (explanatory impact and integration), its likeability... and even its moral appropriateness, but also the mere memorization [23] and retrieval, play a crucial role. Different motives and plans push individual to share beliefs by passing them or adopting them just in order to share them.

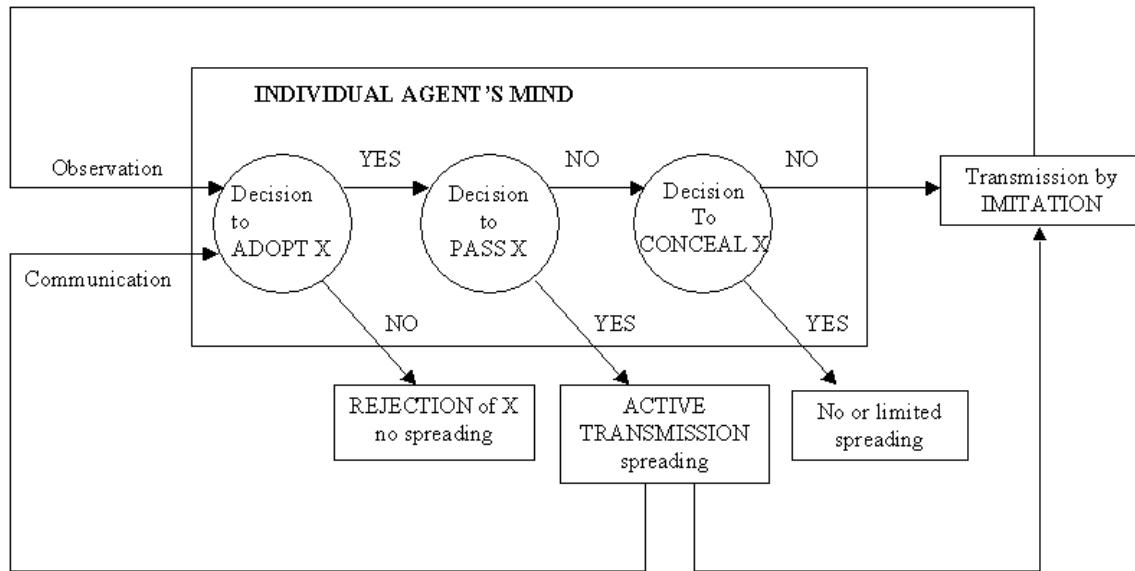
### 7. A more general and sophisticated model: Two crucial 'decisions' and their emergent effects

Let us now look at meme adoption and transmission more in depth, by expanding our simple model (Figure 1) in terms of the Agent's *decisions* and relatively not only to observed behaviors but to meme transmission via actual *communication*. We have in fact to deal not only with observation but also with explicit messages [12] aimed at inducing the addressee to believe or to do something. In such more analytic view of the mental processes involved in meme transmission, we observe that:

the so called 'contagion' can be the result of *decision* processes (to believe or not; to adopt or not), and

the so called spreading can be the result of another possible *decision*: to pass or not to pass such a meme to others.

The agent is very far from being *a passive 'vehicle' of memes*; it can actively decide about receiving them and passing them.



**Figure. 4**

As we can see in this simplified schema of the agent's decision process -relative to candidate memes coming either from the observation of a possible Model or from the communication by a source- the agent (on the basis of its interpretation and motives, see Figure 1) has to decide

- i) whether to adopt or not such a belief, behavior, or method/tool; in case of rejection there will be no spreading of such a meme - at least through this agent; in case of adoption the agent faces another decision (based on other kinds of reasons):
- ii) whether to pass this meme to other agents or not; if the agent intends to transmit this information it will either explicitly communicate (information, prescription, advice, instruction) or make the others observe its own behavior ('behavioral communication' - [12]); while, the decision of not actively transmitting its knowledge means:
- iii) either to actively try to conceal and hid it; or
- iv) not to care about its diffusion and just let the others observe or not and imitate or not it (the same hold for a *lack* of decision of transmitting or not).

It is important to remark -as we did at the end of section 1.3 - that the spreading of a behavior or

of an idea is not necessarily intentional (decided by the agent M); there is some room in this model for forms of 'imitation' (meme adoption) not due to M's decision to spread it around. But there is also room for intentional spreading or suppression [note 16].

## 7.1 Macro-layer effects

If we looked at the spreading of a given behavior (or belief) at the macro (population) layer, we would observe different 'contagion' or propagation properties of those socio-cognitive micro-mechanisms. The speed and the direction of propagation would be different for each mechanism; and also relative to specific memes (in fact, different memes may use different mechanisms and meet different cognitive and motivational problems because of the specific beliefs, norms and motives of the candidate vehicle). Some area of the population could be quite 'viscous' in propagating a given meme, while another could be completely resistant to any invasion. For example, an active role of the individuals in the propaganda of a given meme and in influencing the others predicts (*ceteris paribus*) a higher spreading speed than the mere casual observation of the behavior where the Model is passive and indifferent and adoption is a mere individual choice. The same holds for prescribed behaviors and values diffused either through the decentralized and diffuse social control or through enforcing authorities (or both). Analogously, social monitoring and moral or legal prohibition of given behaviors will make much more *stable* (difficult to change) the customs of a given group and more *uniform* (less variance within the population, and more intolerant culture).

*Different socio-cognitive micro-mechanisms make different predictions as for their emergent effects in meme propagation*

Emergent memetic macro-layer properties are the output of the specific socio-cognitive processing of the memes at the micro-layer.

## 8. Concluding remarks

Knowledge (either practical, deontic or factual) is not acquired and transmitted by mere 'contagion'. It is shared on the basis of adoption and passing, or not shared on the basis of rejection or not passing. Two crucial *decisions*.

It is important to model the various cognitive mechanisms and motives for *adopting* a trait (like the ones examined in section 2., 3. and 4.) or a belief (section 6.). However, not less important is modeling evaluations and motives supporting the decision to share or not a given piece of knowledge. It will also be quite important, for the theory of cultural construction and diffusion, to analyze when and why sharing knowledge or diffusing a behavior is a conscious goal of the agent or is just an unintended functional effect of its behavior [8; 11], and which are the different effects of intentional vs. merely functional diffusion. I just gave some example of those mechanisms, some preliminary model of them and some hint about their different conditions and effects, and different predictions that they allow at the macro-layer. Those models should be much better specified, other mechanisms should be modeled, and in particular their combined and massive effects should be experimentally studied using Multi-Agent based simulation approaches [31]. I just hope to have insinuated the suspicion that to adequately study the social

dynamics and evolution of memes we need somewhat complex 'cognitive' agents and the modeling of meme mental processing [17; 10].

Minds are not the *only* selection environment: the fire of the Alexandria library selected out billions of memes -ideas- that they intended to transmit to us. However, minds are the most relevant selective environment for memes, and for understanding cultural evolution it is necessary to identify (and simulate) the cognitive principles of the success or selection of memes *within minds*.

Memetics can only be cognitive, otherwise it is contradictory and non explanatory [note 17], but - of course - it should not be only cognitive.

## Notes

1. On the notion of 'meme' and on the current debate about its nature and spreading mechanisms, see for example, [19; 4; 1; 20; 28; 29; 37]. For a different and quite interesting approach -that I will not explicitly discuss here- see Sperber [35].
2. Consider for example the spontaneous and unconscious mimic of phonemic, prosodic, and also lexical and grammatical features that we experience while visiting or living in a new linguistic environment, in particular when we are in contact with a dialect close to our own dialect. This is usually a non deliberated but merely automatic mimetic activity, and for sure it is a very relevant mechanism for the spreading and uniformity of linguistic rules and behaviors. Notice that it is based on some cybernetic purposive device (we adapt to a model and adjust on the basis of a match or mismatch) or at least on some 'anticipatory classifier' driven and reinforced by an expectation. For sure we do not evaluate any technical advantage of such (subjective) 'innovation', nor we do so because we have the explicit intention "to be like them; one of them"; this goal is completely implicit in some rule-based system or procedure.
3. For example to its "causal attribution" of the event; for example to and within the activated schemata, frames, or scripts [33].
4. This distinction -very important from a cognitive point of view- is also in Hedstrom's taxonomy [27]: do I repeat/imitate for imitating or for other purposes?
5. For a nice example of different interpretations of the same behavior -with different consequences on spreading- see also [17].
6. To be read as "A Believes that M uses x for the goal p".
7. To be read as "A Believes that M has the goal p" and "A Believes that x is good/useful for p".
8. To be read as "A believes that the new observed/known solution is better than any other solution he knows, for p". Notice that the individual also prefers not searching for a better solution; he believes that either he is not able or that an additional search is too costly. Consider

also that, when one of the motives is curiosity or the novelty is a value, a given solution can be better just because it is new -and until is new.

9. In this work, arguments and distinctions are very good, but the very definition of “rational imitation” is rather unsatisfactory.

10. They follow the authority for one of these reasons: her/his ideas are the true ones, the successful ones, the intelligent ones (competence, prestige in technical (section 2.) or in belief (section 6.) innovation and transmission); s/he is a distinguished person and being like her/him means being accepted by the group or being a distinguished person (section 4.).

11. It is clear that this overlaps with the normative mechanisms, because these are “prescriptions” by group X to belong to X. But the two mechanisms do not completely overlap, since there might be forms of normative adoption which do not concern identity and membership, and forms of adoption for identity recognition or membership about non-normative features.

12. Another very nice example of this goal of “being like them; being recognized as one of them” (intentional imitation), but for completely different reasons and motives, is camouflage. Consider a spy who does not like to be recognized and identified as a foreigner. He will possibly dress and behave like the local population in order not to be noticed. Of course here we have a faked assimilation and adoption of those traits, not a real memetic or cultural propagation. Notice that this kind of imitation-*mimesis* can occur for defensive reasons (ex. in shy and shameful people) or for aggressive reasons (ex. in the spy), like in animals.

13. This crucial goal can also have an independent (and even individual) origin. If I do not want to belong to a given class or group I’m supposed or expected to belong to, I have such a goal not as a consequence of being in another incompatible class or group, but perhaps as a premise of being part or looking for a different identity.

14. Otherwise it would be better to assume as a principle: *Believe something if you do not have reasons for rejecting it.*

15. We cannot prevent ourselves from believing in a source, unless we suspect that there is "something wrong" in it. [7]. In order to reject the information of a source (even of a social one) we must believe that there is "something wrong" in that source.

16. Consider that the notion of ‘decision’ can be interpreted in more or less ‘deliberated’ and conscious way (section 1.1).

17. Tomasello’s book has a rather Vigotskian title: “The cultural origin of human cognition”. I like this perspective; however, the other way around is not less important. I mean; the cognitive origin of human culture!

## Acknowledgments

I would like to thank all people present at the Pontignano workshop for a nice discussion on such topics; a special thank to Maria Miceli and Rosaria Conte for their precious comments. This work has been developed within the *ALFEBIITE* European Project: *A Logical Framework For Ethical Behaviour Between Infohabitants In The Information Trading Economy Of The Universal Information Ecosystem.* - IST- 1999-10298.

## References

- [1] Aunger R. (ed) (2000). *Darwinizing Culture: The status of Memetics as a science.* Cambridge, The Open Press.
- [2] Bandura, A. (1982). Self-efficacy mechanism in human agency. *American Psychologist*, **37**, 122-47.
- [3] Benvenuto, S. (2000) Fashion: Georg Simmel, *Journal of Artificial Societies and Social Simulation* **3**(2), <<http://www.soc.surrey.ac.uk/JASSS/3/2/forum/2.html>>
- [4] Blackmore, S. (1999) *The Meme Machine.* Oxford: OUP.
- [5] Cantwell, J. (1996). Resolving Conflicting Information. Department of Philosophy, Uppsala University, TR. June 20.
- [6] Castelfranchi, C., (1995). Guarantees for autonomy in cognitive agent architecture. In M. Woolridge and N. Jennings (eds.), *Intelligent Agents: Theories, Architectures and Languages.* LNAI, **890**, 56-70, Springer Verlag.
- [7] Castelfranchi, C. (1997) Representation and integration of multiple knowledge sources: issues and questions. In Cantoni, Di Gesu', Setti e Tegolo (eds) *Human & Machine Perception: Information Fusion.* Plenum Press.
- [8] Castelfranchi, C. (1998). Through the minds of the agents, *Journal of Artificial Societies and Social Simulation*, **1**(1), 1998, <<http://www.soc.surrey.ac.uk/JASSS/1/1/contents.html>>
- [9] Castelfranchi, C. (1998). Simulating with Cognitive Agents: The Importance of Cognitive Emergence. In J. Sichman, R. Conte, N. Gilbert (eds.) *Multi-Agent Systems and Agent-Based Simulation.* Springer, Berlin LNAI 1534, 1998, 26-44
- [10] Castelfranchi, C. (2000). Through the Agents' Minds: Cognitive Mediators of Social Action. *Mind and Society.* I.1. 109-140.
- [11] Castelfranchi, C. (2001). The Theory of Social Functions: Challenges for Computational Social Science and Multi-Agent Learning. *Cognitive Systems*, 2001 - in press. Invited talk at SimSoc'97, Cortona, Italy

- [12] Castelfranchi, C. (in preparation) From Implicit Communication to Tacit Agreements. Invited talk for the "Seminar series on Perspectives On Agents: Theories And Technologies". University of Trento and IRST, March 16, 2000. <http://www.cs.unitn.it/~pgiorgio/agents/>
- [13] Cavalli Sforza, L.L. & Feldman, M. (1981). *Cultural Transmission and Evolution. A Quantitative Approach*. Princeton, N.J: Princeton Univ. Press.
- [14] Cecconi, F. & Parisi, D. (1998) Individual versus social survival strategies. *Journal of Artificial Societies and Social Simulation*, **1**(2), <http://www.soc.surrey.ac.uk/JASSS/1/2/1.html>
- [15] Chattoe, E. (1998). Just How (Un)realistic are Evolutionary Algorithms as Representations of Social Processes? *Journal of Artificial Societies and Social Simulation*, **1**(3) <http://www.soc.surrey.ac.uk/JASSS/1/2/3.html>
- [16] Cialdini, R. (1993). *Influence. The Psychology of persuasion*. Q.W. Morrow Co., N.Y.
- [17] Conte, R. (2000) *Memes Through (Social) Minds*. In R. Aunger (ed) *Darwinizing Culture: The status of Memetics as a science*. Cambridge, The Open Press, 2000.
- [18] Conte, R. e Castelfranchi, C. (1995). *Cognitive and social action*. London, UCL Press.
- [19] Dawkins, R. (1976) *The Selfish Gene*. Oxford Univ. Press.
- [20] de Jong, M. (1999). Survival of the institutionally fittest concepts. *Journal of Memetics - Evolutionary Models of Information Transmission*, **3**. [http://jom-emit.cfpm.org/1999/vol3/de\\_jong\\_m.html](http://jom-emit.cfpm.org/1999/vol3/de_jong_m.html)
- [21] Doran, J. (1998) 'Simulating Collective Misbelief', *Journal of Artificial Societies and Social Simulation*, **1**(1). <http://www.soc.surrey.ac.uk/JASSS/1/1/3.html>
- [22] Doyle, J., 1992. Reason maintenance and belief revision: Foundations versus coherence theories. In P. Gardenfors (ed.) *Belief Revision*. Cambridge University Press, Cambridge, UK, pp. 29-51.
- [23] Gabora L. (1998) Autocatalytic Closure in a Cognitive System:. *Psycoloquy*: 9(67) Origin Culture (1) <http://www.cogsci.soton.ac.uk/cgi/psyc/newpsy?9.67>
- [24] Gardenfors, P., 1992. *Belief revision: An introduction*. In P. Gardenfors (ed.) *Belief Revision*. Cambridge University Press, Cambridge, UK, pp. 1-28.
- [25] Gilbert, D.T. 1991. How mental systems believe. *American Psychologist*, **46**, 107-19.
- [26] Harman, G., 1986. *Change in View: Principles of Reasoning*. MIT Press, Cambridge, MA.

- [27] Hedstrom, P. (1998) Rational Imitation. In P.Hedstrom & R. Swedberg (Eds.) *Social Mechanisms : An Analytical Approach to Social Theory* (Studies in Rationality and Social Change), Cambridge University Press. Mass.
- [28] Hull, D.L. (1999) Taking Memetics Seriously: Memetics Will Be What We Make It. Paper presented at the Conference "Can Memes Account for Culture?", Cambridge, UK, 4-5 June, 1999.
- [29] Laland, K.N. & Odling-Smee, J. (1999) The Evolution of the Meme. Paper presented at the Conference "Can Memes Account for Culture?", Cambridge, UK, 4-5 June, 1999.
- [30] Miceli, M. & Castelfranchi, C. (1989). A Cognitive Approach to Values, *Journal for the Theory of Social Behaviour*, **19**(2), 169- 194.
- [31] Moss, S. & Davidsson P. (eds) (2000). *Multi Agent Based Simulation*. Springer.
- [32] Pedone, R. Conte, R. (2000). Imitation and Avoidance in Social Hierarchies. Simulating the Simmel Effect. In S. Moss and P. Davidsson (eds) *Multi Agent Based Simulation*. Springer.
- [33] Schank, R. and Abelson, R. 1977. *Scripts, Plans, Goals and Understanding*. Hillsdale, John Wiley.
- [34] Simmel, G. (1904-1957) "Fashion", *International Quarterly*, **10** (1), October 1904, pp. 130-155, reprinted in *American Journal of Sociology*, **62**(6), May 1957, pp. 541-558.
- [35] Sperber, D., (1991) Epidemiology of Beliefs. *Revista De Occidente*, **125**, 5-& Oct.
- [36] Tomasello, M. (1999). *The Cultural Origins of Human Cognition*, Harvard Univ. Press.
- [37] Wilkins, J. S., (1998). What's in a Meme? Reflections from the perspective of the history and philosophy of evolutionary biology. *Journal of Memetics - Evolutionary Models of Information Transmission*, **2**. [http://jom-emit.cfpm.org/1998/vol2/wilkins\\_js.html](http://jom-emit.cfpm.org/1998/vol2/wilkins_js.html)

## Appendix

### THE CULTURAL ORIGINS OF HUMAN COGNITION

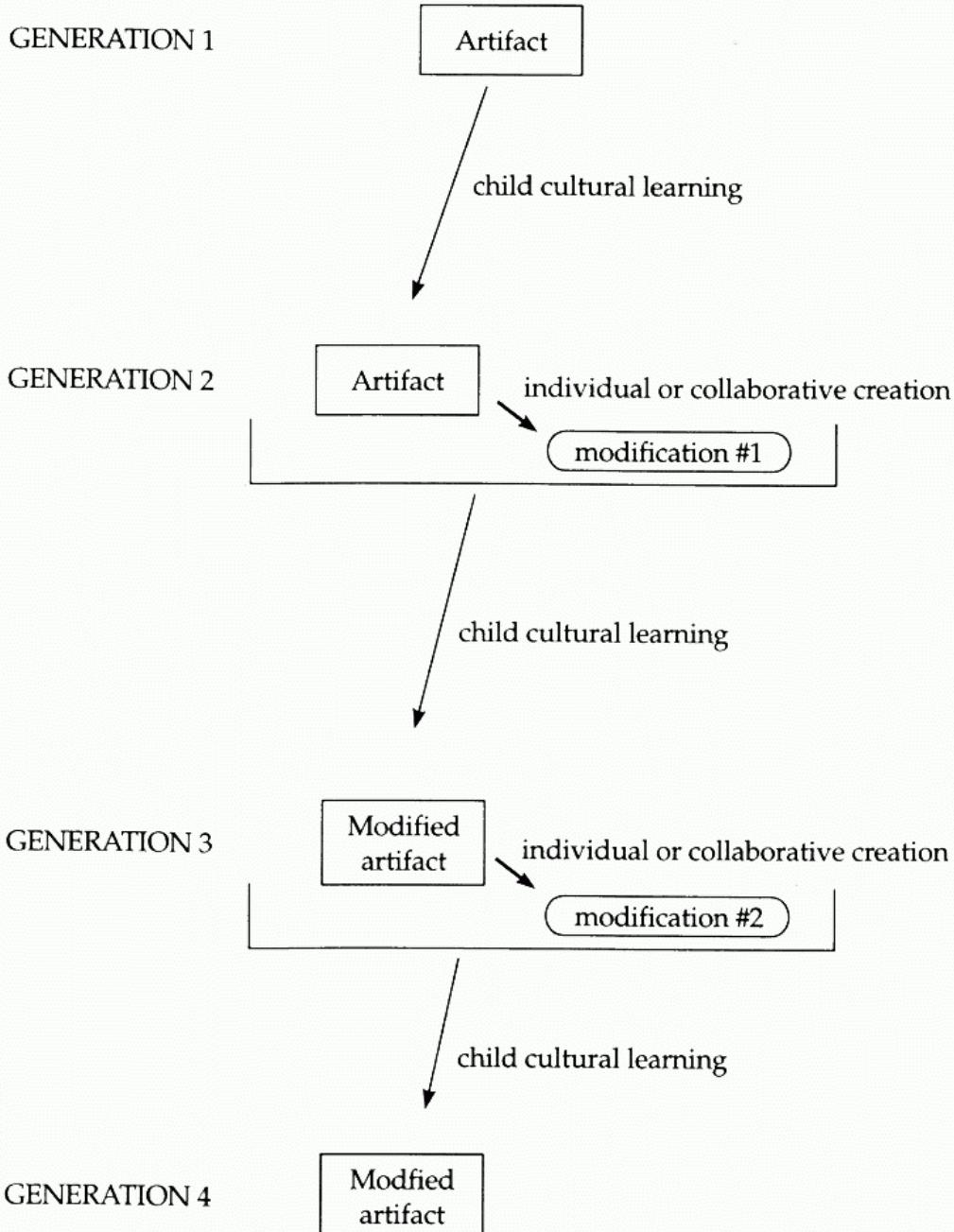


Figure 2.2 A simplified depiction of the ratchet effect working to produce an artifact with accumulating modifications.

# **JUST HOW (UN) REALISTIC ARE EVOLUTIONARY ALGORITHMS AS REPRESENTATIONS OF SOCIAL PROCESSES?**

**By**  
**Edmund Chattoe (1998)**

*Journal of Artificial Societies and Social Simulation* vol. 1, no. 3,  
<<http://www.soc.surrey.ac.uk/JASSS/1/3/2.html>>

To cite articles published in the *Journal of Artificial Societies and Social Simulation*, please reference the above information and include paragraph numbers if necessary

Received: 15-Feb-98    Accepted: 13-Jun-98    Published: 30-Jun-98

## **Abstract**

This paper attempts to illustrate the importance of a coherent behavioral interpretation in applying evolutionary algorithms like Genetic Algorithms and Genetic Programming to the modeling of social processes. It summarizes and draws out the implications of the Neo-Darwinian Synthesis for processes of social evolution and then discusses the extent to which evolutionary algorithms capture the aspects of biological evolution which are relevant to social processes. The paper uses several recent papers in the field as case studies, discussing more and less successful uses of evolutionary algorithms in social science. The key aspects of evolution discussed in the paper are that it is dependent on relative rather than absolute fitness, it does not require global knowledge or a system level teleology, it avoids the credit assignment problem, it does not exclude Lamarckian inheritance and it is both progressive and open ended.

## **Keywords:**

evolutionary algorithms, genetic programming, social evolution, selectionist paradigm

## **Introduction**

### **1.1**

There is now a significant body of work in social science (though predominantly in economics) which can legitimately be referred to as evolutionary. This paper focuses on simulations which use various kinds of evolutionary algorithms<sup>[1]</sup> to represent social and particularly economic processes. There are now enough simulations of this kind that an "existence proof" has clearly been provided. It may now be appropriate to step back a little and try to provide a more general rationale for models of this kind. That is the object of this paper. An attempt to find a coherent interpretation for evolutionary modeling serves three purposes. The first is self evident. It is easier to pursue a research program that has been set in some sort of order, albeit provisional and descriptive. The second is that evolutionary models have so far been carelessly criticized or ignored by orthodox economists and other social scientists.<sup>[2]</sup> Many of the traditional criticisms rest on misinterpretations or take advantage of ambiguities in the interpretations used by researchers in the field. A coherent interpretation may also serve to improve the quality of subsequent debate. Finally, the process of building an interpretation reveals weak spots in the

current simulations and may suggest both immediate solutions and areas for further long term research.

## 1.2

The second section of the paper describes the evolutionary process in biology as it is typically represented by social scientists. We also draw out some of the implications of evolution that are relevant to what follows. The third section summarizes how these evolutionary ideas have been implemented in evolutionary algorithms of various kinds. The fourth section summarizes the "standard" interpretation of evolutionary simulations which has emerged and the way in which it relates to concepts like procedural rationality in economics. The fifth section focuses on a number of specific simulations, including our own work in progress, to refine this interpretation, pointing out its weaknesses and suggesting areas for further development. The sixth section concludes.

## 1.3

A synthetic paper of this kind is inevitably both derivative and speculative. I have provided a quantity of references and footnotes that would normally be excessive in the hope of pre-empting some of the "traditional" criticisms. My goal is not to lay claim to very many of the basic ideas presented in this paper, only to putting them into a logical order and providing an extended critique which I hope will be provocative but helpful. (In addition to criticisms of prevailing approaches, I have also provided descriptions of alternatives wherever possible.) I am also trying to address two audiences and provide something of interest to both. The first audience are those who have not given too much thought to the possibilities of evolutionary models and the second are those who actually work in the field. Obviously keeping both groups interested is a rather delicate balancing act. This raises a final tension in my objectives, between current practice and future possibility. Although the paper focuses on examples from economics, the importance of a correct interpretation of evolution applies to modeling involving evolutionary algorithms in *all* the social sciences, even though in some it is barely used at present. Runciman (1998) discusses the profound lack of enthusiasm for evolutionary models in sociology and this attitude has proved so pervasive that no simulations (at least analogous to those described here) yet appear to exist in the discipline. Rather than providing a review (an exhaustive but essentially reactive description of existing work), this paper attempts to develop a thesis with appropriate illustrations. The thesis is that the coherent interpretation of social evolution (what Runciman calls the "selectionist paradigm") is both possible and desirable. For this purpose, the most appropriate illustrations are the best developed models which currently exist. However, it is hoped that in future models based on a coherent interpretation will increasingly be developed and applied in new areas. This paper is intended as a small contribution to that larger project.

# **Evolutionary Biology and Evolutionary Models**

## 2.1

This section defines social evolutionary processes, summarises the mechanisms of evolutionary biology as they are typically used by social scientists and draws out their implications.

## **The Meaning of "Evolutionary" In Social Science**

### **2.2**

The term "evolutionary" has been much misused and reviled in social science. Four major interpretations can be distinguished. Firstly, there are Social Darwinist theories (Hofstadter 1955, Jones 1980), which are also associated with the ideas of eugenics. Such theories were often mere political rhetoric, and were so imprecise they could easily be used to support or oppose the same policy. They now have little more than historical interest since they made no attempt to justify the analogies they proposed between individuals, states or firms and individuals or species in biology.<sup>[3]</sup> In some cases, they appear to have been based on misunderstandings of what Darwin and subsequent evolutionary biologists actually said. Secondly, there are Sociobiological theories (Wilson 1975, Wilson 1978). These attempt to explain certain behaviors, traditionally regarded as social, in genetic terms. Theories of this kind are not ignored here because they are of poor quality, but because it is generally assumed that there is a realm of distinctively social behavior. Sociobiology may alter the size of this realm and force us to reconsider the explanations of certain behaviors at the margin, but it does not seem likely that it will swallow up the core of social science. Thirdly, there are theories which simply deal with change, including change in the very long term. Institutional economics has often been described as evolutionary in this sense (Hodgson *et al.* 1994). Our reasons for disregarding theories of this kind are more tentative. As with Social Darwinism, *some* of these theories use evolution in a sense that is more rhetorical than substantial (Petr 1982). They stress the (subjectively) functional nature of certain institutions and practices without saying much about how functionality might arise or how functional and non functional practices might be distinguished within a scientific framework. This is not to say that institutional economics has not made an extremely valuable contribution to unorthodox economics, nor that some models originating in the field are not evolutionary in the sense defined in this paper. The main reason for excluding most models of this kind is semantic and pragmatic. Institutional, historical or dynamic economics are perfectly appropriate terms to describe most of these models. By contrast, the models discussed in this paper would be almost impossible to describe in any other way. The final sense of evolutionary, and the one we shall use in this paper, refers to models or simulations which propose a detailed and coherent analogy between a biological evolutionary process and a social process.<sup>[4]</sup> The biological evolutionary process chosen is usually that referred to as the Neo-Darwinian Synthesis (Darwin 1968, Fisher 1930, Ridley 1985, Edey and Johanson 1989) which combines Darwin's theory of natural selection with Mendelian genetics.

## **The Neo-Darwinian Synthesis and Its Implications**

### **2.3**

For the purposes of this paper, this evolutionary theory can be summarized quite briefly. The objects of the theory are individual living creatures, which each consist of a *genotype* (a certain arrangement of genes) and a *phenotype* (the rest of the physical body which was "constructed" by that genotype during the process of gestation.) This construction process involves the genes "encoding" the synthesis of the enzymes and other molecules required to orchestrate the growth

of an individual from a single cell to the point of birth. The environment, which consists of other individuals and physical structures, *selectively retains* individuals depending on the "performance" of their phenotypes.<sup>[5]</sup> This performance involves persisting for long enough to reproduce, thus perpetuating the successful genotype. Two other important mechanisms take place during the reproductive process.<sup>[6]</sup> In sexual reproduction, there is *genetic mixing*, the genes of the offspring contains some of the genes of *both* parents. In addition, the normal duplication of genes results in *mutation* through "copying errors", though this mutation can also result from external processes such as radiation. The functions of these two processes can be seen as the *generation* of novel combinations of genetic material (that may persist better in the environment than those which currently exist) and the *maintenance* of novelty. (We can envisage a situation in which, with a stable environment, genetic convergence will result. Mutation ensures that there is always a minimum of genetic diversity so that a subsequent change in the environment does not immediately result in the extinction of entire populations.)

## 2.4

For the discussion that follows, this process has several important and interdependent implications:

Evolutionary fitness is based on relative and not absolute success. There is no requirement for an "invisible hand" or any global knowledge to exist within the system. Nor is there any requirement for a "minimum standard" of performance before the evolutionary process can begin. Each individual interacts only with other individuals (and physical processes) and persists or ceases accordingly.<sup>[7]</sup>

Evolution avoids the credit assignment problem. Individuals persist or cease on the basis of all their strengths and weaknesses taken together. It is not necessary for them to be capable of making attributions about the particular causes of success or failure.<sup>[8]</sup>

Evolution neither requires nor implies an external teleology. Although individuals may have goals and drives, themselves evolved as responses to the environment, the system as a whole does not have a particular principle for selective retention. (This follows from the local nature of the evolutionary process. There is nothing for such a principle to inhere in.) In not requiring a teleology, evolution avoids an infinite regress of explanation and is self-consistent. However little we know about the actual structure of the first organisms, there is no point at which we are theoretically obliged to postulate a conceptual discontinuity in the evolutionary process.

Evolution is an open ended process in two senses. Because the role of genes is to encode chemical processes, rather than directly "construct" a phenotype by a fixed mechanism, a completely new sequence of genes can result in a completely new enzyme with previously unknown effects. Similarly, at a higher level of description, the phenotypic abilities that individuals have are not fixed: a stick can be used as a digging tool or a club and neither use is strictly determined by the structure of the arm or hand.

Evolution is progressive. This is not meant in a teleological sense, but it follows from open endedness and relative success. Initially, organisms would simply have come into existence and

ceased again. If one was able to make copies of itself (Dawkins 1989), then that would be the only one to persist beyond the life span of a single individual. Once persistence was ensured, with all persisting organisms being copies of the original, then properties such as diversity might have developed additional persistence value. Once diversity existed, then genetic mixing could have persistence value and so on.<sup>[9]</sup>

Evolution does not exclude Lamarckian inheritance. It appears the Lamarckian inheritance - inter-generational genotypic transmission of phenotypic characteristics acquired during a single lifetime<sup>[10]</sup> - has not evolved in biological systems except for some fiercely contested cases among the simplest organisms (McDougall 1927). It is also fairly easy to see why this might be. There is simply no mechanism by which additional arm muscles might be "re-encoded" back into the genes, because they are not directly encoded there in the first place.<sup>[11]</sup> All that is encoded is the synthesis of enzymes and other chemicals that will, taken together, result in a particular set of muscles and other bodily features but *only through a dynamic process of construction*. The completed body does not retain any representation of this construction process, so it cannot be "reverse engineered". However, this is a contingent fact about genes. There is nothing conceptually incoherent about Lamarckian inheritance as part of an evolutionary framework. Indeed the existence of cultural transmission suggests that if the material of the "genotype" is human language, rather than genes, Lamarckian inheritance has indeed evolved and may now dominate genetic inheritance (Moravec 1989).

## 2.5

It is clear that these implications have considerable appeal for unorthodox economics and address serious problems with the orthodox view. The absence of system level teleology and global knowledge is appropriate for emergent phenomena like markets. Absolute success measures are inappropriate<sup>[12]</sup> when agents are reflexive and operating in environments where their actions affect the choices available to others. (It is now becoming clear that these conditions obtain in almost all social processes of any interest.) With the addition of Lamarckian inheritance, evolution is neither incompatible with rational (directed) behavior nor with agents pursuing individual or collective goals.<sup>[13]</sup> Finally, from the institutional perspective, it is encouraging that we can see the current state of social processes as a conceptually smooth progression right back to the most basic forms of social organization.<sup>[14]</sup> Economics has often been accused of "naturalizing" and rendering ahistorical both the attitudes of the affluent middle class and the economic arrangements of late capitalism. By emphasizing process, as institutional economics does for example, the evolutionary approach opposes this tendency. In addition to these direct implications, the evolutionary approach also has potentially appealing rhetorical implications. It is fundamentally pluralistic and dynamic without being teleological. It stresses the descriptive and provisional nature of our understanding and places the understanding of the social scientist in a non privileged position *within* the evolutionary process. Finally, while it does not exclude deliberate or rational action, it situates this action in a context of profound environmental complexity and uncertainty. This context does not arise mysteriously, but naturally from the autonomy of actors and of the physical processes in the environment.

## 2.6

Furthermore, in challenging some of the assumptions of orthodox economics in a constructive way, the evolutionary approach moves it somewhat nearer to the perspectives of the other social sciences. Although sociology, economics and psychology will inevitably continue to differ in their views on the relative importance of social and individual factors, conscious and unconscious processes, rationality and impulse, decision and influence, it is much easier to integrate these insights within a broadly cognitivist framework which views agents (and social scientists) as having limited, diverse and provisional models of the world. It is in part the extremely strong (and occasionally dogmatic) assumptions of individualism, extreme rationality and a positivist treatment of quantitative and strictly behavioral data that have (perhaps unfairly) become identified with economics and effectively separated it from the other social sciences.

## 2.7

In the next section, we move away from biology to a discussion of the way that evolutionary ideas from biology have been used in the development of evolutionary algorithms.

# Evolutionary Algorithms

## 3.1

In this section we will focus on three representative kinds of evolutionary algorithms: Genetic Algorithms (Holland 1975, Holland 1992, Goldberg 1989), Genetic Programming (Koza 1992a, Koza 1994) and Classifier Systems (Holland 1975, Holland 1992, Holland *et al.* 1986, Goldberg 1989).<sup>[15]</sup> In each case, the algorithm involves generating an initial population of individuals and then operating on that population iteratively so that over time the "quality" of the individuals persisting in the population tends to increase. The operations involved and the structure of the individuals in the population are what make the techniques different. Genetic Algorithms (GA) and Genetic Programming (GP) differ mainly in the structure of individuals, while Classifier Systems (CS) not only have a different individual structure from the other two techniques, but different operations too. In the same way that we have abstracted from the intricacies of biological evolution, we shall try to abstract from the implementation details of evolutionary algorithms. Not only are there many different kinds of evolutionary algorithms, but there are many different variants of each kind and many parameter values that may be controlled in each variant. We shall raise these issues only when necessary. For a more detailed overview, we refer the reader to Goldberg (1989), Koza (1992a) and the references in Nissen (1993).

## Genetic Algorithms

## 3.2

In a GA, each individual is represented as a string of numbers which are often binary. These individuals are typically generated randomly.<sup>[16]</sup> The three main operators applied to the population are *reproduction*, whereby more copies are made probabilistically of fitter individuals, *crossover*, whereby two individuals are chosen probabilistically on the basis of

fitness, "cut" at the same point and the tails of the strings "flipped", and *mutation*, whereby a single position in an individual is changed randomly to another syntactically acceptable value. (If the numbers in the string are binary for example, then 1 is flipped to 0 and vice versa.) Fitness is defined by a function, called the *fitness function*, which maps all syntactically acceptable strings into a corresponding fitness. Crossover corresponds to one process of genetic mixing in sexual reproduction, while mutation serves the same diversity maintenance function as it does in biological evolution. While the technicalities are rather complex, it is intuitively clear why the population fitness gradually improves. Reproduction results in a higher proportion of fitter individuals in the population. It is these that probabilistically form the subjects for subsequent crossover and mutation. Unfit individuals which result will not persist, but any that are better than the current individuals will in turn come to dominate the population. (This is where relative performance is important.) Even if the initial population is randomly generated, there will always be one individual that performs the tiniest bit better than the others and this will be enough to start the evolutionary process off.

### 3.3

While the similarity between a GA and biological evolution is fairly clear, there are at least two important differences even at a conceptual level. Firstly, the existence of a fitness function provides both an "invisible hand" and a teleology for the system. The fitness function allows the fitnesses of all individuals in the population to be compared and both proportionate and rank based fitness functions rely on this "global" knowledge. Whatever properties of an individual the fitness function indicates as "desirable" are the properties that will become more prevalent in the population. It is this which provides the teleology for the system. Secondly, no distinction is made between the genotype and phenotype of the individual. The fitness function is able to evaluate an individual directly on a genotypic basis, rather than through some phenotype to which that genotype gives rise.

### 3.4

Both of these differences can be explained by the "standard interpretation" of the GA as an *instrumental* function optimizer.<sup>[17]</sup> If the point of a GA is to serve as a tool allowing the programmer to find the optimal value of some fixed function, then global knowledge and a designed fitness function are both appropriate and necessary to that task. Rather than modeling natural selection, the programmer is attempting husbandry, incorporating an exogenous fitness function into the program in the same way that animal breeders measure their success by the amount of extra meat on a new strain of cow relative to the current herd.<sup>[18]</sup> However, for the purposes of this paper, we are interested in the extent to which evolutionary algorithms are capable of serving as *descriptive* models of social processes. (In such cases, the assumptions of global knowledge and a shared teleology are probably not appropriate.) In the same way that we have described evolutionary models in general as those which specify a coherent analogy with biological evolution, so we can envisage simulations based on coherent analogies with evolutionary *algorithms*. It is to these analogies we turn in section four, but for now, we continue with the technical description of evolutionary algorithms.

## Genetic Programming

### 3.5

GP also typically makes use of the operators of reproduction, crossover and mutation. However, the individuals in the population are *programs* in some (simplified) programming language. A commonly used language consists of the set of possible trees for a specified set of operators (internal nodes) and terminals (leaf nodes).<sup>[19]</sup> For example, the set of operators could be  $\{+ | - | * | / \}$  and the set of terminals  $\{0 | 1\}$ . This language would produce a set of trees involving simple mathematical expressions of arbitrary depth.

### 3.6

Again, it is clear that this technique has a simple instrumental interpretation and GP has indeed been used for pattern induction and function estimation (Koza 1992b). However, there has also been a significant increase in representational richness in at least two senses that prove to be relevant for descriptive models of social processes. Firstly, the tree representation allows for individuals of arbitrary depth.<sup>[20]</sup> Secondly, the interpretation of a GP tree is decomposable in a way that a GA string is not. If the operators are chosen to have "obvious" meanings, so that + refers to the addition operation, for example, then GP trees can be parsed by a human almost as easily as by a computer. Decomposability is demonstrated by the fact that an individual can be presented with a syntactically correct GP tree fragment like  $(+ (+ 3 5) ...)$  and parse it to  $(+ 8 ...)$ . By contrast a GA string fragment  $(... 0 1 0 ...)$  makes no sense unless one knows its exact position in the string and even then it is necessary to refer to the coding of the fitness function.<sup>[21]</sup> This decomposability is not purely a matter of comprehensibility, it also corresponds to a specific problem with the GA as an instrumental optimizer. Although there is no technical reason why a GA string should not encode an arbitrarily complex structure, the more complex it gets in practice, the greater the danger that optimization will be dramatically impaired by the unequal importance of different string positions in producing fit solutions. Instead of truly parallel search, the GA will tend to optimize the most important positions first and then progressively optimize the less important ones. (This is the problem of *epistasis* Goldberg 1989: 46-48).<sup>[22]</sup>

### 3.7

Despite this increase in representational richness, GP techniques are still used predominantly in an instrumental way (Koza 1992a). They are set up to assume global knowledge, no phenotype/genotype distinction and the exogenous teleology of the fitness function. However, the richer representation possible in Genetic Programming has drawn attention to the fact that even for instrumental applications:

It might be appropriate to evaluate the fitness of individuals on the basis of repeated program outcomes and that these outcomes need not be separable;

Terminals and operators could just as easily be interpreted as *actions* in a simulated environment.

### 3.8

An obvious example of the former is evolving a program to generate random numbers (Koza 1991). In this case, the fitness must measure the statistical properties of *sequences* of numbers and can do nothing with single instances. As an example of the latter, consider evolving a program to serve as a controller allowing a robot to perform a simple task like finding litter and picking it up (Koza 1992c). In this case, the "output" of the GP can correspond to an *action* in a simulated environment like "Turn Left" or "Pick Up" rather than simply generating a number. (For instrumental applications, the programmer can *interpret* a number as indicating an action, but for descriptive simulations avoiding an external teleology, it is required that the interpretation be encoded in the program itself.<sup>[23]</sup>) In addition, we are now making a concrete distinction between genotype and phenotype. The program is the genotype, but the execution process and actual outcome of executing the program in the simulated (or real)<sup>[24]</sup> world involves the phenotype.

Trivially, if the indicated action is "Move Forward" but the robot is facing a wall, the phenotypic outcome will be a dented robot! Although both of these applications are still instrumental, it is much shorter step, at least in the latter case, from asking how we might get a robot to optimize its performance of certain tasks given its abilities, to asking what behavior we might expect to observe in an agent with certain abilities, when it is exposed not to teleological selection, but to selection by the physical environment, by other agents, or a combination of the two.<sup>[25]</sup> We consider various answers to this question in section five.<sup>[26]</sup>

## Classifier Systems

### 3.9

The final class of evolutionary algorithms we will discuss in this section are Classifier Systems (CS).<sup>[27]</sup> In a CS, the individuals consist of IF ... THEN ... rules (condition-action pairs).<sup>[28]</sup> The architecture of the CS is rather more complicated than those of the GA and GP. It begins when suitably encoded input from the environment is added to the "message list" of the CS. This input may satisfy the conditional part of one or more rules in the "rule base". Each rule has an associated strength, which reflects its past success at contributing to the generation of "appropriate" output. Eligible rules compete in an "activation auction", a probabilistic process in which fitter rules are better able to make high "bids" for activation on the basis of past performance. Once it has been established which rules are to be activated, they are copied to the message list and the combination of inputs and rules is "executed". (The combination of inputs and rules may generate outputs to the message list which in turn make new rules eligible.) As well as outputs to the message list, some rules will presumably generate outputs to the environment and these are also carried out. After this, the message list is purged. The outputs to the environment generate environmental feedback, which is used to alter the weights on rules.<sup>[29]</sup> Finally, a GA is used from time to time to weed out lightweight rules in the rule base and generate new ones probabilistically from those with greatest weights.

### **3.10**

Despite its apparent complexity, the CS actually functions in a rather similar way to the GA and GP. The population of the rule base should gradually improve in quality as time passes. The activation auction corresponds to the probabilistic application of genetic operators like crossover and reproduction to determine which individuals are actually "used" for tasks. The weights on rules can be seen as an alternative to duplicating fitter rules through reproduction in the GA or GP population. (One rule with a weight of 1 and four with weights of 0 are equivalent to a population of any size consisting only of multiple copies of the first rule. The analogy is not perfect. In the latter case, the rule diversity has really been lost, while in the former case the rules are still "available" if the weights should change. We will return to this point.) The "environment" can either be some instrumental task like summarizing a set of cases, or a descriptive simulation of a real environment. As with the GP, the existence of rules which generate outputs to the message list as well as the environment allows for output behaviors of arbitrary complexity. The most important difference between the CS and GA/GP is that the CS explicitly produces strategies that consist of *sets* of rules which deal with a particular environment collectively. By contrast, even though GA and GP produce diverse populations, the idea is that only the fittest individual in that population will be "used" at any one time.<sup>[30]</sup> This difference supports descriptive uses of CS and raises questions about the flexibility of the responses of simple GA and GP to systematic environmental variation. Although a single program in a GP *can* cover all the cases in an arbitrary rule base, the resulting program will be very large, unwieldy and perhaps difficult to interpret.<sup>[31]</sup> In addition, the majority of the program will be redundant most of the time.

We will return to this issue in section five.

### **3.11**

To sum up, it can be seen that even in their most simplified forms, evolutionary algorithms are a creditable abstraction of the biological evolutionary process. However, some caution has to be exercised in using them as models of social processes. In particular, their origin as instrumental optimization techniques may be incompatible with desirable features of biological evolution like the absence of global knowledge and external teleology. In sections four and five of this paper we will focus on a number of specific simulations to see what they reveal about the potential and limitations of evolutionary algorithms as representations of *social* processes.

## **The "Standard Interpretation" Of Evolutionary Simulations**

### **4.1**

Although the simulations differ in detail, it is possible to present a "standard interpretation" of genuine evolutionary simulations which forms the basis for discussion in this section. (Section five attempts to generalize and develop this interpretation.) The standard interpretation draws on several different strands of thought from unorthodox economics. Two of these will be addressed now and others will emerge as the discussion proceeds. As has already been remarked, many of the insights which follow are taken for granted in sociology and psychology.

## 4.2

Probably the most important ideas underpinning these simulations are that rationality is genuinely bounded and procedural (Simon 1972, Simon 1976, Simon 1978, Simon 1981). Unfortunately, until relatively recently, the idea of bounded rationality seems to have been reincorporated into orthodox theory as a generalization of standard economic rationality rather than an ontologically different theory of human decision.<sup>[32]</sup> For example, the observation that decision making is time consuming and cognitively difficult formed the basis for models in which agents were supposed to optimize their allocation of time to the decision process (Winston 1987). The same logic was used to transform the observation that individuals and organizations tend to use "rules of thumb" into an optimizing theory (Baumol and Quandt 1964). Simon himself is responsible for two of the most damaging criticisms of this approach.<sup>[33]</sup> In the first place, it involves an infinite regress of explanation. Does the agent include time spent deliberating about how long to deliberate as part of the calculations for the deliberation? Is the rule of thumb chosen using a rule of thumb? If so, how is *that* rule of thumb chosen? In each case, what appears *prima facie* to be an explanation actually involves postponing any real grounding for the theory? Lest it be thought that this is no more than the general problem of reductionist science, Simon further points out that the regress is not leading to explanations of equal or increasing simplicity and power, which would be helpful,<sup>[34]</sup> but to explanations of increasing difficulty, which is not. It is intuitive that the rule of thumb for *choosing* rules of thumb must be at least minimally more complex than the rules which it must choose between since it must be capable of not only "executing" those rules but evaluating them. A related point is that information requirements increase for each level of regress. If the rules of thumb contain any parameters which must be tuned for effective use, the rule of thumb for choosing must have *its* parameters tuned to take account of different values for the rules of thumb it is supposed to decide between.<sup>[35]</sup>

## 4.3

The solution is obvious, but extremely different from the orthodox view. Instead of rationality consisting of a "disembodied observer" who is able to look down on mental states from whatever height is required to get the complete picture and act "rationally", to some extent, rationality *must* be "embedded" in the unfolding of those mental states to avoid an infinite regress. The metaphor of the observer is a useful one. The fact that a human observer is embedded in the world certainly does not preclude them having some choice over which direction to look in, though parts of any view will be obscured by trees, fog or whatever. The observer can also change their point of view by walking or going up in a balloon, for example, and this may clear some obscurities, though it will almost certainly introduce others. However, at the very least, any observer is typically prevented from observing the state of their own forehead! If a mirror is used, this observation is possible, but at the expense of any view of what is going on behind the mirror. This example is a purely heuristic one. It shows that any embedded system has at least one "blind spot" in addition to its obscurities, but says nothing about how big it is. The nature and scale of obscurities and blind spots in the mental process is an open question, but we can be almost certain that rational choice theory, which postulates no genuine blind spots at all, is incorrect. Furthermore, the existence of even the smallest blind spot changes the whole nature of the decision process, since any decision can no longer be absolutely known to be rational. This

difficulty is more serious than that of mere probabilistic risk because there can also be no knowledge of the degree of uncertainty induced by the blind spot, since its "size" and "scope" cannot be determined either.<sup>[36]</sup>

#### 4.4

The implication of this embeddedness is that at some level, as Simon argued, rationality can only consist of "unselfconsciously" following procedures. Two procedures can be compared, but they are compared by something that is itself a procedure and cannot be rationally justified beyond a certain point.<sup>[37]</sup> (Wittgenstein 1953 makes the same point in his discussion of "language games".) One cannot step outside the whole mental process and "straighten it out".<sup>[38]</sup> At best, from any particular mental point of view, one can make comparisons and attempt to remove inconsistencies. (Always allowing that changing things to remove one local inconsistency may introduce another somewhere else.<sup>[39]</sup>)

#### 4.5

This interpretation of bounded, embedded or procedural rationality strongly suggests a computational approach. Although a computer program can modify the state of its variables, and in some cases even its own code, at any point it must take the variables and state of the program as given when executing a particular instruction. Thus, in addition to the general advantages of computer simulation (Doran and Gilbert 1994, Chattoe 1996), there is also a conceptual similarity between programs and mental processes.

#### 4.6

A second related theme, which is suggested by the absence of rational justification for procedures, is that agent models of the world are something quite distinct from the world itself (Chattoe 1994, Moss and Edmonds 1994). Agents do not, as orthodox economics sometimes suggests, all have knowledge of the correct model of the world or even of the same model with divergent opinions about its parameters.<sup>[40]</sup> It is only in the context of an evolutionary theory, which assumes diversity, that the enormity of this assumption becomes clear. Everyday experience overwhelmingly suggests that agents differ not only in their preferences, but in their calculative techniques, their views on achieving goals and so on. Situations where beliefs or goals based on anything other than direct perception are widely shared are not the norm, but a fascinating field for social research.

#### 4.7

The analogy between this view of agency and the description of biological evolution in the second section is fairly clear. Actors, which may be individuals, families, social groups, firms or governments, can be viewed as having both a genotype, which consists of their more or less *persistent* mental processes and operating procedures, and a phenotype, which consists of all the other "physical" aspects of their functioning, including the impact which they can have on the environment. The genotype is distinct from the phenotype, but "constructs" it, since we assume that actors are intentional.<sup>[41]</sup> Phenotypes, and through them genotypes, experience selective retention by the environment. In addition, the processes of genetic mixing and mutation also

have their analogues. Genetic mixing occurs during the process of learning by imitation and mutation can result either from incorrect copying of a genotype or random (undirected) experimentation. (These processes will be analyzed in more detail in section five.)

#### 4.8

The implications of biological evolution sketched out in the second section are also congruent with this analogy. In the great majority of cases, "fitness" in social situations must be measured in relative and not absolute terms. The complexity of the environment (consisting of the autonomous and interacting behavior of diverse individuals as well as purely physical processes) means that credit assignment for particular modifications to the genotype is extremely hard to implement. This affects the type of learning and adaptation mechanisms we can expect to observe. The *ceteris paribus* requirement (to which orthodox economics often appeals) can seldom even be maintained in the genotype of a single individual. In particular, the refinement of cognitive models involves the acquisition and organization of new knowledge. This means that cognitive agents cannot easily "step into the same river twice" as far as their perception of the environment is concerned. This in turn means that agents will perform new actions, based on new inferences, and influence the environment in new ways, requiring other agents to modify their practices accordingly. This gives a "direction" to the social process and removes the costless reversibility which is typically assumed in orthodox economic models.<sup>[42]</sup>

#### 4.9

This leads on to the fact that social processes are also both open ended and progressive. Simple arrangements for buying and selling, fighting or building settlements gradually diversify and increase in complexity as knowledge is created and organized by agents. Over time, genuinely new products, institutions and technologies are incorporated into the social process.<sup>[43]</sup> Open endedness reminds us that economic systems, like biological ones, do not imply or require an external teleology, even though it is almost a matter of definition that autonomous agents should have *internal* goals. This point needs some clarification, since in some contexts, particularly the interaction of firms in a market, the absence of an external teleology is somewhat obscured by "market rhetoric".<sup>[44]</sup> It is certainly true that there are many laws and customs, more or less well observed, designed to regulate the function of markets and other social institutions, but not that they reflect any coherent or agreed plan to support "efficiency" or indeed any unitary goal, nor that they can be expected to ensure any particular social behavior such as profit maximization in competitive firms.<sup>[45]</sup> Indeed, the structure of institutions like the market often reflects some of the interests of the very "players" who are supposed to be "regulated" by them.<sup>[46]</sup> Although firms will adapt to the structure of the market place, this structure is not coherent, binding or permanent enough to constitute a teleology, any more than it would be valuable to talk about other individuals constituting an external teleology.<sup>[47]</sup>

#### 4.10

Finally, although we have so far focused on a pure evolutionary process, we are not obliged to exclude various mechanisms of Lamarckian inheritance in social systems, because it is much easier to see how phenotypic effects can be re-encoded into the genotype when the material of

the genotype is everyday language rather than DNA. This process of re-encoding is precisely what we mean by directed or "rational" adaptation, as opposed to evolutionary adoption. It is important to stress both that evolutionary models are *not* intended to exclude rational or deliberate behavior and that the evolutionary view of bounded rationality is actually straightforwardly compatible with many forms of learning. Many critics of evolutionary models (Penrose 1952) accuse them of asserting that human behavior is either random or genetically determined. This is simply incorrect. These same critics also fail to recognize that the technical reasons why Lamarckian inheritance cannot occur in biological systems are not necessarily binding on social systems. Finally, while it is true that evolutionary models can easily incorporate learning as Lamarckian inheritance, attempts to incorporate evolutionary behavior into "rational" models have so far been much less successful. Two of these attempts, the GA models of Arifovic and the use of replicator dynamics in game theory, will be discussed subsequently. On the basis of the foregoing discussion, two of the most important general questions raised by evolutionary models are:

How much rationality is actually required for the level of organization we observe in social processes?

How can social evolutionary processes *interact* with rational decision?

#### **4.11**

On the basis of the evolutionary models discussed in this paper, the answer to the first question is considerably less than orthodox economics has previously suggested. The answer to the second question is much less clear cut and one of the reasons for attempting a coherent interpretation of evolutionary models is precisely to make the boundaries between evolutionary adoption and rational adaptation clearer. (If economics is guilty of over-playing the role of deliberate action, then perhaps sociology and social psychology are guilty of under-playing it.) Before moving on to a detailed discussion of models in the next section, the last purpose of this section is to examine two interdependent aspects of the analogy between biological and economic systems that are traditionally regarded as problematic.

#### **4.12**

These aspects are the relationship between the measurement of individual "lifetimes" and the objects of evolutionary selection. For biological evolution, in the absence of Lamarckian inheritance, there is a unique relationship between a genotype and a corresponding phenotype. Once we allow Lamarckian inheritance, this is no longer true and a particular individual can no longer be identified with a fixed genotype. However, this does not, as Penrose suggests in the context of industrial evolution, raise any difficulty with the "identity" of firms, nor are we obliged to treat each change in the genotype as the creation of a "new" firm.<sup>[48]</sup> Any such argument fails to understand the definition of Lamarckian inheritance. It is precisely a change in the genotype *during the lifetime of the individual*. One of the things that is often understated in evolutionary modeling is the empirical integrity of the objects of selection which typically form the basis for models. It is trivial that individuals display the properties of organisms: autonomy, internal structure, a well defined boundary and mechanisms of self-maintenance. However, these

properties are also very much in evidence in organizations and social groups of all kinds, and particularly in firms with actual or potential competitors. The boundary is maintained by the distinction between employees and others, the use of "non disclosure" agreements for departing employees, the importance of tacit knowledge and "on the job training" and by more intangible phenomena like "company pride". The combination of a correct interpretation of Lamarckian inheritance and the observed integrity of individuals and groups resolves the apparent difficulty. The objects of selection really are organizations and individuals and their lifetimes really are the periods for which they persist. This resolution does however raise a subsidiary issue, which is that very few social processes appear to take place over the real life span of groups or individuals.<sup>[49]</sup> If this were actually so, then the role of evolutionary mechanisms would of necessity be considerably weakened. The answer is that although we have stressed Lamarckian inheritance as a way of fitting *directed* learning into an evolutionary framework, there is no conceptual reason why it shouldn't equally apply to the effects of undirected learning within the individual lifetime. All that remains is to distinguish more clearly what we mean by directed and undirected learning and that is a task postponed until later in the next section.

## Focusing and Developing the Standard Interpretation

### 5.1

Ironically, but perhaps not surprisingly, two of the best known kinds of evolutionary models are the least satisfactory from the point of view of the foregoing discussion. This is because they attempt to reincorporate evolutionary ideas into an orthodox economic framework without due regard for their logical coherence.

### How Not To Use Evolutionary Algorithms

### 5.2

The first set of models is GA based and has been developed by Arifovic and others (Arifovic 1990, Arifovic 1994, Arifovic 1995, Arifovic 1996, Arifovic and Eaton 1995, Bullard and Duffy 1994). In these models, pricing or other decisions are coded as GA strings which then undergo the usual operators of reproduction, crossover and mutation, along with an additional *election* operator which will be explained shortly. Arifovic (1990) draws attention to two different interpretations of the GA, which will be useful in what follows. The first could be called the *mental interpretation*, in which GA strings represent different possibilities in the mind of a single agent or organization. Under this interpretation, individuals carry out the operators on their own mental contents, using the fitness function to pick the string which actually generates an action in any particular period. The second is the *population interpretation*, which is the one that Arifovic favors and uses subsequently. Under this interpretation, each individual consists of a single string and the GA consists of a *population* of individuals. The first difficulty with these models is the criteria by which they should be judged. The behavioral interpretation provided is cursory and it appears that the objective of these models is to simply to obtain or "explain" convergence to the equilibria indicated by analytical models. In particular, Arifovic (1990) is quite specific about introducing the election operator to deal with what she describes as the "problem" of non-convergence.

### 5.3

For now, we will assume that Arifovic intends her model to be interpreted descriptively. The interpretation of crossover as learning by imitation and mutation as copying error or random experimentation are fairly straightforward. However, the absence of a genotype/phenotype distinction and the choice of crossover as the mixing operator have important implications. It appears that one agent can look directly into the head of another and perceive what it intends to do.<sup>[50]</sup> Furthermore, crossover imposes a lot of "structure" on the imitation process, ensuring that it always produces new strings that are syntactically correct. Thus agents can not only read each other's minds, but do so with such clarity that they always imitate fully formed "concepts" despite the highly structured encoding of the GA string. (This structure may also cause the problems with mutation that lead to the introduction of the election operator. Mutation at some positions may have far more effect on the fitness of the string than at others. This is undesirable purely from the point of view of design.) The reproduction operator produces larger numbers of fitter strategies in the population. This operator can be interpreted in two ways, but both are problematic. The first is that the environment imposes the removal of less fit firms and, by implication, the increase in the population of fitter firms. However, while it is easy enough to see why unfit firms might be removed by bankruptcy, it is much harder to say what properties replacement firms would have, and in particular how they would be able to imitate the practices of only fitter firms. (The argument usually given is that new firms adopt market "best practice" but this assumption needs more support. In particular, if this option is available to new firms, why is it not open to current firms performing poorly?) Nothing is said about how the time scales implied by this interpretation can be reconciled with the speed of environmental change in real markets. The other possible interpretation is that reproduction represents imitation of the whole strategy of another firm on the basis of fitness. The two difficulties with this interpretation are firstly, that observing utility is even more implausible than observing intentions<sup>[51]</sup> and even if we grant this as a possibility, secondly, there seems to be an strong implication of global knowledge to ensure that people imitate in proportion to the overall distribution of fitness. Another possibility is that firms could "learn" the fitness distribution in the market, but it is not clear that they could do this fast enough in a dynamic situation. (Both interpretations of the reproduction operator also suffer from the absence of a clear genotype/phenotype distinction, since it is not obvious whether we are talking about the success and failure of real firms or of the strategies adopted by those firms.)

### 5.4

The interpretation of the election operator is similarly ambiguous. This operator is applied to the process of crossover. Both the crossover products are assigned a *potential fitness*, an expectation of how they are likely to perform on the basis of their performance with past data. If the potential fitness of either crossover products exceeds the actual fitness of the "parents", then it will replace that parent in the population. Full crossover is far harder to justify than a "half crossover" (imitation) and without that, there will not be a pair of crossover products or parents. We *can* apply the same algorithm to the decision as to whether or not to incorporate an imitated section of string from another agent, but it is still not clear what it is about the organization of the environment that ensures that agents are presented with imitation opportunities in proportion to their fitness. (Again, a real phenotype might provide a solution, with firms tending to imitate

those displaying similar patterns of *outward* behavior or observable features like size and age. This is the approach taken in Chattoe and Gilbert 1997.) In fact, it seems that if agents are so well informed about fitness and the intentions of other agents, they should be able to decide for themselves whether or not to imitate, regardless of the state of the other firms they are presented with. Thus election is in danger of rendering reproduction redundant and turning the GA into a hill-climbing algorithm with all the attendant drawbacks such as non global convergence.

## 5.5

The other class of well known evolutionary models are those in game theory using the so called replicator dynamics.<sup>[52]</sup> This is a rapidly expanding field (Weibull 1995, Vega-Redondo 1996) and there is not space to provide a detailed survey here. The main point that is relevant to the current discussion is that the behavioral rationale provided for these models bears even less cursory inspection than that in the Arifovic models. (Further evidence for this is provided by the fact that, as far as I know, there have never been attempts to fit game-theoretic models of this kind to any data or even to suggest social processes from which these data might be collected.) Replicator dynamics originated in theoretical biology, where they were intended to model shifts in the proportions of genes or species of differing fitness. In extremely simple terms, a gene or species will increase (decrease) its share in the population if its own fitness is greater (less) than the average fitness and the change will be proportional to the difference between the average and the particular gene/species fitness. All models of this kind face the same difficulty. Replicator dynamics is based on the assumption of an underlying biological process which justifies its mathematical form.<sup>[53]</sup> For example, it is assumed that species display exponential population growth when they are not resource constrained. In applying such models to social processes, the results will only be useful if the same assumptions apply. Without exception, replicator dynamics models justify themselves with reference to discursive behavioral arguments about imitation and observation errors and not by reference to the other underlying assumptions of the original model. This produces a number of problems:

How do individuals get information about average fitness, or even observe the fitnesses of other individuals? Usually these models simply substitute an unexplained rational theory of imitation for a rational theory of individual game play.

How do individuals correlate game strategies with decision processes? In far too many models, this problem is assumed away. Either the individual just *is* their strategy, or else individuals are both telepathic and know the full set of possible decision processes. The latter assumption is required to ensure only legal mutations.

Biological models assume competition for scarce resources like food and shelter which directly impact on reproductive success through physical mechanisms. The possibility of making correlations between certain independently observable and measurable skills (like foraging rates) and reproductive success makes the concept of selection by reproductive success both non tautologous and potentially falsifiable, though this is very hard work outside the laboratory. In games, the resource "emerges" only in game play and its only value to the player is that it enhances the likelihood of strategy propagation. While a strategy like "co-operate" is observable and can be correlated with reproductive success (though with the reservation that this success

must actually be correlated with all possible strategy combinations), this correlation is meaningless because the only "measure" of the strategies performance is precisely the same "utility" which determines that success.<sup>[54]</sup>

## 5.6

What lessons can be learnt from these unsatisfactory models? The first is that the interpretation of evolutionary models *does* matter on at least two levels. On the methodological level, very few unorthodox economists (or non economists come to that) subscribe to the "as if" argument for modeling, which states that it is irrelevant whether a model is behaviorally plausible as long as it predicts adequately. In any event, it has no force in areas like evolutionary game theory where predictive models are not available. It is not enough for those who believe in procedural rationality to know *what* happens, they also want to know *how* it happens. In that sense, an interpretation is just a more complete specification of the *process* by which evolution could take place in a social rather than biological context. On the practical level, although the Arifovic models *are* used to make limited sense of experimental data, the behavioral interpretations used are in the main either inconsistent or implausible. This could be avoided if the behavioral interpretation of the model was more than a rhetorical defense of the selected formalism. This brings us on to the second lesson, which is that a working knowledge of the history and implications of evolutionary biology and evolutionary algorithms appears to be important in their sensible application to social science.<sup>[55]</sup>

## Simulations Based On the Mental Interpretation

## 5.7

In this section I will discuss a series of models which appear to achieve a more plausible behavioral interpretation in the use of evolutionary algorithms. Simulations using the mental interpretation and based on Genetic Programming have been produced by Chattoe and Gilbert (1996), Chattoe and Gilbert (1997), Dosi *et al.* (1994), Edmonds (1997) and Edmonds and Moss (1996) among others.<sup>[56]</sup> These models have been applied to the behavior of households learning to budget, firms setting prices in a market and modeling individuals who face an environment where having the same world view as everyone else is undesirable. Despite these differences, we can identify two sets of common questions which must be addressed by each of these models.

## The Interpretation of the Agent Decision Process

## 5.8

All the questions in this group depend on how much detail can we provide about the "mental process" corresponding to an internal evolutionary algorithm. It is clear that for any problem, an individual or organization will have a number of candidate solutions which seem feasible. The reproduction operator corresponds to a process by which some candidate solutions become "more favored" in the population. The crossover operator corresponds to a process in which one candidate solution is modified by the incorporation of a sub tree of material from another. The mutation operator corresponds to a process by which some part of a candidate solution is

replaced by a *novel* sub tree or "innovation". All three operators require more detailed interpretation.

### **The Interpretation of Reproduction**

#### **5.9**

How is the process of reproduction represented in the state of the population? We have already discussed in passing that the weights assigned to Classifier System rules can be seen as corresponding to the duplication of fitter strategies in a traditional GA. However, the behavioral meaning of strategy duplication is much less clear than that for some sort of weighting or "confidence measure" attached to particular strategies. In particular, there are empirical features of human cognitive abilities which do not correspond well to the duplication interpretation. The most important of these is the existence of long term memory. Although it is true that humans can hold rather few strategies "in mind" at any one time, they are extremely good at remembering strategies which they have already implemented or observed elsewhere. Thus the traditional GA or GP architecture, which implies parallel processing of operators applied to many candidate solutions is probably not appropriate for models of individual decision making.<sup>[57]</sup> A behaviorally plausible alternative exists, in the shape of the GENITOR algorithm (Whitley 1989). In this algorithm, candidate strategies are permanently ranked by fitness. In each generation, instead of parallel processing, strategies are chosen proportionate to their ranking and a single operator is applied.<sup>[58]</sup> Offspring of crossover and mutation are returned to the population at the "appropriate" positions indicated by their fitness. They either replace the strategies with the nearest fitness or are inserted into the ranking so that the least fit strategy "falls off the bottom". This approach has three advantages. Firstly, it explicitly avoids duplicates if strategies replace those with the nearest fitness or can easily be made to do so under the insertion strategy. The "importance" of strategies is represented by their position in the ranking rather by the existence of duplicates and insertion does not require the re-evaluation of fitnesses for all other strategies.<sup>[59]</sup> Secondly, this ranking also remains relatively stable because operators are applied one at a time. (The flipping from strategy to strategy displayed by non converged traditional GA or GP is not self-evidently plausible as a representation of human behavior.) Finally, the algorithm relies on the "active consideration" of no more than two strategies at any one time, all other strategies are simply stored in memory. However, there is a drawback to this approach, for which detailed discussion will be postponed. In an instrumental GA, the fitness function is known and fixed so new strategies can be "placed" in the ranking immediately. It is not clear what should be done to them if fitness evaluation requires interaction with the environment. We will return to this point.

### **The Interpretation of Crossover**

#### **5.10**

Is crossover to be proportional to some measure of fitness? If so, what measure? In the GENITOR algorithm we have already discussed, crossover is more likely to occur among solutions with higher ranking. In an organization, this could be interpreted as reflecting the number of people preferring and arguing for a particular strategy. In an individual, it could

reflect the fact that better strategies are more likely to "spring to mind" or remain in the forefront of memory. One thing to bear in mind is that once we become interested in descriptive models, it is more appropriate to consider what is realistic than what produces rapid convergence and this applies just as much to the parameters of the simulation as it does to the selection of the evolutionary algorithm. (We may even become interested in distinctive *failures* of convergence if these can be identified empirically.) Simulations are a very useful technique for comparing different possibilities in this context. This point is just as relevant to the assumption of strictly proportional fitness as it is to the size of the population of candidate solutions discussed in the last section. Just how much do we lose if the population is 5 rather than 50?<sup>[60]</sup> How much slower would convergence be if solutions were only ranked as good, average and poor or if solutions were ranked with probabilistic accuracy?

## 5.11

If the exogenous fitness function used in the Arifovic and replicator dynamics models is regarded as imposing too much structure on the evolutionary process, then at least two alternatives exist. The first appears not to have received much attention. This would involve each individual using an internal fitness function which is nevertheless regarded as being "infallible" by that individual. Even if individuals have identical internal fitness functions, they avoid the problem of external teleology, but this approach also allows for the fact that agents may simply have different goals.<sup>[61]</sup> In addition, it is possible to investigate how the environment selects fitness functions at the same time as it selects strategies which are fitter relative to those fitness functions it selects.<sup>[62]</sup> One of the reasons why the "internal fitness function" approach has not received much attention is that it goes against an important assumption of most bounded rationality models, which is simply that agents cannot accurately rank the fitnesses of strategies *a priori* except in trivial environments. Instead, they can only evaluate strategies in use and make comparisons on the basis of what has already happened. This is the basis of the Dosi *et al.* (1994) simulation, in which firms use cumulated profits to choose GP strategies for price setting on a probabilistic basis, with a minimum probability for previously untried strategies arising from crossover or mutation on the current population. (Depending on how mutation is defined, it is capable of generating whole new strategies.) Strategies which fail on any one of three conditions will be replaced, either with new strategies or crossovers/mutations of current strategies. The first failure condition is that the probability for a strategy being picked should fall below a minimum value. The second is that a strategy should generate a negative profit. The third is that it should indicate a price outside a rather broad range of "sensible" prices which runs between 0 and some large positive value. This model suggests a number of points. Firstly, its functioning is rather similar to the GENITOR algorithm in that the methods for generating replacement strategies tend to preclude duplication and force out the worst strategies disproportionately. Secondly, there is negligible calculation involved in the decision process. All that is required is a running total of cumulated profits for each current strategy and maximum and minimum values for sensible prices. Finally, the model draws attention to the distinction between syntactic (genotypic) and semantic (phenotypic) evaluation of strategies. It is possible to reject a strategy because it generates what is believed to be a "silly" price without reference to what happens when it is tried out in the market. By contrast, trying it is the only way that a strategy can be found to be unprofitable. Although we have already cast doubt on the plausibility of internal fitness functions which are capable of ranking strategies, it is much more plausible that there

should be internal *filtering* processes that cut down the size of the strategy space based on general domain knowledge (McCain 1992).

## The Interpretation of Mutation

### 5.12

Does mutation correspond to "transcription error" or experimentation? This is a relatively minor point, but it affects the fitness attributed to new strategies. In the Dosi *et al.* model, new strategies resulting from crossover are assigned the average of the cumulated profits of both parents. (This is the nearest that firms come to doing any calculation.) If mutation is interpreted as transcription error, then the mutant should be attributed the same cumulated profit as its parent since by definition such errors would be corrected if they were noticed. This could have unfortunate effects if a mutant was much poorer than a parent which had already cumulated a large profit. (In fact, this is one potential drawback with the Dosi *et al.* model, that firms have an infinite memory for cumulated profits, thus leading to possible lock in. On the other hand, as we have already remarked, evolutionary algorithms which produce too much strategy vacillation may be empirically implausible.<sup>[63]</sup>) By contrast, if mutation corresponds to experimentation then all mutants should be treated as untried strategies or perhaps receive the parental fitness penalized according to some distance metric between parent and offspring.

### 5.13

This section suggests that with care, coherent interpretations can be produced for all the traditional genetic operators. These can be fitted together within the framework of the GENITOR algorithm, with some modifications to allow for the fact that strategies cannot be evaluated for actual fitness without trial. Strategies remain ranked by fitness (cumulated profit for example) at all times. New strategies are assigned a minimum fitness (or probability of being chosen) reflecting their untried status. Crossover products receive the average fitness of their parents. Mutation products are regarded as untried.<sup>[64]</sup> The parents are chosen probabilistically on the basis of their ranking as is the strategy which is actually to be implemented. Tried strategies that yield negative profits or produce "silly" prices may still remain on the list as a mechanism to prevent their being reintroduced, but they will have no possibility of being chosen as parents or of actually being implemented. (Firms will reject crossovers, mutations and novel strategies if they have already been tried within institutional memory or are currently "under discussion".) This algorithm has a number of desirable features. Firstly, the probability distribution for strategy implementation need not be the same as the probability distribution to select parents for the genetic operators and neither needs to be equivalent to the institutional memory of the firm. For example, the implemented rule could just be the top ranked one at all times. This would make firm behavior very stable. For equivalence to the Dosi *et al.* model, parent selection would take place from the set of all strategies which had actually been tried and achieved a cumulated profit greater than the minimum for untried strategies while implementation would take place from the set of all strategies including the untried ones. Secondly, the evolutionary process is "driven" by the firm's view of its own institutional memory, without this making it a rational choice process. While the GENITOR algorithm performs one genetic operator every period, a firm will only attempt to identify new strategies when it is unhappy (in some sense to be defined)

with the set which it already has. If the firm always wants to entertain four possibilities, then initially it will have four untried strategies, but as it tries them, they will get ranked. Only strategies which perform so poorly that they are probabilistically less likely to be chosen than untried strategies (which are assumed to have a fixed minimum probability of being chosen reflecting their "curiosity" or "novelty") will encourage the firm to cast about for new ones. This also produces a measure of realistic stability into the behavior of the firm. Thirdly, this algorithm not only incorporates the coherent interpretation of the operators, but takes into account earlier comments about the distinction between institutional memory and active consideration, the behavioral implausibility of duplicate strategies, the limits on what firms can know and calculate and the possibilities for syntactic reasoning about strategies even when semantic reasoning is not possible. It should be made clear that this algorithm is only being proposed for further investigation rather than advocated as the *correct* model of evolutionary decision. It does however seem to have a number of desirable features which may make it worthy of further consideration.

## 5.14

In the next section, we turn from discussion of the decision processes of the agent to their interaction with other agents and physical processes in the environment.

### The Interpretation of the Environment

## 5.15

So far, individuals under the mental interpretation have just been concerned with evaluating strategies they have generated "internally". Depending on the nature of these strategies, it is also necessary to develop a coherent interpretation of the individual's relationship with the environment of other individuals and physical processes.

## 5.16

The first issue arising in this context is a general one, concerning the open endedness of biological and social evolution and the way this compares with the functioning of evolutionary algorithms. *Prima facie*, it may appear that the problem with evolutionary algorithms is that they have a fixed set of terminals and operators while biological evolution does not. In fact, this is not the case. Within certain limits, like the coining of new scientific terms, any human language also has a more or less fixed set of components, though it is obviously far larger than the typical GP language. It is the *combination* and *structuring* of these components that can accommodate both the richness of sonnets and the pragmatism of operating instructions. However, neither richness nor pragmatism can be attributed purely to the combinations of words themselves. Instead, it results from some comparison between the words and some other aspect of reality. In the case of operating instructions, the words are compared with perceptions of the thing to be operated and feedback on whether implementing the instruction produces the desired result. In the case of a sonnet, the words are "compared" with memories of summer days and sweet spring showers to see if they evoke sensations appropriate to what the poet is talking about. In each case, it is the *interplay* of genotypic and phenotypic aspects that gives rise to richness. In an instrumental GP,

the fitness function just summarizes the environment. In a descriptive GP, open endedness arises not just from enlarging the GP vocabulary, but equally from enriching the environmental consequences of phenotypic behavior to which it gives rise. For example, instead of the terminals "pick up box", "pick up cone" and "pick up sphere" in a simple blocks world, we might have the terminal "pick up nearest" which, when applied, has consequences which depend on what the nearest object is. Of course, the environmental interpretation of "pick up nearest" is not immediately accessible to the agent who carries out the operation. In fact, for any sort of "pick up" terminal, it consists of the "physics" of the environment. All the individual observes is that "pick up nearest" usually works very well when the nearest thing is a pen, but results in a hernia when the nearest thing is a safe! It is on this basis that the individual may start to organize their perceptions and make inferences. (For example learning not to attempt "pick up nearest" on "big things".) This view of the relationship between agent and environment is described as qualitative physics and is essential to the existence of genuine novelty in the environment.<sup>[65]</sup> In fact, as far as we can tell, the environment is not completely open ended but follows "natural" or "social" laws. The fact that nitric acid mixed with glycerin produces nitro-glycerin was "known" to the physical environment but not to Alfred Nobel, who blew himself up several times finding out! Thus, building open ended evolutionary algorithms is, perhaps paradoxically, a matter of increasing the sophistication of the environment and not of the agent.<sup>[66]</sup>

## 5.17

This observation has an immediate application to the issue of what individuals are supposed to be able to infer or observe about each other. In the first place, it is important to distinguish *behavioral* imitation (where one firm simply imitates the price set by another in the last period for example) from *cognitive* imitation (where one firm is actually able to incorporate some aspect of another firm's decision process into its own). More research is needed in the extent to which firms are capable of making inferences about each other and actually do so.<sup>[67]</sup> On one hand, a great deal of information about a firm is extremely hard to observe, even if it is not actively kept secret. Against this can be set the incentives for success and the considerable skill which humans have at making inferences. For example, one firm may infer that another is testing a new production process secretly by observing that it has started purchasing a certain sort of raw material in small quantities.<sup>[68]</sup> It may make *this* inference by observing the delivery van of a specialist chemical company. In the light of these difficulties, transcription errors are a far more likely explanation of mutation than deliberate experimentation during the process of imitation. It is also clear that the syntactic "interpretability" of GP trees is important not just for the simulator, but also in ensuring that a firm "knows" when it has obtained a coherent piece of "intelligence", even before it can place a value on it. By contrast, GA models have to assume syntax preserving operators like crossover so that firms don't have to address the issue of interpretation. The qualitative physics perspective is relevant both in the process of making inferences as part of imitation - inference involves "filling in the gaps" using correlations that have been acquired from past experience - and in trying to assess who it would be appropriate to imitate. Again, more research could perhaps be done on what knowledge about firms can be regarded as available in the market, but we can be sure that even if the profits accruing to individual strategies are not available, firms will be attempting to build up proxies that define their competitors in terms of age, share dividends, recent expansion or recruitment, firm size and so on.<sup>[69]</sup> All of these proxies can be defined in relative and not absolute terms.

## **5.18**

The final difficulty in interpreting the environment comes in representing what individuals believe about the effects of their actions. Arifovic assumes by omission that firms don't have any beliefs about this and perhaps, since the only option they have is to set their own prices, this is appropriate. However, it may raise a problem with the election operator since this requires a firm to assume market stability in two interdependent respects. Firstly, the profit accruing to a strategy in the previous period must be positively correlated with its profit in the present period. Secondly, the impact which one firm has on the environment must not be so great that the profit accruing to it, had it used the strategy it is now testing for election, would have been completely different to the one it actually observed for the strategy it *did* use. These problems are avoided in the Arifovic models by the extreme simplicity of the environment and the effective "invisible hand" of the fitness function, but in real markets such assumptions of stability might not be appropriate. The same problem occurs with more force in a GP model of the El Farol Bar (Arthur 1994) developed by Edmonds (1997). In this model, each agent is keen to visit the bar only if it is not too crowded. Agents make use of mental models which allow them to decide about whether to go to the bar each night, based on previous experience. The preferences of the agents require the evolution of mental model diversity, since if *everyone* believes the bar will not be too crowded, they will all go and make it too crowded. In Edmonds (1997), despite the fact that agents are evaluating sophisticated GP strategies which take account of the past actions and utterances of other individuals, they perform evaluations on the assumption that what *they* do has no impact on the actions of those other agents! The kind of comparisons which it is appropriate for individuals to make will depend on the particular structure of the model, but it is important to devise comparisons which are both consistent and plausible. Again, the qualitative physics approach suggests that individuals will be attempting to *identify* stable comparisons which can then be used as parts of their decision process. We may chose to simplify the model by allowing only "appropriate" comparisons, on the basis of "external" observations of the model properties, but we should not ignore the issue altogether.

### **Simulations Based On The Population Interpretation**

## **5.19**

Population based models, which attribute a single strategy to each individual are obviously more appropriate to modeling norms or habits rather than sets of strategies which have any rational component in their selection. Many of the issues connected with models based on the population interpretation, such as imitation and the significance of the genotype/phenotype distinction have already been discussed. One obvious area for further research is models which combine evolution at the individual and population levels.<sup>[70]</sup> At their simplest, however, individuals may only consist of a single action such as "co-operate" or "defect".

## **5.20**

It should not be thought that replicator dynamics is the only technique for representing population level models. Apart from the reservations already expressed, it is straightforward to build models which more closely approximate to a coherent social evolutionary process. For

example, we can view the payoff to games in terms of "energy" rather than "utility".<sup>[71]</sup> This is used up at a steady rate and any agent reaching a zero energy level is removed from the population. (This approach is used in the "sugarscape" described by Epstein and Axtell 1996.) Depending on whether there is more to agents than particular game strategies, performance based imitation can take place on the basis of proxies like age and involve behavioral imitation or inference of the genotype with imitation and mutation. It is not yet known whether models of this kind would arrive at different equilibria than the replicator dynamics models and this remains an unexplored field of study.

## Syntactic and Semantic Evaluation

### 5.21

In addition to developing coherent interpretations of the decision processes of individuals and their interactions with the environment, there is also a more general issue of interpretation which needs discussion. This is the relationship between directed and undirected learning and the role of syntactic and semantic analysis of strategies. Syntactic analysis involves being able to say something about the value of a strategy purely on the basis of its syntax. An example was provided by a GP tree which indicates a negative price. This is obviously unsatisfactory and does not need to be tried out in the market. By contrast, semantic analysis involves being able to say something about the "meaning" of a GP tree, for example that the reason why a cost-plus pricing strategy is failing to show a profit is that the definition of the cost element fails to take account of depreciation in the form of intermittent costs for machine replacement. The obvious difficulties with semantic analysis have meant that descriptive GP models have ignored it so far. This has meant that in some sense, learning by alterations to GP strategies has been largely undirected. Although whole strategies are chosen according to past performance, the components which make them up are not selected on this basis, or at best only very indirectly. (Firms must implicitly assume that sub trees of better firms - where better is defined by their criteria on whether or not to imitate - are themselves on average likely to be better, but this is much weaker than being able to identify and imitate only fitter sub trees.) By contrast, directed learning involves being able to evaluate the choices one is presented with in some way and pick the better ones directly.

### 5.22

Although the development of models which include semantic analysis of GP strategies is a major project, one possible direction is indicated by the discussion of qualitative physics in the previous section.<sup>[72]</sup> In the same way that what is lacking in open ended GP models is not more operators and terminals but a richer environment, what is lacking for the firm to perform semantic analysis is not some "special" GP language but adequate *internal* structure to the model of the firm.<sup>[73]</sup> For example, if "costs" are categorized as actions which result in a reduction in the current account, then a comparison of this set with the set of terminals in the current GP strategy will reveal the absence of any terminal connected with machine replacement costs. Semantic analysis of strategies thus presupposes that the operators and terminals possess additional attributes which give them semantic content.

## 5.23

Semantic analysis also raises two other interdependent issues. The first is that figuring out what a GP tree means can also be a taxing problem for the simulator as well as for the individual that is using it. Some of this difficulty can be removed by sensible choice of operators and terminals and avoidance of very large trees.<sup>[74]</sup> (Combinations of operators and terminals which are easy to interpret individually but hard to interpret as a tree may provide *prima facie* evidence against those combinations occurring in real decision processes given our evolved linguistic competence. Mixing logical and arithmetical operators in the same tree provides an obvious example.) However, some difficulty is bound to remain and this results from the fact that GP trees are single entities. They can be contrasted with CS rule bases in which each rule is simple to understand but the combined effect can be very complex, particularly if rules are able to alter the state of the message list so that new rules become eligible to fire. There is no simple resolution for this problem, which is no less than the long running knowledge representation debate from Artificial Intelligence (Brachman and Levesque 1985, Sloman 1985). However, research into ADFs may also provide some suggestions as to how useful sub trees can be identified and propagated more effectively.<sup>[75]</sup>

## Conclusions

### 6.1

This paper has had to draw in a wide range of ideas and examples to present a coherent framework for social evolution. However, to conclude effectively, it is possible to decompose the resulting framework into three constituents: the process of biological evolution, the instantiation of the evolutionary process in an evolutionary algorithm and the implications of both of these constituents for evolutionary models in social science.

### 6.2

On the first count, it matters very much whether one has an effective understanding of evolutionary biology in producing models of social evolution, both in recognizing the parallels between social and biological evolution and, equally importantly, acknowledging their differences. In this paper, I have focused particularly on three aspects of this understanding. Firstly, the importance of a proper understanding of Lamarckian inheritance, stressing that its absence in biology is only a contingent fact rather than a theoretical or logical necessity. Lamarckian inheritance not only fits smoothly into the *social* evolutionary framework but, correctly interpreted, resolves difficulties about the identity of units of selection and the likely speed of evolutionary change. Secondly, there is a need to avoid spurious teleological reasoning. Although it is true that social institutions, including the market, exert some selection pressure on individuals and organizations, it is not at all clear how strong that pressure is and whether it should be seen as congruent with any desirable social goals. Finally, it is important to be clear about the behavioral interpretation of processes of "genetic transfer" essential to evolutionary models, so that agents are assumed to perform actions that they seem likely to be capable of in the real world. There seems little point in moving from rational models of decision to equally implausible rational models of imitation.

### 6.3

On the second count, it is equally important to understand the strengths and weaknesses of evolutionary algorithms and to recognize their many variants. The majority of these have been developed with instrumental objectives in mind so caution must be used in applying them to social processes. This is particularly true of the fitness function which imposes an external teleology on the instrumental GA in a way that may not be appropriate for descriptive models. However, on a more positive note, the quest for rapid optimization techniques has led to many creative uses of the biological analogy in generating variant algorithms which have not yet been fully exploited in models of social evolution. One illustration is provided by the discussion of the GENITOR algorithm, a less well known alternative to the traditional "Holland type" GA, which is argued to be a more behaviorally plausible representation of an evolutionary decision process at least under the mental interpretation. Two other obvious areas of research which have remained unexplored to date are the possibility of inducing reusable subroutines in the decision process using the ADF approach in GP (Koza 1994) and attempts to implement any semantic (rather than purely syntactic) modification processes for decision trees, based on some background knowledge in the agent. It also surprising, given the relative ease of interpretation for single rules compared to GA strings or GP trees, that genuinely social models based on CS have not been more extensively developed. (It is also possible that new developments in the modeling of social evolution will inspire new instrumental approaches.)

### 6.4

On the final count, despite the emphasis on economic models in this paper, it is clear that with certain reservations, evolutionary modeling is appropriate to the other social sciences too. The most important dimension of difference between the disciplines is the degree of deliberation assumed to be involved in social action and the role of others in the decision process. However, it is clear that imitation and mutation processes are if anything even more appropriate for representing mechanisms of social influence or conformity. Since most descriptive models of social evolution dispense with the global fitness function, they can be used equally well for rather economicistic approaches involving a well defined utility function as for more social approaches based on reference points or aspiration levels. It has already been remarked that the more "cognitivist" approach is capable of drawing the different social science disciplines more closely together, at least in principle. This suggests a whole new field of modeling in which attempts are made to *integrate* the insights of the different disciplines within a simulation framework. (One example is provided by Chattoe and Gilbert 1997). Another relevant area of research that has not really been touched on in this paper is the use of *neural networks* to model the evolution of social processes. Like the use of evolutionary algorithms, this research is mainly concentrated in unorthodox economics (Margarita 1992, Beltratti *et al.* 1996), but there are also simulations which can be seen as modeling "low level sociality" and could be assigned to economics, sociology, anthropology or even archaeology (Parisi 1997, Pedone and Parisi 1997). The reason for neglecting these approaches is not that neural networks are an inappropriate way of representing agent competencies, that depends very much on the behavior concerned and on disciplinary preferences for models of conscious deliberate action or instinctive, behaviorist and socially contextual response. Rather, the difficulty is that there is little analysis of how the evolved neural networks do what they do. By contrast, far from being a black box, some

accessible and conscious (though plainly not fully rational) mental process is widely believed to be what makes social action distinctively social. Be this as it may, neural network models draw attention to the fact that social evolution does not need to imply a rational or cognitivist view of agents.

## 6.5

In addition to these individual observations about the three components of the synthesis, an overarching conclusion can tentatively be drawn: that a coherent and complete interpretation of social evolution can at least be sketched. Furthermore, it is possible to move the debate on from the conceptual or methodological level (where some of the misunderstandings originated with contemporaries of Darwin) to the implementation of some concrete models which have been illustrated in this paper. Applications of evolutionary algorithms in new areas can do nothing but sharpen the debate further.

## 6.6

Finally, however, it seems appropriate to conclude a speculative paper such as this not with a definitive conclusion but rather with two unanswered questions, one empirical and one theoretical:

Just how much teleology is there in social organizations like markets? On one hand, it is clear that markets do not enforce profit maximization. On the other, it seems to be constitutive of a competitive market that firms must satisfy the minimum requirement that they cover costs, at least in the long run. Because, at least theoretically, bounded rationality permits us to envisage a world in which everybody has a completely different cognitive model, we can see the considerable social advantages in agreements to set up institutions binding everyone participating in them to relatively similar behavior. What can we say about the status of evolutionary processes taking place within these institutions? The answer has to be empirical rather than ideological. Diversity provides some guidance to the evolutionary pressure being exerted, though with the qualification that firms may respond to this pressure by specializing and deliberately occupying different niches. Simulations which explicitly model the processes by which social structures like markets come to be legislated might provide another fruitful field for future research.

To what extent should evolutionary algorithms (GA strings, CS rule bases and GP trees) even be *expected* to represent the whole process of decision? In particular, should we be thinking about the development of *hybrid* systems in which the processes of identifying and testing regularities to serve as operators and terminals would be carried out in parallel and using different techniques to the process of evolving good strategies based on those operators and terminals? Neural networks are an obvious technology for identifying patterns in data and could be seen to correspond to the innate (and largely unconscious) human ability to recognize patterns.<sup>[76]</sup> GP technology can also be used in the same way. In particular, we can envisage a firm devoting its "Statistics Department" to the task of predicting what other firms (or "the market") will do. If these predictions ever attain an adequate quality they will be incorporated into the strategy of the firm that has developed them. (Of course, this may not work for long if the other firms are doing

the same!) This self prediction technique can also be applied instrumentally and descriptively to the task of understanding GP trees. The task of a subsidiary GP would be to predict the behavior of a complex GP tree, either using the same language or a simplified one. The subsidiary GP is rewarded for the quality of its prediction but heavily penalized for tree depth thus enforcing simplification. Instrumentally, this technique may be added to the automated syntactic simplification traditionally practiced on GP trees, where (\* 1 (...)) is replaced by (...) for example (Koza 1992a). Descriptively we can imagine firms continuously trying to identify predictable sub trees in their strategies using both syntactic and semantic techniques.

## Acknowledgements

I would like to thank Bruce Edmonds and Luigi Marengo for helpful discussions of their models which are described here (Edmonds 1997, Dosi *et al.* 1994). Any errors of interpretation remain my responsibility. My thanks are also due to the editor and two anonymous referees for their helpful comments.

## Notes

<sup>1</sup> This term includes, but is not restricted to, Genetic Algorithms, Evolutionary Strategies, Genetic Programming and Classifier Systems. For further useful discussion see Nissen (1993).

<sup>2</sup> Orthodox is used here in the sense defined by Nelson and Winter (1982).

<sup>3</sup> There are, however, some fascinating exceptions (Bagehot 1887).

<sup>4</sup> The earliest paper satisfying this definition is generally agreed to be Alchian (1950).

<sup>5</sup> Although genetic "fitness" is typically associated with numbers of live offspring, the tendency to reproduce as often as possible may be an effect as much as a cause of evolutionary selection, which merely requires persistence. Although the first species that began to reproduce rapidly would gain an advantage against species which did not, no species could gain any further advantage once rapid reproduction was universal.

<sup>6</sup> This is a simplification. There are actually several processes by which genetic mixing and mutation take place, corresponding to various things which can happen to the chromosomes (Weaver and Hedrick 1992: 87-94). In asexual reproduction there is no genetic mixing. However, asexual reproduction is limited to simple creatures and sexual reproduction can itself be seen as an evolved mechanism by which genetic mixing can take place if it is beneficial.

<sup>7</sup> The validity of the assumption that social behavior is no more than the sum of individual interactions remains contested in social science. Although economics favors a purely individualistic interpretation, sociologists sometimes appear to argue that norms and other social influences are more than the sum of individual mental contents.

<sup>8</sup> We can thus see the ability to make these attributions and cultural (rather than genetic) transmission of behaviors as evolved mechanisms which reduce the wastefulness of the

evolutionary process. Instead of "throwing away" a whole individual, it is possible for that individual to both identify and modify particular aspects of behavior that are persistence threatening. We would expect these features to be both a cause and a consequence of increasing complexity. As individuals become more complex and the potential for diversity increases, it becomes more wasteful to "throw away" whole individuals. At the same time, self awareness and cultural transmission dramatically speed up processes of experimentation in relation to the environment, which were previously restricted to one "try" per lifetime.

<sup>9</sup> Note that *progressive* increases in complexity are not implied. That depends on whether a species with a given level of complexity persists robustly against all simpler organisms and their subsequent developments. So far, humanity appears to be in this position, but only so far.

<sup>10</sup> The traditional example is that the children of hard working blacksmiths might be born with extra muscles. But equally the children of amputees might be born with missing limbs.

<sup>11</sup> Lamarckism arose before genes were discovered. It persisted, rather uneasily, when there still appeared to be a simple mapping between genes and "traits" like blue eyes, but has not survived the discovery of detailed biochemical mechanisms of development.

<sup>12</sup> They may even be meaningless (Tintner 1941a, Tintner 1941b) or incalculable (Thomas 1993).

<sup>13</sup> It will be recalled that the evolutionary process implies nothing about whether competitive or co-operative strategies will tend to persist more effectively (Kropotkin 1939).

<sup>14</sup> Conceptual smoothness merely implies that we do not need to postulate any uncaused causes or mysterious jumps. It is no more than "good science" writ large. Although we will never know what went through the mind of the person who invented the wheel at the crucial moment, we can be confident that it was something to do with what they knew and had experienced, seen or perhaps been told. Although the discovery of fire may have caused an enormous discontinuity in social practices, there is no conceptual discontinuity implied by its discovery and application.

<sup>15</sup> These were also chosen because they are the techniques used in most of the evolutionary modeling done in social science to date. Other related techniques include Evolutionary Programming (Fogel *et al.* 1966, Fogel 1991), Evolution Strategies (Rechenberg 1973, Schwefel 1981) and Simulated Annealing (Davis 1987).

<sup>16</sup> Random generation is compatible with the conceptual smoothness of the evolutionary process in obliging the population of individuals to begin in a state of maximum disorder.

<sup>17</sup> This interpretation has however been challenged by De Jong (1992).

<sup>18</sup> This analogy raises an interesting point. It is *assumed* that the meat yield of cows is a purely genetic matter, just as it is assumed that an instrumental GA process can have a fixed fitness function attributed to it, as it clearly can in the simplest cases. If it turns out that meat yields are affected by the "social" organization of the herd or that the fitnesses of population members actually affect each other, then the instrumental approach is an *empirically* invalid choice.

<sup>19</sup> One reason for this is that the tree representation translates directly into a nested bracket representation suitable for list processing languages such as LISP.

<sup>20</sup> There are GA techniques which make use of variable length representations but they are typically unwieldy. See Harvey (1992) for further discussion.

<sup>21</sup> It is hard to say conclusively that a GP tree is more expressive than a GA string, since any arbitrarily complex encoding for the GA can be hidden away in the fitness function. However, it is intuitive that the GP tree and the fixed "meanings" of its operators *taken together* are likely to be more economical than such an arbitrary encoding.

<sup>22</sup> It is clear that this is a drawback for an instrumental GA which should ideally solve a problem as quickly as possible. However, it may not be a drawback for a GA used descriptively since it corresponds *prima facie* to human behavior in focusing on important issues first. In practice, this analogy does not bear close inspection and other unrealistic aspects of the simple GA, such as the exogenous fitness function, are far more damaging to its use as a descriptive model.

<sup>23</sup> This can either mean that the program is "self executing" or that a process of interpretation and execution is *internal* to the individual. Either interpretation also requires some distinction between the genotype and phenotype.

<sup>24</sup> We can distinguish three interpretations of the robot controller example. In the first, both the environment and the controller are "abstract" and exist purely within the GP. In the second, the controller remains "abstract" but there is also an attempt to simulate a real environment in which the controller will evolve. In the third, the GP is actually a program inside a real robot, operating in a real environment.

<sup>25</sup> Social science seems to give inadequate general attention to differing degrees of the social. Although it rightly has no interest in situations where a single individual interacts purely with a physical environment, it offers little by way of guidance as to how we might expect behavior to change as we move from agents interacting predominantly with an environment to agents predominantly interacting with each other.

<sup>26</sup> An interesting paper using this approach, applied to zoology rather than social science, is Koza *et al.* (1992).

<sup>27</sup> This description follows Nissen (1993).

<sup>28</sup> The structure of individuals could be subsumed straightforwardly into a GP, but not the operations which are carried out on those individuals.

<sup>29</sup> For the technicalities of this process, the reader is referred to Holland *et al.* (1986). Rules which fire and result in a good outcome "share out" the positive feedback, as do those which result in negative feedback. This sharing discourages bloating of the rule base.

<sup>30</sup> A GA which generates diverse populations to solve problems collectively is discussed in Smith *et al.* (1992). It is an interesting question whether the syntax of GP makes diverse populations redundant, impractical or neither.

<sup>31</sup> Although the GP may produce a program that is equivalent to (rule1 AND rule2 AND rule3), it has no endogenous process to ensure hierarchical comprehensibility, so even if the operators are designed for easy interpretation, there is no guarantee that the overall structure will be easy to interpret. Furthermore, as the depth of the GP tree increases, so does the number of equivalent trees. These can be seen as potential drawbacks of GP representations which may be addressed by current research into program modularization through ADFs (Koza 1994).

<sup>32</sup> The prevailing view that bounded rationality involves no more than applying rational principles to the process of cognition seems both incompatible with the bulk of what Simon wrote and incoherent on closer inspection, for precisely the reasons that Simon gives.

<sup>33</sup> In addition to the obvious point that these models are empirically very implausible!

<sup>34</sup> An example is provided by the history of atomic theory, where compounds were explained in terms of molecules, composed of atoms, composed of neutrons, protons and electrons, composed of quarks, composed of ...

<sup>35</sup> There may also be a signal extraction or credit assignment problem when an agent co-varies rules and meta-rules while trying to make sense of the environment.

<sup>36</sup> For a simple illustration, consider an individual who has a gun with a laser sight trained on him by a distant assassin. The assassin is so far away that they cannot be seen directly, but the reflection of the laser sight can. If the reflection can be seen near the target, the rational action is to take cover, otherwise it is to try and move to locate the assassin. Unfortunately, if the assassin trains the laser on some part of the victim's forehead, which is the best way to be sure of a kill, the only way the victim can see it is by holding up a mirror which, we suppose, is of such a shape and size that using it to look for the reflection will block the beam. Even if the blind spot is tiny, the victim will certainly die, despite having a rational decision process, because they have no way of simultaneously or sequentially observing the location of the reflection and the effect which the mirror is having on the beam. Admittedly, this example relies on an assumption about the shape of the mirror, but recall that its object was only to explain why even the smallest blind spot can destroy the possibility for rational action as economics defines it. In fact, one could argue that the shape of the mirror is consistent with defining some part of the forehead as a proper blind spot in the first place. A blind spot is more than somewhere you can't see when not looking in the correct direction! (Even if the victim had a second mirror of the same size and shape, and was very dexterous, they still could not use that mirror to observe the effect the laser was having on the back of the first mirror because it would then be the second mirror blocking the beam.) These ideas are developed further in work on "autopoietic systems" (Varela 1979, 1991).

<sup>37</sup> However, they can be justified in *other* ways, for example as biologically evolved competencies. Unfortunately, this seems rather an admission of defeat as far as social science is concerned.

<sup>38</sup> This raises another interesting issue. Although one agent is hampered by the difficulties of obtaining adequate information about another, it does not suffer from any *conceptually necessary* blind spot in observing other agents. (It is an open question whether the blind spot of the first agent will definitely impair its understanding of the second.) It is truly the case that others may know us better than we know ourselves from a logical point of view!

<sup>39</sup> The same logic applies to any models constructed by the social scientist. This view also implies a coherentist rather than positivist notion of truth.

<sup>40</sup> Note that *common* knowledge of this shared correct model, which enables it to form the basis for rational action, is an even stronger assumption than that agents all merely happen to have the correct model (Parikh 1990).

<sup>41</sup> There is an obvious but very important difference between the assumption that individuals *do* what they intend to do and that what *happens* is what agents intend. The fact that it is possible to miss this difference is illustrated by the argument between Alchian (1950, 1953) and Penrose (1952, 1953).

<sup>42</sup> Although we are concentrating on cognitive irreversibility, it is obvious that irreversibility relating to autonomous physical processes (Georgescu-Roegen 1971) can also be part of the same framework.

<sup>43</sup> As Nelson and Winter (1982) have pointed out, orthodox economic theory struggles with genuine novelty. We will have return to this issue as the same problem appears to beset simple evolutionary algorithms.

<sup>44</sup> An example is provided by the discussion of "big players" in Koppl and Langlois (1994).

<sup>45</sup> Friedman (1953) argues that the assumption that firms are profit maximisers can be justified by the fact that the market will tend to eliminate those firms that are not. This pseudo-Darwinian argument is still widely believed, despite being convincingly refuted by Witt and Perske (1982) and Chiappori (1984) among others. Ironically, the error in Friedman's reasoning is one that originates with Herbert Spencer over a hundred years ago!

<sup>46</sup> Although the abstract theory of the market may imply a universal framework of law standardizing the behavior of firms, a more detailed view reveals a rich structure of compliance, evasion, detection, political action, punishment and resistance. A question for the future is the extent to which more or less generally agreed but not universal practices *can* be said to constitute an external teleology and whether or not different degrees of agreement can be detected in the dynamics of different social processes.

<sup>47</sup> The minimal condition for market persistence is rather similar to that in biological systems, that the organism should "cover costs". Discussion of the widely held but mistaken belief that evolution produces optimal behavior can be found in Hodgson (1991).

<sup>48</sup> There are some extremely interesting and widespread developments in industrial organization which can be considered in evolutionary terms. The first of these is the existence of firms which develop by merger and buy out rather than pure production. The second is the existence of franchises and chains which really do reproduce branches whose operating practices may or may not be appropriate to their locales. These chains have to "trade off" economies resulting from common practices against loss of sales from local idiosyncrasies. They also have to consider the "carrying capacity" of the environment in siting their new branches.

<sup>49</sup> We should not perhaps presume on this issue, economic socialization is still inadequately explored and although long run historical analyses of the nature of the firm do exist, these are largely discursive.

<sup>50</sup> Another interpretation is that if firms know that the encoding is common and that the fitness function is one to one, they can work back from what firms did to what their GA string must be, although this raises issues about timing. Perhaps telepathy is a more behaviorally plausible assumption after all!

<sup>51</sup> If we assume sociable agents then intentions *are* transmissible, but it is not clear how accurate measures of utility could be transmitted, even with the most active desire to communicate. Transmission of information about money amounts is obviously not helpful!

<sup>52</sup> It should be noted that not all game theoretic evolutionary models are based on replicator dynamics.

<sup>53</sup> Other differences between economic and biological games are discussed by Selten (1993).

<sup>54</sup> One could view replicator dynamics as a purely instrumental technique on this basis. Once it is assumed that all individuals are utility maximisers and it is utility which determines propagation, the attainment of game equilibria is a foregone conclusion. If, as many users of replicator dynamics argue, its important contribution is to show *which* equilibrium occurs, then it is no longer possible to argue that behavioral assumptions about the process are irrelevant.

<sup>55</sup> One could diagnose the convergence problems in the Arifovic models knowing nothing at all about its interpretation.

<sup>56</sup> There are a number of models based on GA which avoid one or more of the shortcomings of the Arifovic models, for example Curzon Price (1997), Lomborg (1992) and Vriend (1994). Space considerations preclude their detailed discussion here, but some of the comments on GP models will also raise issues appropriate to those using GA techniques. In what follows, GP terminology will be used for simplicity.

<sup>57</sup> Edmonds (1997) uses a population of 40! Two additional comments can be made on this observation. The first is that because most of the cost of executing an instrumental GP comes in evaluating strategies, the actual coding of the GP involves deliberately *not* evaluating duplicate strategies, but using a hash table to weed out strategies that have already been tried (Koza 1992a). Even using this technique cannot avoid the wastefulness of strategies which are semantically equivalent but syntactically different. The second is that the number of strategies which can be "borne in mind" is far greater in a firm, where these would correspond to the views of particular individuals in the firm. The weights of strategies could then correspond to the number of people arguing for a particular view or their importance in the hierarchy.

<sup>58</sup> Rank based selection has the advantage that selection pressure does not drop as the population converges (Whitley 1989). It is an interesting question whether this also applies in social situations.

<sup>59</sup> This is trivial for a static fitness function, but not for a descriptive simulation in which fitness may change.

<sup>60</sup> This issue is discussed from an instrumental perspective by Reeves (1993).

<sup>61</sup> One interesting possibility is that common knowledge may be mimicked by the fact that individuals tend to project their own models onto others. If, in fact, everybody did have the same model, this would be as good as common knowledge for the purposes of deciding what to do. Of course, it would fail miserably if models differed significantly. Perhaps this is why people have wars over religion rather than food.

<sup>62</sup> In implementing a version of the Dosi *et al.* (1994) model, I followed them in assuming that firms selected strategies probabilistically on the basis of cumulative profits. When the strategies simply consisted of fixed prices rather than GP trees, all the firms quickly learnt to set the maximum price they could and rapidly priced themselves out of the market. Although there was an equilibrium with everyone charging the maximum permissible price, it was never observed because it required that all firms initially charged that price and did not deviate. This is an example where the fitness function for firms was not a complete representation of the restrictions on the market and thus prevented any firms from persisting. In this case, firms valued positive profit without limit but placed no value on market share at all.

<sup>63</sup> It is also the case that retention of profits *does* make firms far less susceptible to whatever discipline the market imposes and also perhaps more able to influence the terms on which market discipline is applied. This point seems to receive inadequate attention from evolutionary market apologists. Since this escape from discipline reflects the very success of firms, perhaps it is true as Marx suggested that competition, if not capitalism contains the seeds of its own destruction!

<sup>64</sup> They can be assigned the fitness of their parents providing that firms do not have an infinite memory for cumulated profit.

<sup>65</sup> The relevance of qualitative physics is illustrated in a paper by Sims (1991). This describes a simulation in which artificial organisms constructed of rigid blocks and joints are "evolved" to

perform simple tasks like "capturing" food. Rather than seeing the properties of joints and blocks as attributes of the organism, as classical Artificial Intelligence might, these attributes are modeled as functions of environmental factors such as "gravity" and "mass". The result is that the behavior of arbitrary evolved combinations of blocks and joints is always defined.

<sup>66</sup> This conclusion is also reached in the Artificial Life literature (Langton 1989, Langton *et al.* 1991, Langton 1993). Complex environments have another interesting effect. Because evolutionary algorithms are very good at optimization, they can often produce strategies which, in being self-evidently silly, reveal something about the inadequacy of the environmental specification. Developing the simulation thus becomes a co-evolutionary process, with behaviorally implausible strategies sometimes revealing unrealistic assumptions about the agent and sometimes about the environment.

<sup>67</sup> This distinction is often badly muddled in replicator dynamics models. If individuals just are their strategies then there is no difficulty, except that this is a very unrealistic view of individuals. If there is a difference between the strategy and the actions it produces, then strong knowledge assumptions are required to avoid worrying about how inferences can be made from actions back to strategies.

<sup>68</sup> In one firm where I worked, employee research group membership was not listed in the internal phone directory so nobody outside the company could get an overall picture of the amount of research being done in different areas!

<sup>69</sup> In Chattoe and Gilbert (1997), agents learn how to budget by evolving budgeting plans individualistically, but they are also able to observe the consumption patterns of other agents. The result is a co-evolution of effective budgeting plans and stratified lifestyles based on income.

<sup>70</sup> There is an analogy here with the highly instrumentally effective Island Model class of Genetic Algorithms in which relatively small populations evolve in parallel, but transmit their best strategies at random to other populations from time to time (Gordon and Whitley 1993).

<sup>71</sup> Another way of seeing the issue of profits reducing the competitive pressure on firms is to ask whether profit tends to "evaporate" or not. Individuals can never achieve more than a certain level of energy or wakefulness and this continues to drain away whatever they do.

<sup>72</sup> This program is proceeding for instrumental GP in the development of ADF techniques (Koza 1994) which may provide interesting insights which descriptive models can use.

<sup>73</sup> There is an obvious role here for the sort of computational organization theory models devised by Carley and others (Carley and Prietula 1994).

<sup>74</sup> Large trees are behaviorally implausible as well as almost impossible to interpret. An additional danger, suggested by the Dosi *et al.* results on price tracking is that a large GP tree may just become a lookup table for the state of the present environment. There has been little research so far on pulling GP trees that are apparently successful out of one environment and putting them in another.

<sup>75</sup> Another consequence of semantic analysis is the possibility of analogical reasoning. If a sub tree has a "meaning" attached to it, like "total costs" it becomes possible to substitute one sub tree for another in a way that is somewhat directed, as it relies on similarities of meaning. This also applies at the level of whole GP trees, which may only produce a number as output, but it is the use to which the organization puts that number which determines its meaning. Thus a tree that solves one environmental problem may be more likely to solve an analogous problem.

<sup>76</sup> Work of this kind could make use of research such as that by Sun and Bookman (1993) on the integration of neural and symbolic processing.

## References

- ALCHIAN, Armen A. (1950) 'Uncertainty, Evolution and Economic Theory', *Journal of Political Economy*, 58(3), June, pp. 211-222.
- ALCHIAN, Armen A. (1953) 'Comment: Biological Analogies in the Theory of the Firm', *American Economic Review*, 43(3), Part 1, September, pp. 600-603.
- ARIFOVIC, Jasmina (1990) 'Learning by Genetic Algorithms in Economic Environments', Working Paper 90-001, Economics Research Program, Santa Fe Institute, New Mexico, October.
- ARIFOVIC, Jasmina (1994) 'Genetic Algorithm Learning and the Cobweb Model', *Journal of Economic Dynamics and Control*, 18(1), Special Issue on Computer Science and Economics, January, pp. 3-28.
- ARIFOVIC, Jasmina (1995) 'Genetic Algorithms and Inflationary Economies', *Journal of Monetary Economics*, 36(1), August, pp. 219-243.
- ARIFOVIC, Jasmina (1996) 'The Behavior of the Exchange Rate in the Genetic Algorithm and Experimental Economies', *Journal of Political Economy*, 104(3), June, pp. 510-541.
- ARIFOVIC, Jasmina and EATON, Curtis (1995) 'Coordination via Genetic Learning', *Computational Economics*, 8, pp. 181-203.
- ARTHUR, Brian (1994) 'Inductive Reasoning and Bounded Rationality', *American Economic Association Papers*, 84, pp. 406-411.
- BAGEHOT, Walter (1887) *Physics and Politics or Thoughts on the Application of the Principles of 'Natural Selection' and 'Inheritance' to Political Society*, eighth edition (London: Kegan Paul, Trench and Company).
- BAUMOL, William J. and QUANDT, Richard E. (1964) 'Rules of Thumb and Optimally Imperfect Decisions', *American Economic Review*, 64(2), Part 1, March, pp. 23-46.
- BELTRATTI, A., MARGARITA, S. and TERNA, P. (1996) *Neural Networks for Economic and Financial Modelling* (London: International Thompson Computer Press).

BRACHMAN, R. and LEVESQUE, Hector J. (eds.) (1985) *Readings in Knowledge Representation* (Los Altos, CA: Morgan Kaufmann).

BULLARD, James and DUFFY, John (1994) 'A Model of Learning and Emulation with Artificial Adaptive Agents', Draft Paper, Research Department, Federal Reserve Bank of St Louis, MO and Department of Economics, University of Pittsburgh, May.

CARLEY, Kathleen M. and PRIETULA, Michael (eds.) (1994) *Computational Organization Theory* (Hillsdale, NJ: Lawrence Erlbaum Associates).

CHATTOE, Edmund (1994) 'Some Thoughts on the Methodology of Computational Economics', paper presented at the conference *Computational Economics*, Amsterdam, 8-11 June.

CHATTOE, Edmund (1996) 'Why Are We Simulating Anyway? Some Answers From Economics', in Troitzsch, Klaus G., Mueller, Ulrich, Gilbert, Nigel and Doran, Jim E. (eds.) *Social Science Microsimulation* (Berlin: Springer-Verlag), pp. 78-104.

CHATTOE, Edmund and GILBERT, Nigel (1996) 'The Simulation of Budgetary Decision Making and Mechanisms of Social Evolution', paper presented at ISA '96: Fourth International Social Science Methodology Conference, Colchester, Essex.

CHATTOE, Edmund and GILBERT, Nigel (1997) 'A Simulation of Adaptation Mechanisms in Budgetary Decision Making', in Conte, Rosaria, Hegselmann, Rainer and Terna, Pietro (eds.) *Simulating Social Phenomena* (Berlin: Springer-Verlag), pp. 401-418.

CHIAPPORI, Pierre-Andre (1984) 'Sélection Naturelle et Rationalité Absolue des Entreprises', *Revue Economique*, 35(1), January, pp. 87-108.

CURZON PRICE, Tony (1997) 'Using Coevolutionary Programming to Simulate Strategic Behaviour in Markets', *Journal of Evolutionary Economics*, 7, pp. 219-254.

DARWIN, Charles (1968) *On The Origin of Species by Means of Natural Selection, or the Preservation of Favoured Races in the Struggle for Life* (Harmondsworth: Penguin).

DAVIS, Lawrence D. (ed.) (1987) *Genetic Algorithms and Simulated Annealing* (London: Pitman).

DAWKINS, Richard (1989) 'The Evolution of Evolvability' in Langton, Christopher G. (ed.) *Artificial Life: The Proceedings of an Interdisciplinary Workshop on the Synthesis and Simulation of Living Systems, Los Alamos, New Mexico, September 1987*, SFI Studies in the Sciences of Complexity Volume VI (Redwood City, CA: Addison-Wesley), pp. 201-220.

DE JONG, Kenneth (1992) 'Are Genetic Algorithms Optimizers?', in Männer, Reinhard and Manderick, Bernard (eds.) *Parallel Problem Solving from Nature, 2, Proceedings of the Second Conference, Brussels, Belgium, 28-30 September 1992* (Amsterdam: North-Holland), pp. 3-13.

DORAN, Jim and GILBERT, Nigel (1994) 'Simulating Societies: An Introduction', in Gilbert, Nigel and Doran, Jim (eds.) *Simulating Societies: The Computer Simulation of Social Phenomena* (London: UCL Press), pp. 1-18.

DOSI, Giovanni, MARENKO, Luigi and VALENTE, Marco (1994) 'Norms as Emergent Properties of Adaptive Learning: The Case of Economic Routines', paper presented at the Royal Economic Society Conference, Exeter, 28-31 March.

EDEY, Maitland A. and JOHANSON, Donald C. (1989) *Blueprints: Solving the Mystery of Evolution* (Oxford: Oxford University Press).

EDMONDS, Bruce (1997) 'Gossip, Sexual Recombination and the El Farol Bar: Modelling the Emergence of Heterogeneity', Centre for Policy Modelling Discussion Paper Number CPM-97-31, Manchester Metropolitan University, <http://www.cpm.mmu.ac.uk/cpmrep31.html>.

EDMONDS, Bruce and MOSS, Scott (1996) 'Modelling Bounded Rationality using Evolutionary Techniques', Centre for Policy Modelling Discussion Paper Number CPM-96-10, Manchester Metropolitan University, <http://www.cpm.mmu.ac.uk/cpmrep10.html>.

EPSTEIN, J. M. and AXTELL, R. (1996) *Growing Artificial Societies: Social Science from the Bottom Up* (Cambridge, MA: Brookings Institution/MIT Press).

FISHER, R. A. (1930) *The Genetical Theory of Natural Selection* (Oxford: Clarendon Press).

FOGEL, L. J., OWENS, A. J. and WALSH, M. J. (1966) *Artificial Intelligence Through Simulated Evolution* (New York, NY: John Wiley and Sons).

FOGEL, D. B. (1991) 'The Evolution of Intelligent Decision Making in Gaming', *Cybernetics and Systems*, 22, pp. 223-236.

FRIEDMAN, Milton (1953) 'The Methodology of Positive Economics', in *Essays in Positive Economics* (Chicago, IL: University of Chicago Press), pp. 4-14.

GEORGESCU-ROEGEN, Nicholas (1971) *The Entropy Law and the Economic Process* (Cambridge, MA: Harvard University Press).

GOLDBERG, David E. (1989) *Genetic Algorithms in Search, Optimization and Machine Learning* (Reading, MA: Addison-Wesley).

GORDON, V. Scott and WHITLEY, Darrell (1993) 'Serial and Parallel Genetic Algorithms as Function Optimizers', in Forrest, Stephanie (ed.) *Proceedings of the Fifth International Conference on Genetic Algorithms, University of Illinois at Urbana-Champaign, 17-23 July 1993* (San Mateo, CA: Morgan Kaufmann), pp. 177-183.

HARVEY, Inman (1992) 'The SAGA Cross: The Mechanics of Recombination for Species for Variable Length Genotypes', in Männer, Reinhard and Manderick, Bernard (eds.) *Parallel*

*Problem Solving from Nature, 2, Proceedings of the Second Conference, Brussels, Belgium, 28-30 September 1992* (Amsterdam: North-Holland), pp. 269-278.

HODGSON, Geoffrey M. (1991) 'Economic Evolution: Intervention Contra Pangloss', *Journal of Economic Issues*, 25(2), June, pp. 519-533.

HODGSON, Geoffrey M., SAMUELS, W. J. and TOOL, Marc R. (eds.) (1994) *The Elgar Companion to Institutional and Evolutionary Economics* (Aldershot: Edward Elgar).

HOFSTADTER, Richard (1955) *Social Darwinism in American Thought* (Boston, MA: Beacon Press).

HOLLAND, John H. (1975) *Adaptation in Natural and Artificial Systems: An Introductory Analysis with Applications to Biology, Control and Artificial Intelligence* (Ann Arbor, MI: University of Michigan Press).

HOLLAND, John H. (1992) *Adaptation in Natural and Artificial Systems: An Introductory Analysis with Applications to Biology, Control, and Artificial Intelligence*, second edition (Cambridge, MA: The MIT Press/A Bradford Book).

HOLLAND, John H., HOLYOAK, K. J., NISBETT, R. E. and THAGARD, P. R. (1986) *Induction: Processes of Inference, Learning, and Discovery* (Cambridge, MA: MIT Press).

JONES, Greta (1980) *Social Darwinism and English Thought: The Interaction Between Biological and Social Theory* (Brighton: Harvester).

KOPPL, Roger and LANGLOIS, Richard N. (1994) 'When Do Ideas Matter? A Study in the Natural Selection of Social Games', *Advances in Austrian Economics*, 1, pp. 81-104.

KOZA, John R. (1991) 'Evolving a Computer Program to Generate Random Numbers Using the Genetic Programming Paradigm', in Belew, Richard K. and Booker, Lashon B. (eds.) *Proceedings of the Fourth International Conference on Genetic Algorithms, UCSD, San Diego 13-16 July 1991* (San Mateo, CA: Morgan Kaufmann), pp. 37-44.

KOZA, John R. (1992a) *Genetic Programming: On the Programming of Computers by Means of Natural Selection and Genetics* (Cambridge, MA: MIT Press/A Bradford Book).

KOZA, John R. (1992b) 'A Genetic Approach to Econometric Modelling', in Bourgine, Paul and Walliser, Bernard (eds.) *Economics and Cognitive Science, Selected Papers from the Second International Conference on Economics and Artificial Intelligence, Paris, 4-6 July 1990* (Oxford: Pergamon), pp. 57-75.

KOZA, John R. (1992c) 'Evolution and Coevolution of Computer Programs to Control Independently Acting Agents', in Meyer, Jean-Arcady and Wilson, Stewart W. (eds.) *From Animals to Animats, Proceedings of the First International Conference on Simulation of*

*Adaptive Behaviour (SAB 91), Paris, 24-28 September 1991* (Cambridge, MA: The MIT Press/A Bradford Book), pp. 366-375.

KOZA, John R. (1994) *Genetic Programming II: Automatic Discovery of Reusable Programs* (Cambridge, MA: The MIT Press/A Bradford Book).

KOZA, John R., RICE, James P. and ROUGHGARDEN, J. (1992) 'Evolution of Food Foraging Strategies for the Caribbean *Anolis* Lizard Using Genetic Programming', Working Paper 92-06-028, Santa Fe Institute, New Mexico.

KROPOTKIN, Prince Petr (1939) *Mutual Aid: A Factor in Evolution* (Harmondsworth: Penguin).

LANGTON, Christopher G. (ed.) (1989) *Artificial Life: The Proceedings of an Interdisciplinary Workshop on the Synthesis and Simulation of Living Systems, Los Alamos, New Mexico, September 1987*, SFI Studies in the Sciences of Complexity, Proceedings Volume VI (Redwood City, CA: Addison-Wesley).

LANGTON, Langton, Christopher G., TAYLOR, Charles, FARMER, J. Doyne and RASSMUSSEN, Steen (eds.) (1991) *Artificial Life II: Proceedings of the Workshop on Artificial Life, Santa Fe, New Mexico, February 1990*, SFI Studies in the Sciences of Complexity, Proceedings Volume X (Redwood City, CA: Addison-Wesley).

LANGTON, Langton, Christopher G. (ed.) (1993) *Artificial Life III: Proceedings of the Workshop on Artificial Life, Santa Fe, New Mexico, June 1992*, SFI Studies in the Sciences of Complexity, Proceedings Volume XVII (Redwood City, CA: Addison-Wesley).

LOMBORG, Bjorn (1992) 'Cooperation in the Iterated Prisoner's Dilemma', Papers on Economics and Evolution, Number 9302, edited by the European Study Group for Evolutionary Economics, Max Planck Institute, Evolutionary Economics Unit, Jena, Germany.

MARGARITA, Sergio (1992) 'Genetic Neural Networks for Financial Markets: Some Results', in Neumann, Bernd (ed.) *Proceedings of the 10th European Conference on Artificial Intelligence, Vienna, Austria, 3-7 August 1992* (New York, NY: John Wiley and Sons), pp. 211-213.

MCCAIN, Roger A. (1992) *A Framework for Cognitive Economics* (Westport, CT: Praeger).

MCDougall, W. (1927) 'An Experiment for Testing the Hypothesis of Lamarck', *British Journal of Psychology*, 17, pp. 267-304.

MORAVEC, Hans (1989) 'Human Culture: A Genetic Takeover Underway', in Langton, Christopher G. (ed.) *Artificial Life: The Proceedings of an Interdisciplinary Workshop on the Synthesis and Simulation of Living Systems, Los Alamos, New Mexico, September 1987*, SFI Studies in the Sciences of Complexity Volume VI (Redwood City, CA: Addison-Wesley), pp. 167-199.

MOSS, Scott and EDMONDS, Bruce (1994) 'Modelling Learning as Modelling', CPM Report 3, Centre for Policy Modelling, Manchester Metropolitan University,  
<http://www.cpm.mmu.ac.uk/cpmrep03.html>.

NELSON, Richard R. and WINTER, Sidney G. (1982) *An Evolutionary Theory of Economic Change* (Cambridge, MA: Belknap/Harvard University Press).

NISSEN, Volker (1993) 'Evolutionary Algorithms in Management Science - An Overview and List of References', Papers on Economics and Evolution, Number 9303, edited by the European Study Group for Evolutionary Economics, Max Planck Institute, Evolutionary Economics Unit, Jena, Germany.

PARIKH, Rohit (1990) 'Recent Issues in Reasoning about Knowledge', in Parikh, Rohit (ed.) *Theoretical Aspects of Reasoning about Knowledge: Proceedings of the Third Conference (TARK 3)* (San Mateo, CA: Morgan Kaufmann), pp. 3-9.

PARISI, D. (1997) 'What to Do With a Surplus', in Conte, Rosaria, Hegselmann, Rainer and Terna, Pietro (eds.) *Simulating Social Phenomena* (Berlin: Springer-Verlag), pp. 133-151.

PEDONE, R. and PARISI, D. (1997) 'In What Kind of Social Groups can "Altruistic" Behaviors Evolve?', in Conte, Rosaria, Hegselmann, Rainer and Terna, Pietro (eds.) *Simulating Social Phenomena* (Berlin: Springer-Verlag), pp. 195-201.

PENROSE, Elizabeth Tilton (1952) 'Biological Analogies in the Theory of the Firm', *American Economic Review*, 42(4), December, pp. 804-819.

PENROSE, Elizabeth Tilton (1953) 'Rejoinder: Biological Analogies in the Theory of The Firm', *American Economic Review*, 43(3), Part 1, September, pp. 603-607.

PETR, Jerry L. (1982) 'Economic Evolution and Economic Policy: Is Reaganomics a Sustainable Force?', *Journal of Economic Issues*, 16(4), December, pp. 1005-1012.

RECHENBERG, I. (1973) *Evolutionsstrategie: Optimierung Technischer Systeme nach Prinzipien der Biologischen Evolution* (Stuttgart: Frommann-Holzboog).

REEVES, Colin R. (1993) 'Using Genetic Algorithms with Small Populations', in Forrest, Stephanie (ed.) *Proceedings of the Fifth International Conference on Genetic Algorithms, University of Illinois at Urbana-Champaign, 17-23 July 1993* (San Mateo, CA: Morgan Kaufmann), pp. 92-99.

RIDLEY, Mark (1985) *The Problems of Evolution* (Oxford: Oxford University Press).

RUNCIMAN, W. G. (1998) 'The Selectionist Paradigm and Its Implications for Sociology', *Sociology*, 32(1), February, pp. 163-188.

SCHWEFEL, Hans-Paul (1981) *Numerical Optimization of Computer Models* (New York: John Wiley and Sons).

SELTEN, Reinhard (1993) 'Evolution, Learning, and Economic Behaviour', *Games and Economic Behaviour*, 3(1), February, pp. 3-24.

SIMON, Herbert A. (1972) 'Models of Bounded Rationality', in McGuire, C. B. and Radner, Roy (eds.) *Decision and Organization: A Volume in Honour of Jacob Marshack* (Amsterdam: North-Holland), pp. 161-176.

SIMON, Herbert A. (1976) 'From Substantive to Procedural Rationality', in Latsis, Spiro J. (ed.) (1976) *Method and Appraisal in Economics* (Cambridge: Cambridge University Press), pp. 129-148.

SIMON, Herbert A. (1978) 'Rationality as a Process and as a Product of Thought', *American Economic Review*, 68(2), May, pp. 1-16.

SIMON, Herbert A. (1981) *Models of Bounded Rationality*, two volumes (Cambridge, MA: MIT Press).

SIMS, Karl (1991) 'Interactive Evolution for Computer Graphics', *Computer Graphics*,

SIGGRAPH Conference Proceedings, 25(4), pp. 343-350.

SLOMAN, Aaron (1985) 'Why We Need Many Knowledge Representation Formalisms', CSRP 052, Cognitive Science Research Paper, University of Sussex.

SMITH, Robert Elliot, FORREST, Stephanie and PERELSON, Alan S. (1992) 'Searching for Diverse Cooperative Populations with Genetic Algorithms', TCGA Report Number 92002, The Clearinghouse for Genetic Algorithms, Department of Engineering Mechanics, University of Alabama (Tuscaloosa), 19 March.

SUN, Ron and BOOKMAN, L. (eds.) (1993) *Computational Architectures Integrating Neural and Symbolic Processing* (Boston, MA: Kluwer Academic Publishers).

THOMAS, J. (1993) 'Non-Computable Rational Expectations Equilibria', *Mathematical Social Science*, 25, pp. 133-142.

TINTNER, Gerhard (1941a) 'The Theory of Choice Under Subjective Risk and Uncertainty', *Econometrica*, 9(3-4), July-October, pp. 298-304.

TINTNER, Gerhard (1941b) 'The Pure Theory of Production Under Technological Risk and Uncertainty', *Econometrica*, 9(3-4), July-October, pp. 305-312.

VARELA, Francisco J. (1979) *Principles of Biological Autonomy* (New York, NY: Elsevier North Holland).

VARELA, Francisco J., THOMPSON, Evan and ROSCH, Eleanor (1991) *The Embodied Mind: Cognitive Science and Human Experience* (Cambridge, MA: MIT Press).

VEGA-REDONDO, Fernando (1996) *Evolution, Games, and Economic Behaviour* (Oxford: Oxford University Press).

VRIEND, Nicolaas J. (1994) 'Self-Organized Markets in a Decentralized Economy', Working Paper 94-03-013, Economics Research Program, Santa Fe Institute, New Mexico.

WEAVER, Robert F. and HEDRICK, Philip W. (1992) *Genetics*, second edition (Dubuque, IA: William C. Brown).

WEIBULL, Jörgen W. (1995) *Evolutionary Game Theory* (Cambridge, MA: MIT Press).

WILSON, Edward O. (1975) *Sociobiology: The New Synthesis* (Cambridge, MA: Harvard University Press).

WILSON, Edward O. (1978) *On Human Nature* (Cambridge, MA: Harvard University Press).

WINSTON, Gordon C. (1987) 'Activity Choice: A New Approach to Economic Behaviour', *Journal of Economic Behaviour and Organization*, 8(4), December, pp. 567-585.

WITT, Ulrich and PERSKE, Joachim (1982) *SMS - A Program Package for Simulating and Gaming of Stochastic Market Processes and Learning Behaviour*, Lecture Notes in Economics and Mathematical Systems, Number 202 (Berlin: Springer-Verlag).

WHITLEY, Darrell (1989) 'The GENITOR Algorithm and Selection Pressure: Why Rank-Based Allocation of Reproductive Trials is Best', in Schaffer, J. David (ed.) *The Proceedings of the Third International Conference on Genetic Algorithms, George Mason University, 4-7 June 1989* (San Mateo, CA: Morgan Kaufmann), pp. 116-121.

WITTGENSTEIN, Ludwig (1953) *Philosophical Investigations* (Oxford: Basil Blackwell).

# **VIRTUAL URBAN LEGENDS: INVESTIGATING THE ECOLOGY OF THE WORLD WIDE WEB**

**By**  
**Edmund Chattoe**

International Conference: 25-27 March 1998, Bristol, UK  
IRISS '98: Conference Papers

## **Introduction**

This paper traces the evolution of several ecological phenomena connected with the spread of replicating messages on the World Wide Web (web) and Internet (net). The paper begins by discussing some basic ecological and evolutionary concepts and then turns to analyzing common features of replicating messages which have adapted them to survive in these particular 'ecologies'. Specifically, replicating messages survive best when they take advantage of the altruistic and 'information sharing' ethos of the virtual world, when they deal with issues of universal interest (most commonly the maintenance of the web and net themselves), when there is a steady influx of non immunized transmitters, and when it is difficult to falsify the content of the message definitively. (Certain very successful replicating messages share features of urban legends as suggested by Woolgar and Russell in their analysis of computer virus stories. In particular, a strong belief that some of these stories 'ought' to be true dramatically decreases the effectiveness of resources like hoax warning web pages.) These features can be illustrated by specific reference to the developments of the replicator known variously as xxx-1, Join the Crew, Win a Holiday or Good News. Good News is interesting because it is the warning and not the (probably) fictitious virus which constitutes the replicator. The paper concludes by drawing out the more general implications of studies of replicating messages. Firstly, the dialogues about these messages which occur on newsgroups are a very interesting source of secondary data about the transmission process which can be considered from the perspective of Conversation Analysis and Social Construction. Secondly, it may be appropriate to consider web phenomena like replicating messages from an 'ecological' perspective. Thirdly, the web serves as an extremely useful research tool for the study of replicators. (It may help to operationalise research on the spread of the memes suggested by Dawkins.) The paper concludes with some thoughts on the design and ethics of a 'tracer' or replicator deliberately released for research purposes.

## **Evolution, Memes and Messages**

The genetic theory of evolution by natural selection (Darwin 1968, Fisher 1930, Edey and Johanson 1989) has three important elements for the purposes of this paper:

**The Distinction between Genotype and Phenotype:** We consider a population of individuals, each of which consists of a *genotype* and a *phenotype*. The genotype is a set of 'instructions' for constructing the phenotype. In biology, the genotype is the DNA which organizes chemical reactions to produce enzymes and other chemical compounds which in turn organize the division of cells and construction of the body during gestation.

**Mechanisms of Genetic Variation and Propagation:** The genotype is transmitted through reproduction. Certain mechanisms ensure that some genotypic traits are propagated relatively intact between parents and offspring - for example eye or hair color in families. Other mechanisms induce mixing of genotypes - for example children having aspects of **both** parents in their appearance - or the introduction of the generation of novelty through mutation as can be observed in the breeding of new roses or pedigree cats and dogs. Mixing of these two kinds of mechanisms - propagation and variation - is essential to continuing evolutionary success. Too little genetic novelty and variation can lead to inbreeding, observed in isolated communities, while too much, in the form of birth defects, results from such harmful environmental effects as excessive radiation. At any time, in a population of individuals, there is always a certain amount of genetic diversity leading to phenotypic diversity.

**Processes of Environmental Selection:** Because the genotype constructs the phenotype, variations in the genotype lead to variations in the phenotype. For example, a mutation in the genes may lead to an individual with an abnormally long neck. This may be helpful or harmful to the persistence of the individual and this in turn to its ability to reproduce others of its kind. In the case of a longer neck, we might expect a grazing animal to reach food not available to others of the same species. Over time, phenotypic aspects which are helpful to individuals will persist and spread, while those which are not will tend to be removed from the population. This process is caused by the environment. It does not reflect any deliberation by the creatures concerned.

As Dawkins (1976, 1983) has pointed out, this theory has two interesting implications: The first is that, rather than seeing the genes as mere passive carriers of genetic information, we can also regard them as 'selfish' puppet masters, building our bodies for the sole purpose of propagating themselves. Obviously, this rhetorical device must be treated with a certain amount of caution, the genes do not 'decide' or 'intend' to do anything in any accepted sense. Nonetheless, it can be useful to reverse the usual perspective and consider bodies as a mere by-product of genes rather than the other way around. (This reversal is also encountered in two everyday jokes: 'Mankind is water's way of getting from place to place' and 'Computers will never try to wipe out mankind. We are their sex organs.') This shift in perspective gives rise to one particular insight, which is that evolutionary biologists have tended to see certain processes as *causes* of natural selection when they can perhaps more usefully be seen as *effects*. Darwin deduced the 'struggle for existence' from the observed geometric increase in populations which are not resource limited. However, from a genetic point of view, the inclination to reproduce as frequently as possible might itself evolve as a way of reducing the species effects of individual death and securing a greater proportion of resources from other species which did not have this trait. Dawkins argues that the minimal condition for evolution is just *persistence* and that all this requires in turn is *replication* or *self-reproduction* when individuals have a finite life span. Thus, we see the 'replicators' proposed by Dawkins, as any objects which are 'able' to copy themselves or 'get themselves' copied in such a way as to persist in a particular environment.

The second implication is that processes involving replicators need not be restricted to biological systems. Dawkins coins the term 'meme' for a *social* replicator, the examples he uses being tunes or jokes. A 'successful' replicator in this context is a tune or joke which can 'get itself' transmitted continuously. An 'unsuccessful' one is not passed on and is forgotten. These ideas also arise in a number of other contexts. Lynch (1996) provides extensive discussion of belief transmission

processes of this kind. More general application of ideas of this kind have also given rise to the discipline of Memetics. Brunvand (1984) has collected many examples of 'urban legends' which, it is argued, propagate not because they are 'true' or necessarily 'useful' but because they identify our archetypal fears and concerns in such an effective way that they tend to be passed on readily. (From the selfish replicator perspective, this gives a whole new meaning to 'playing on our fears'.) Although much work is done in presenting these and other extraordinary narratives as 'factual' (Wooffitt 1992), the transmission of urban legends can be distinguished from related processes which are more generally treated *as* factual, such as the propagation of 'factoids' (such as The Legend of Tailor Pink, Petronius Arbiter, Time Traveler and Monkeys and Typewriters.) in everyday conversation or the persistence of erroneous 'standard interpretations' in academic literature (These distinctions themselves suggest another interesting analogy with evolution in biology. There is no rigid distinction between many 'observably different' species at the genotypic level. Instead, species are defined functionally by their ability to interbreed. Although it is straightforward to distinguish a 'factoid' from an 'urban legend' in terms of the function it serves in communication, the beliefs that it supports and the work that is done to warrant it, from the replicator perspective the two kinds of messages are functionally equivalent. What separates them is the inability to warrant an urban legend in the same way as one does an academic interpretation.)

In the next section, we turn to the possibilities for applying these ideas to the study of communication on the web and net.

## Propagation Processes in the Virtual World

The mechanisms by which DNA replicates during cell division and reproduction provide both stability and variation in subsequent copies of the original genetic material. Stability is provided by the mechanical process of transcription which ensures that copies of DNA strings are typically very close to the originals. By contrast, strings of genetic material can become physically 'tangled' with the result that the 'tails' of two strings become attached to different 'heads'. These processes have analogues in the transmission of web and net information. The possibilities of 'cut and paste', forwarding and reply of emails, provide obvious opportunities for highly accurate transcription. At the same time, the transmitter has the ability to add, alter or delete information without this necessarily being visible to subsequent recipients. This alteration may be with deliberate intent to alter the content of the message or this may arise inadvertently, as when, in altering the formatting, a section of the message is unintentionally deleted. (If the reader is unconvinced they should do an Alta Vista search for a piece of text like 'infinite binary loop' from Text 1 below. The number of variations is surprising.) The same is true of information provided on web pages. (However, the situation is slightly different here, since if the person who copies the 'message' has no interest in modifying it or claiming it as their own, they may simply provide a link to the original page.)

We can thus distinguish between different kinds or levels of replication. Whole messages may simply be forwarded or transmitted verbatim. Particular parts of a message may be separated from the original and continue to propagate, for example in an exchange of emails forming a 'conversation'. (This process can also be seen in the way that quotations 'come adrift' from the larger works which produced them. In time, these quotations may take on a 'life of their own'

isolated from the qualifications of context.) Finally, a message may be propagated in a way that continues its 'tone' or 'subject' matter, while not necessarily retaining too much of the original text. (The obvious example of this is the children's game of Chinese Whispers, in which children pass a whispered message in a circle for the fun of hearing it strangely transformed.) These processes have historical analogues in the study of other kinds of messages and it is interesting to see how the mechanical nature of the transmission process affects the balance of propagation and variation. Hand written manuscripts of any length can often be dated fairly accurately, at least in relative terms, by studying interpolations, idiosyncrasies of the copyist, responses to damage in the original, and so on. By contrast, there is little variability in the text between copies of modern printed books, at least for the same edition. Things like photocopying and scanning are at an intermediate level of variability. It is possible, but tiresome, to produce composite images which can then be transmitted. Entertaining computer applications such as the 'Dancing Baby' are widely transmitted, but with no variation at all.

From the research point of view, there are several reasons why studies of replicators on the web and net are particularly appropriate:

**Volume of Traffic and Speed of Transmission:** Although it is possible to study hand written manuscripts or the ways that stories or historical events change their published form over time, data of this kind is very 'thin' in both spatial and temporal terms and the records, as well as the messages they contain, may be susceptible to selection mechanisms. By subscribing to lists, newsgroups, archives and so on, researchers may expose themselves to as large a volume of messages as they can handle. In addition, because of the rapidity of transmission, experiments on the spread of new ideas can be done over a realistic time scale

**Compatibility and Anonymity:** One of the defining features of both web and net is their high levels of compatibility in message transmission. If one wished to study the transmission of photocopies, it would be necessary to seek out and become a part of some social group in which photocopies were routinely made, such as a large office. Even here, the number of 'instances' would be small and there would be no non-intrusive way to ensure that one was able to observe a representative sample. Any findings which resulted would not necessarily be useful in other organisations. By contrast, the relative anonymity of the virtual world means that one can easily observe transactions on many different newsgroups as well as collecting more anecdotal data in the normal course of being a web or net user.

**Accessibility:** One of the most remarkable features of the web and newsgroups is the fact they are now indexed down to the level of individual words by search engines. This permits particular strings of text to be identified and located for the purposes of studying speed of transmission of replicators. Because of the enormous numbers of sites involved, it may even be reasonable to hope that there is some statistical reliability in any findings, for example that the number of instances of a particular message which have found their way into archives are an accurate measure of the volume that were transmitted but not stored.

These dimensions also allow us to compare web or net research with other forms of research that are traditionally done. Historical or documentary research suffers from much lower accessibility and speed of transmission. Laboratory experiments have far higher accessibility but anonymity

(complete or even partial) is often considered unethical. Participant observation or ethnography also has advantages in respect of accessibility but their status in term of anonymity is unclear. In the next section, we turn to an analysis of some 'typical' replicators encountered on the net and elsewhere.

## Replicators I: Chain Letters

Even before the existence of the web, the 'chain letter' was a widely known instance of a replicating message (Goodenough and Dawkins 1994). These letters shaded into various kinds of deliberate pyramid selling schemes, the best known being the disastrous 'Ponzi Scheme' which is believed to have exacerbated civil unrest in Albania. (It is an interesting issue whether such a scheme could have had such a dramatic effect outside a highly insular former communist country. The analogy with viral 'resistance' is obvious.) The basic idea behind these letters is that one sends a dollar, or some stamps, or a photograph to a certain number of people, and at some time in the future one is deluged with many instances of the same. (What differentiates a deliberate pyramid selling scheme is whether all copies of the letter involve sending something to the originator.) These letters encourage replication and 'repay' the efforts of duplication with promises of (spurious) reward. Brief reflection makes it clear that only a very few people in the pyramid can show a substantial net gain from the scheme and only at the expense of many losers, the system being closed. Legislation against pyramid selling and ease of falsification - one only has to participate once to realize that the flood of inward mail is not forthcoming - has resulted in these paper letters becoming widely known and unpopular.

What is interesting is that the chain letters which have survived, both on and off the net, are those which actually promise *less* tangible benefits, but also incur lower costs. Two well known examples are the so called 'Grapefruit of Love' and the request for business cards or postcards (accounts differ) to be sent to 'Craig Shergold'. In the first case, the message is a basically uplifting one, that someone who cares about you has sent the message on with positive thoughts, and that you should in turn pass it on to someone who you think needs it. (In many cases, as with jokes, this message is passed from friend to friend rather than via strangers. This increases its transmission rate as does the fact that the message will be sent more often to those most likely to appreciate it and perhaps in turn to other 'like-minded' individuals.) In the second case, the request is that cards should be sent to a UK cancer patient to cheer him up. More generally, chain letters have increasingly stressed intangibles like 'good luck' rather than more concrete rewards. They thus render themselves immune from obvious falsification and tap into a human tendency to 'superstition', spotting patterns where none exist. If I fail to pass on a chain letter and shortly afterwards something bad happens, I may connect the two events and be more susceptible in future. Replicators offering intangibles may also benefit from the fact that while only some of the population is interested in stamps, photographs and perhaps even free money, there are few who would not seek 'luck' or 'happiness' which they can interpret as they will.

Chain letters raise several issues that are also relevant to other classes of replicator. Firstly, they support the argument put forward by Dawkins that persistence is more important than many copies being made of the same message. Both of the above replicators typically contain no instructions about message duplication, thus lowering the minimum costs of compliance, but they will continue to persist as long as they generate *on average* at least one copy per recipient.

Another benefit of this strategy from the replicator perspective is that it partially conceals the nature of the message *as a replicator*. The emergent implications of a message like 'Please pass this on to a concerned friend' are far less clear than the implications of 'Duplicate this message ten times and send it to ten friends.' (As before, multiple duplication is *only one* evolved response to a population which has already been extensively immunized. We shall discuss others shortly.) Secondly, we can see that different sorts of replicator may be more or less effective at jumping from one medium to another. Because the net has very low costs of transmission, we would have expected certain forms of the standard chain letter to transfer to the web very readily, as has been observed with the Craig Shergold message, but not for replicators originating on the web to shift to other media. In fact, paper chain letters of the traditional form now seem to be all but extinct, perhaps for this reason. Thirdly, we offer the perhaps surprising hypothesis that chain letters offering less obvious personal benefits may actually survive better. Apart from the points already made about falsification and generality of interest, altruistic replicators are also less likely to produce discouraging responses from third parties. Not only did chain letters themselves become unpopular, but, to a lesser extent, so did people who passed them on. This applied not only to people who passed them on knowingly, but also reflected badly on the naiveté of those who passed them on *unknowingly*. Since a highly negative response from a third party will not only discourage an individual from passing on a particular chain letter, but *anything* they subsequently recognize as such, replicators have an 'incentive' to avoid producing these responses and also to avoid identifying themselves. (A really extreme response may even take the critic 'out of circuit' altogether with regard to all those observing the outburst thus breaking up the transmission network.) This brings us on to the final point, which is that there is no consensus on the valuation of replicating messages. Certain replicators, like computer viruses, do concrete damage and are widely disapproved, while others have been reputed to fill up computer memory or slow down transmission of 'more useful' messages. (It would be interesting to know if these justifications are themselves urban legends.) By contrast, disapproval of certain replicators like lists of jokes - barring cases where they are sent to people who just don't like them - seems to depend entirely on their recognition *as* replicators which are regarded as a 'waste of time' or 'nuisance'. Perhaps some people resent being taken advantage of by pieces of text? This paper thus tries to remain neutral about whether messages of different kinds *should* replicate and merely concentrate on how they do it.

## Replicators II: Virus Warnings

So far the discussion has applied equally well to paper replicators and to those variants which have found their way onto the Internet. Now we turn to replicators which seem particularly suited to the distinctive properties of the virtual world. In doing so, we hope to learn something about those properties. Text 1 is a well developed instance of the standard warning message about the 'Good Times' (sometimes 'Good News') virus. This is a very robust class of replicating message that has been circulating on the net for at least three years under a variety of different names (Join The Crew, xxx-1, Win a Holiday, Penpal Greetings, Deyeendra).

Please take note of this internet virus.

There is computer a virus that is being sent across the Internet. If you receive an E-mail message with the subject line "Good Times", DO NOT read the message.

DELETE it immediately. Some miscreant is sending email under the title "Good Times" nationwide, if you get anything like this, DON'T DOWN LOAD THE FILE!!!! It has a virus that rewrites your hard drive, obliterating anything on it.

Please be careful and forward this mail to anyone you care about.

The FCC released a warning last Wednesday concerning a matter of major importance to any regular user of the Internet. Apparently a new computer virus has been engineered by a user of AMERICA ON LINE that is unparalleled in its destructive capability. Other more well-known viruses such as "Stoned", "Airwolf" and "Michaelangelo" pale in comparison to the prospects of this newest creation by a warped mentality. What makes this virus so terrifying, said the FCC, is the fact that no program needs to be exchanged for a new computer to be infected. It can be spread through existing the email systems of the Internet.

Once a computer is infected, one of several things can happen. If the computer contains a hard drive, that will most likely be destroyed. If the program is not stopped, the computer's processor will be placed in an the nth-complexity infinite binary loop-which can severely damage the processor if left running that way too long.

Unfortunately, most novice computer users will not realize what is happening until it is far too late. Luckily there is one sure means of detecting what is now known as the "Good Times" virus. It always travels to new computers the same way in a text email message with the subject line reading "Good Times". Avoiding infection is easy once the file has been received simply by NOT READING IT!. The act of loading the file into the mail server's ASCII buffer causes the "Good Times" mainline program to initialize and execute.

The program is highly intelligent. It will send copies of itself to everyone whose email address is contained in a receive-mail file or a sent-mail file, if it can find one. It will then proceed to trash the computer it is running on.

The bottom line is: If you receive a file with the subject line "Good Times", delete it immediately! Do not read it! Rest assured that whoever's name was on the "From" line was surely struck by the virus. Warn your friends and local system users of this newest threat to the Internet! It could save them a lot of time and money.

**DO NOT DOWNLOAD ANY FILE NAMES PKZIP300 REGARDLESS OF THE EXTENSION**

We work closely with the military and received this message from a reliable source in DC this morning. A NEW Trojan Horse Virus has emerged on the Internet with the name PKZIP300.ZIP, so named as to give the impression that this file is a new version of the PKZIP software used to "ZIP" (compress) files. DO NOT DOWNLOAD this file under any circumstances!!! If you install or expand this file, the virus WILL wipe your hard drive clean and affect modems at 14.4 and higher. This is an extremely destructive virus and there is NOT yet a way of cleaning up this one.

**REPEAT: DO NOT DOWNLOAD ANY FILE NAMES PKZIP300 REGARDLESS OF THE EXTENSION.**

### **Text 1. A Particularly Horrible Virus Warning**

This message displays a number of interesting features:

**Essential Components of a Replicator:** Replicators appear to have three essential components. The first is some sort of content, in this case, a warning about the viruses 'Good News' and 'PKZIP300'. The second component is an attempt to *warrant* the content of the message thus motivating belief in the reader. The third component is a separate motivation to transmit and/or multiply the message. In this example and those that follow, warranting text is marked in green and motivations to replicate in blue.

**Non Falsifiability and Self-Concealment:** It has already been mentioned that replicators may be more effective if they offer benefits that are less tangible since they may benefit both from human 'superstition' and the difficulty of effective falsification. This class of replicators is particularly interesting in this respect because it is the *warning* and not the virus itself that constitutes the replicator. In fact, as far as anyone can establish, there is no virus corresponding to any replicators in this class. Thus this replicator survives by eluding conclusive falsification and it does this by making use of a nonexistent referent. Although the message can be disregarded for technical reasons, these will not be effective in preventing non-technical users from transmitting it. This is particularly the case because, for reason discussed in Woolgar and Russell (1990), users are rather inclined to believe in computer viruses as just 'punishment' for electronic promiscuity. This replicator also actively discourages falsification by claiming that reading the infected message is precisely the way that the virus is spread. Thus anybody receiving an email with one of the 'threatening' titles will have a strong incentive not to investigate the message itself. This gives rise to another property of this class of replicators which is that they are frequently renamed, thus concealing themselves and extending their lifespans. Although all these messages are functionally equivalent, a person who has already been affected by xxx-1 may not spot the similarity with Good News particularly in the face of other minor variations occurring in transmission. Interestingly, this renaming may also foster the illusion that there are always new viruses cropping up and this may maintain willingness to pass on all messages in this class.

**Technical Incoherence and Implausibility:** The pieces of text marked in red and pink suggest that this replicator is intended to be transmitted by those who are relatively technically incompetent. Such novices form an increasing proportion of those on-line and are thus an effective vector. The message makes no distinction between an email message in plain text, a file (perhaps an attachment), a program and a virus as a component of a file or program and the discussion of the effects of the virus maintains this confusion. Some of the warnings are either meaningless - the delightful and sinister sounding 'nth-complexity infinite binary loop' - or implausible. Most operating systems and email programs simply do not support programs or macros that would allow the hard drive to be reformatted or search the mailboxes to compile lists of addresses. (The nearest we have so far to a virus that can cross platforms or affect machines with different operating systems are various kinds of macro viruses. However, these are similarly limited by the functionality of the application from within which they operate.) Finally, it is not at all clear what a virus program could do to a modem and the message is wise not to say!

**Structural Incoherence:** In general, the form of all these messages is rather strange, with repetitions, changes of tack, awkward constructions and inconsistencies suggesting that they have been augmented and modified repeatedly. Subjectively, the style is far worse than and different from that of a typical 'designed' email message with a single author. Although we have

not yet discussed the extent to which these messages may be deliberately designed, it appears that they continue to change and evolve after they have been 'released' in ways that were unforeseen by the original designer. To this extent, evolution and design are not incompatible as understandings of the spread of replicating messages. (More generally, replicators are subject to evolution to the extent that their implications are not foreseen. Even the most cunning designer of replicators is unlikely to be aware of all the ramifications of a given text so it will still be subjected to some selection pressure.) The replicator in Text 1 actually consists of two separate warnings that have become grafted together. The PKZIP warning is much more coherent than the other and more likely to be based on 'fact', but it is not clear why the two messages are traveling together since they refer to completely different infections. (In biology, there is a concept of 'evolutionary piggybacking'; one gene may continue to be associated with another provided it causes no harm. In this case, of course, two warnings may actually be more impressive and effective than one.) In this case, one can even speculate on where the two messages are joined, using knowledge of the postulated structure of a replicator. It seems likely that the sentence which begins 'Warn your friends...' was once a separate paragraph, since without it, the PKZIP part of the message which follows lacks an exhortation to reproduce and is thus 'incomplete'.

**Warranting and Motivation to Propagate:** These can be seen as two stages in getting the replicator transmitted. The first is that the content of the message should be believed. The second is that, given this belief, it should seem like an appropriate thing to pass the message on. Warrants internal to the message are usually provided both directly, using references to 'authorities' like the FCC (US Federal Communications Commission) or 'the military' (perhaps surprisingly) and indirectly, by adopting a quasi-technical tone that is a strange mixture of bulletin and anecdote. (In Text 2 below, for example, the writer of the message is supposed to know all the bad things that happened to the victim, even down to the header of the damaging file, but at the same time they can provide no better identification than 'someone'. Urban legends often provide this strange mixture of detail and vagueness.) The motivation to transmit is almost invariably couched in altruistic terms, urging the reader to pass the message on to friends or colleagues, thus enhancing its external warrant as coming from someone known to the next recipient.

Text 2 is another example of a replicator of the same class which shares almost no genetic material with Text 1 although it conveys a very similar message. (One possibility for the generation of new variants is that stories warning of viruses may also be told 'offline' with individuals then retelling the story into an email 'from scratch' as far as they recall it. One would expect the technical details to suffer particularly badly from this process, but it will be hard to study it.) It has some interesting additional features. Firstly, it has an external warrant from Professor Smith (name altered for confidentiality) who may or may not exist and have passed the message on in good faith. More generally, the headers of these messages, showing that they have been extensively passed on, may also act to reinforce their authenticity though they can be faked relatively easily. Secondly, this replicator conflates at least four warnings, suggesting that composite messages are more effective. Thirdly, there is a new twist to the ability of replicators to avoid falsification. Even with the dire threats of infection if one reads messages with the 'hazardous' titles, it may still occur to people that they don't actually seem to be getting any actual messages only the warnings. A virus purportedly called 'Returned or Unable to Deliver' gains effectiveness because all users, especially novices, will send messages that 'bounce' and

thus generate their own spurious confirmations of the existence of this replicator. If a novice sees a virus warning before reading one of these bounce messages, they may forever after see such messages as evidence of the virus being so effectively discouraged from checking. Finally, there is also a new twist on the exhortation to duplicate which is that this is '... an [sic] new virus ...' which '... not many people know about ...' Such additions seem to rely on a general inability to perceive the emergent implications of receiving replicating message copies personally. It is highly unlikely that one will really be among the first to receive such a message and the header may even directly refute that belief but there seems to be a tendency to take email at local 'face value'. (The same can be said of pyramid selling schemes where individuals fail to ask what will happen if everyone follows the same logic that they themselves have.)

The following came to Professor John Smith and he wanted to share this information with everyone as soon as possible. Please read it and be careful what you open.

-----Original Message-----

Sent: Wednesday, August 27, 1997 5:49 PM

Subject: Computer virus alert

WARNING!!! The following information was received yesterday from IBM. Please share it with anyone that might access the Internet.

There are several new viruses being sent via e-mail entitled "JOIN THE CREW," PENPAL GREETINGS," and "RETURNED OR UNABLE TO DELIVER."

If you receive an e-mail entitled "JOIN THE CREW," DO NOT OPEN IT. It will erase EVERYTHING on your hard drive! Send this message out to as many people as you can ... this is an new virus, and not many people know about it yet.

There is a dangerous virus propagating across the Internet through an e-mail message entitled "PENPAL GREETINGS," DO NOT DOWN LOAD ANY MESSAGE ENTITLED "PENPAL GREETINGS." Delete it without reading it. This message appears to be a friendly letter asking you if you are interested in a penpal. By the time you read the letter, it is too late. The Trojan Horse virus will have already infected the boot sector of your hard drive, destroying all of the data present. It is a self-replicating virus. Once the message is read, it will AUTOMATICALLY forward itself to anyone whose e-mail address is present in YOUR mailbox! This virus will destroy your hard drive and holds the potential to destroy the hard drive of anyone whose mail is in your in-box, and anyone whose mail is in their in-box, and so on, and so on. If this virus keeps getting passed, it has the potential to do a great deal of damage to computer networks worldwide!!!

DO NOT open or even look at any mail that is entitled "Returned or Unable to Deliver." This virus will attach itself to your computer components and render them useless. Immediately delete any mail items that say this. AOL has said this is a very dangerous virus, and there is NO remedy for it at this time.

Please forward this message to all of your on-line friends. ASAP.  
Here is another virus notice.

Subject: Virus Warning

IMPORTANT!!!!

It is essential that this problem be reconciled as soon as possible.

A few hours ago, Someone opened an E-mail that had the subject heading of ao4free.com

Within seconds of opening it, a window appeared and began to display all files that were being deleted ... The user immediately shut down the computer, but it was too late.

This virus wiped all out. It ate the Anti-Virus Software that comes with the Windows '95 Program along with F-Prot AVS. Neither was able to detect it. Please be careful and send this to as many people as possible, so maybe this new virus can be eliminated.

Please pass this on, ..... to anyone you know.

----- End Forwarded Message -----

### **Text 2. Another Horrible Virus Warning**

In the next section, we turn to evolved 'responses' to replicators which can be seen as immunizing the population and reducing further transmission.

### **Replicators III: Immunization Responses**

In this section, we consider the evolved responses to particularly 'irritating' replicators. One difficulty is that there is relatively little permanence to newsgroup postings and individual emails so these will not provide long term immunity. This is also true because the numbers of web and net users is continuously growing as is the proportion of novices. Although there are many web pages devoted to debunking common hoax messages, there are many more people using email than have web access and the same applies to the possession of the relevant skills. How have immunizing replicators co-evolved to compete with the original replicators? Texts 3 and 4 provide two examples. Because they are emails, they affect the same constituency that is transmitting the replicators and will continue to propagate in the same way. Interestingly, neither message *relies* on exhortations to transmit although the second makes these exhortations in an obviously ironic way. Instead, transmission appears to take place through a combination of intrinsic humor and satirizing the pre-existing replicators. We would expect these immunizers to die out without a steady stream of the original replicators to feed on.

\*\*\*\*\* Badtimes Virus Alert \*\*\*\*\*

If you receive an e-mail with a subject of "Badtimes", delete it immediately WITHOUT reading it.

This is the most dangerous eMail virus yet. It will re-write your hard drive. Not only that, but it will scramble any disks that are even close to your computer.

It will recalibrate your refrigerator's coolness setting so all your icecream goes melty.

It will demagnetize the strips on all your credit cards, screw up the tracking

on your VCR and use subspace field harmonics to scratch any CD's you try to play.

It will give your ex-boyfriend/girlfriend your new phone number.

It will mix Kool-aid into your fishtank. It will drink all your beer and leave its socks out on the coffee table when there's company coming over.

It will put a dead kitten in the back pocket of your good suit pants and hide your car keys when you are late for work.

Badtimes will make you fall in love with a penguin. It will give you nightmares about circus midgets. It will pour sugar in your gas tank and shave off both your eyebrows while dating your current boyfriend/girlfriend behind your back and billing the dinner and hotel room to your Visa card.

It will seduce your grandmother. It does not matter if she is dead, such is the power of Badtimes, it reaches out beyond the grave to sully those things we hold most dear.

It moves your car randomly around parking lots so you can't find it. It will kick your dog. It will leave libidinous messages on your boss's voicemail in your voice! It is insidious and subtle. It is dangerous and terrifying to behold. It is also a rather interesting shade of mauve.

Badtimes will give you Dutch Elm disease. It will leave the toilet seat up.

It will make a batch of Methamphetamine in your bathtub and then leave bacon cooking on the stove while it goes out to chase grade-schoolers with your new snowblower.

These are just a few signs ... Just be very careful!

### Text 3. Badtimes Immunizer

Rather than a dry warning about specific replicators which would date rapidly, Badtimes satirizes the tone and content of the whole class of replicators and makes claims that are so self-evidently ridiculous that they may provoke skepticism about similar claims couched in more technical terms. Compare the 'subspace field harmonics' in Badtimes with the 'nth-complexity infinite binary loop' in the Good News message. Again, it is not clear whether what we should assume about the amount of deliberation that goes into these immunizers. It may be that some or all of them were rationally 'constructed' with the specific intention of reducing the transmission of replicators, or that they were created simply as a joke that has turned out to serve that 'purpose'. This kind of study illustrates the problems of relating intentions to consequences (Merton 1936).

\*\*\*\*\*  
WARNING, CAUTION, DANGER, AND BEWARE!

Gullibility Virus Spreading over the Internet!  
\*\*\*\*\*

WASHINGTON, D.C.--The Institute for the Investigation of Irregular Internet Phenomena announced today that many Internet users are becoming infected by a new virus that causes them to believe without question every groundless story, legend, and dire warning that shows up in their inbox or on their browser. The Gullibility Virus, as it is called, apparently makes people believe and forward copies of silly hoaxes relating to cookie recipes, email

viruses, taxes on modems, and get-rich-quick schemes.

"These are not just readers of tabloids or people who buy lottery tickets based on fortune cookie numbers", a spokesman said. "Most are otherwise normal people, who would laugh at the same stories if told to them by a stranger on a street corner". However, once these same people become infected with the Gullibility Virus, they believe anything they read on the Internet.

"My immunity to tall tales and bizarre claims is all gone", reported one weeping victim. "I believe every warning message and sick child story my friends forward to me, even though most of the messages are anonymous." Another victim, now in remission, added, "When I first heard about Good Times, I just accepted it without question. After all, there were dozens of other recipients on the mail header, so I thought the virus must be true". It was a long time, the victim said, before she could stand up at a Hoaxees Anonymous meeting and state, "My name is Jane, and I've been hoaxed". Now, however, she is spreading the word. "Challenge and check whatever you read," she says.

Internet users are urged to examine themselves for symptoms of the virus, which include the following:

The willingness to believe improbable stories without thinking.

The urge to forward multiple copies of such stories to others.

A lack of desire to take three minutes to check to see if a story is true.

T. C. is an example of someone recently infected. He told one reporter, "I read on the Net that the major ingredient in almost all shampoos makes your hair fall out, so I've stopped using shampoo". When told about the Gullibility Virus, T. C. said he would stop reading email, so that he would not become infected.

Anyone with symptoms like these is urged to seek help immediately. Experts recommend that at the first feelings of gullibility, Internet users rush to their favorite search engine and look up the item tempting them to thoughtless credence. Most hoaxes, legends, and tall tales have been widely discussed and exposed by the Internet community.

\*\*\*\*\*  
This message is so important, we're sending it anonymously! Forward it to all your friends right away! Don't think about it! This is not a chain letter! This story is true! Don't check it out! This story is so timely, there is no date on it! This story is so important, we're using lots of exclamation points! Lots!! For every message you forward to some unsuspecting person, the Home for the Hopelessly Gullible will donate ten cents to itself. (If you wonder how the Home will know you are forwarding these messages all over creation, you're obviously thinking too much.)

\*\*\*\*\*

#### **Text 4. Gullibility Virus Immunizer**

The Gullibility Virus also satirizes the standard form and content of a replicating message rather closely - thus increasing its effectiveness - but at the same time highlights and makes explicit the

sort of 'tricks' used by replicators to ensure their propagation. In addition, it provides motivation, through ridicule, to avoid being associated with the victims described, even though they are obviously fictitious. Notice also that the three virus 'symptoms' described correspond exactly with the essential components of a replicator postulated in this paper: content - 'A lack of desire to take three minutes to check to see if a story is true' - assent - 'The willingness to believe improbable stories without thinking' - and reproduction - 'The urge to forward multiple copies of such stories to others.' There is not room to pursue the analysis of co-evolution in this paper, but the point of this section is to draw intention to the fact that the analogy with biological evolution is not a 'shallow' one in that it deals not only with the evolution of simple replicators, but also with more sophisticated processes like symbiosis (where the survival of the immunizers depends on the survival of replicators) and co-evolution. Co-evolution is the process by which the nature of the immunizers continues to change to take account of the nature of the replicators and vice versa (Thompson 1994).

## Conclusions and Future Research

In a single paper, it has not been possible to do more than introduce a memic analysis of messages on the web and net while attempting to sketch out a few of the implications. However, even this cursory analysis suggests that it may be useful to take a memic view of messages as manipulating their senders rather than vice versa. This is particularly the case where the relations between intentions and outcomes are not straightforward. There are four main areas in which the ideas here can be developed further:

**Qualitative Analysis:** Although I have analyzed a few replicator texts in a rather simplistic way, these texts and the dialogues about the correctness (or otherwise) of their transmission will bear considerable more detailed investigation from the perspective of Conversation Analysis, particularly making use of a much larger sample and using rigorous Qualitative Data Analysis. In particular, how do these messages do the task of getting transmitted and how do individuals argue for and against their transmission? What evidence do they bring to bear when the structure of the replicators actually seems to discourage falsification? Conversation Analysis proceeds from a Social Constructionist perspective which emphasizes the importance of competing accounts rather than the existence of objective facts of the matter. Regardless of whether this view is a correct representation of social processes - which of course one can only assert from a **non** Social Constructionist perspective - or an essential precondition to the use of CA, it is still the case that discussion of messages about potentially nonexistent viruses comes as close to a paradigm case for CA as it is possible to get. There really is nothing beyond the text of interaction to refer to.

**Quantitative Analysis:** So far, the claims that certain features of the replicator help or hinder its transmission have been very much subjective. This discussion needs to be followed up by the collection of a very much larger and more representative sample of instances from the web and net and attempts both to record the spread of different replicators over time and to demonstrate the evolution of variants for a single replicator.

**Experimental Analysis:** One very effective way to investigate the spread of replicators would be to deliberately design and release a replicator which combined the best insights about

effective transmission with ease of subsequent tracing. (There are also other experimental possibilities like an online questionnaire about beliefs and attitudes connected to replicators.) Ironically, too much detailed public discussion of this tracer might lower its effectiveness, but it is worth pointing out here that such an action raises ethical issues both because it contravenes the prevailing culture of the web and net in spreading unsolicited messages but also because it may involve deliberate deception of the researched. On the other hand, being 'open' about the function of the replicator, providing a contact email address on the message and so on might actually make the replicator far more effective. Indeed, it might become **too** effective so that the researcher was swamped with data. (This is in itself part of the urban legend of replicators, that the hospital mentioned in the 'Craig Shergold' replicator was swamped with cards.) Perhaps this message, or a summary of it, appended to a request for instances might itself constitute an effective replicator. I therefore ask, at this point, that anyone who receives messages of this kind or one they think they might be replicators to forward them to Edmund Chattoe with the header 'Replicator Instance' so I can filter them automatically. (I will doubtless make the amount of email I receive manageable by putting this request at the *end* of a rather long paper rather than at the beginning!) This raises another interesting issue which is that research in this area obviously contravenes the positivist agenda of distinguishing between researcher and researched. This paper, available as it is on the web, may well become part of the general debate on the transmission of replicators as was part of my intention in writing it. Conversely, in appealing for data, I will be changing the behavior of those who might transmit replicators, again not entirely unintentionally!

**Simulation:** Although there are techniques for both qualitative and quantitative analysis of existing and introduced replicators on the web, it is generally hard to synthesize these approaches. One possibility is to use computer simulation to build a model of an email network in which the mental states of transmitters or non-transmitters of messages are explicitly represented in simplified form so that the relation of intentions to replicator outcomes can also be studied. In such a model, some agents may themselves have belief sets about what will be credible and construct their messages accordingly, while others respond to messages 'unreflectingly' based only on what they themselves believe. Such simulations can also be used as investigative tools to suggest further hypotheses for testing on the web and net.

Finally, although this paper has chosen a research topic that is apparently very specific, it is possible that the insights gained will be far more general. How can marketers and advertisers produce replicators that will spread effectively by word of mouth? What are the memic properties of ideologies and cultural norms that tend to ensure their propagation and persistence? What do replicators reveal about the distinctions between 'objective' and socially constructed facts? What is the relation between intentions and outcomes and between socially evolved and deliberate or rational behavior?

## Bibliography

Brunvand, Jan Harold (1984) *The Choking Doberman and Other 'New' Urban Legends* (New York, NY: W. W. Norton).

Darwin, Charles (1968) *On The Origin of Species by Means of Natural Selection, or the Preservation of Favoured Races in the Struggle for Life*, reprint of the first edition (Harmondsworth: Penguin).

Dawkins, Richard (1976) *The Selfish Gene* (Oxford: Oxford University Press).

Dawkins, Richard (1983) *The Extended Phenotype: The Gene as a Unit of Selection* (Oxford: Oxford University Press).

Edey, Maitland A. and Johanson, Donald C. (1989) *Blueprints: Solving the Mystery of Evolution* (Oxford: Oxford University Press).

Fisher, R. A. (1930) *The Genetical Theory of Natural Selection* (Oxford: Clarendon Press).

Goodenough, Oliver R. and Dawkins, Richard (1994) 'The 'St. Jude' Mind Virus', *Nature*, vol. 371, 1st September 1994, pp. 23-24.

Lynch, Aaron (1996) *Thought Contagion: How Belief Spreads Through Society* (New York, NY: Basic Books/HarperCollins).

Merton, R. (1936) 'The Unanticipated Consequences of Purposive Social Action', *American Sociological Review*, pp. 894-904.

Thompson, John N. (1994) *The Coevolutionary Process* (Chicago, IL: The University of Chicago Press).

Wooffitt, Rob C. (1992) *Telling Tales of the Unexpected: The Organisation of Factual Discourse* (London: Harvester Wheatsheaf).

Woolgar, Steve and Russell, Geoff (1990) 'The Social Basis of Computer Viruses' Discussion Paper, Centre for Research into Innovation, Culture and Technology (CRICT), Brunel University, December.

## **Links to Some Hoax Information Sources**

- Department of Energy Computer Incident Advisory Capability
- Symantec Anti Virus Research Center
- McAfee Associates Virus Hoax List
- The Urban Legends Web Site
- Urban Legends Reference Pages
- Datafellows Hoax Warnings

## **Acknowledgements**

Perhaps unusually, this paper had a very specific genesis, which was the presentation and subsequent discussion of Woolgar and Russell (1990) at a conference in Bournemouth. I can say

that without hearing that paper, this one would never have been written and I am very grateful to Steve Woolgar for provoking my interest in this topic.

### **The Author: Edmund Chattoe**

*Centre for Research on Simulation in the Social Sciences (CRESS), University of Surrey,  
Department of Sociology, University of Surrey, GU2 5XH.*

*Edmund Chattoe is Associate Director of CRESS (The Centre for Research on Simulation in the Social Sciences) based in the Department of Sociology at the University of Surrey. This Centre aims to promote computer simulation among social scientists by offering training, advice and information. His research interests include simulation techniques, processes of social evolution, the application of Artificial Intelligence and Cognitive Science to social science and modeling decision making which involves both "social" and "rational" components.*

# PARASITE ECOLOGY AND THE EVOLUTION OF RELIGION

By  
Ben Cullen

Dept of Archaeology and Palaeoecology  
Queen's University of Belfast  
BELFAST BT7 1NN  
NORTHERN IRELAND

**Abstract:** It is argued that the blanket view of religion as a disease, advocated by Dawkins, is inconsistent with the principles of parasite ecology. These principles state that vertically transmitted parasites evolve towards benign, symbiotic states, while horizontally transmitted parasites increase their virulence. Most of the world's established religions are transmitted vertically, from parents to children, and are therefore expected to be benign towards their hosts. Yet, certain horizontally transmitted cults, such as the Aum Shinrikyo, seem to effectively exploit their hosts in a way similar to an infectious disease.

## Foreword by the editor

On December 29, 1995, while preparing this paper for publication, Ben Cullen unexpectedly died, apparently from an arrhythmia of the heart. This paper was one of two wonderful presentations which he had given at the conference "Einstein Meets Magritte" the previous summer. The unfinished manuscript consisted basically of the text he had read aloud during his presentation, which means that it lacked the structure typical of texts meant for publication. Therefore, I have edited the manuscript by dividing it into paragraphs and sections, eliminating references to photographs projected during the presentation, adding an abstract and bibliography, and correcting various small errors.

Let me elaborate why I find it important for Ben's work to be brought to the attention of the public. An Australian born in 1964, Ben Cullen graduated from the University of Sydney in 1987, with a degree in Prehistoric Archeology. After a stay as Research Fellow at the University of Wales, Lampeter, he moved to the Queen's University of Belfast a few months before his death. He had just defended his PhD thesis in archeology, in which he developed his "Cultural Virus Theory" (Cullen, 1993, 1995c, 1996), a stimulating new model of the evolution of culture, reminiscent of, but distinct from, the memetics view proposed by Dawkins (1989).

I met him in 1992 in London at a conference, where we were both presenting models of cultural evolution. We were obviously on the same wave length and had several long discussions after the lectures. I was impressed by his sharp intellect and the depth of his knowledge and understanding. Afterwards, we kept in touch by exchanging papers. Just a few weeks before his death, he sent me a preprint (Cullen, 1995) which he asked me to "really read, not like some others", which I interpreted to mean that many people misinterpret his ideas. There does not seem to be much danger of misinterpretation with the following paper, which is a jewel of clarity and simplicity.

*Francis Heylighen*

## **Introduction**

There has been much recent publicity of the idea of religion and chain letters being diseases of the mind, or human information parasites as they are sometimes termed. Prominent amongst these have been pieces by the biologist Richard Dawkins in the London press, and also a piece about the St. Jude Chain Letter in a letter published recently in the leading scientific journal *Nature*. (Goodenough & Dawkins, 1994). Headlines such as the one about the 'St Jude' mind virus have not excited a great deal of controversy, presumably because few people publicly identify with chain letters.

However, other headlines, such as "Is Religion Just A Disease?" in the Daily Telegraph (Dawkins, 1993b), have excited a great deal of vociferous debate, and pricked many egos. Richard Dawkins's view of religion (1993a,b) as a disease not surprisingly offended members of the Church of England, and various bishops wrote pieces condemning both Dawkins and his view. But most objections to the "religion as a disease" idea have come from the Church acting in self-defense, guarding its cultural capital so to speak. Most defenses have come from theologians who submit the general idea that Dawkins is discussing spiritual matters beyond his understanding, a predictable theologian response. Very few objections have come from the sciences, human or otherwise. Few articles have explicitly disputed the scientific credentials of the Dawkins view, or the neo-Darwinian logic behind it.

In this paper I intend to go against this trend, and argue against the "religion as a disease" position, on the grounds that Dawkins has not manipulated neo-Darwinian logic consistently. In other words, I will argue that Dawkins has erected a highly personal view of religion which is in complete conflict with both his own selfish gene theories (Dawkins, 1989), and a growing body of parasite research.

## **Principles of Parasite Ecology**

Let me enumerate the basic principles emerging in recent parasite ecology, the rules which the Dawkins view of religion are in direct conflict with, yet also the rules which Dawkins himself helped to articulate.

The first rule is that *vertically transmitted* parasites, namely, those that tend to pass from parent to child, evolve toward decreasing virulence, ultimately approaching benign or symbiotic configurations. Vertically transmitted parasites find themselves in the same sequence of bodies as the genes of their hosts, generation after generation, and the two therefore share a common future. They are in the same evolutionary boat, so to speak, and if the boat sinks, both parties drown.

One classic example of this type of relationship is that between wood-boring ambrosia beetles and bacteria, an example which Dawkins has featured in discussions of this nature. The bacteria are entirely dependent upon the beetles for food and shelter, and live their entire lives inside beetle tissue, being carried into the next generation inside the eggs which the females lay. This route to the future, however, ensures that they pass from each parent beetle to the beetle's offspring, and they are, therefore, vertically transmitted. Thus the rule would predict that these

bacteria should be symbiotic, and indeed they are, as without them the beetles would be unable to reproduce. This is because male beetles can only develop from eggs in which the bacteria are present. Without the bacteria the beetles would be a race of female individuals who could only reproduce through sexual reproduction, a difficult situation at the best of times.

This general rule linking symbiosis with long term host/parasite cohabitation and vertical transmission also finds support elsewhere, particularly in the field of "Darwinian medicine" (Nesse & Williams, 1995), and in the work of researchers such as Paul Ewald (1990, 1995). For instance, it has been shown that new and less virulent strains of the AIDS virus are proving more successful in human populations where condoms are frequently used, and where viruses therefore have to spend longer periods of time in each host (Ewald, 1994).

Rule number 2 is simply the reverse of Rule number 1. Disease is often connected to *horizontal modes of transmission*, where parasites are transmitted indiscriminately with respect to host relationships. Horizontally transmitted parasites are as likely to be passed on between families as within families. Parasites of this kind tend to evolve toward a typical "disease" configuration. The presence of the parasite significantly alters the host body or behavior in such a way as to improve the parasite's chances of reproduction, and decrease that of the host. Curiously, this may involve enhancing the host's chances of survival, while preventing its reproduction, in an attempt to produce a thriving parasite factory. Some barnacles, for example, parasitize crabs by targeting the crab's sex organs first, and the vital organs last, so maintaining a living resource for the parasite for as long as possible.

One of the most spectacular parasites of all is not an animal, but a plant which hunts other plants. It too is horizontally transmitted, and conforms to the rule. The corpse flower, or *Rafflesia*, is a family of plants so called because many of them exude a smell of rotting meat, mimicking the smell, colour and texture of a corpse. At 30 to 40 inches or about 1 metre across, it is the largest known single flower, and one of the most extraordinary predators of the plant world. It is an inhabitant of the rainforests of Borneo and Sumatra. The corpse flower has no leaves, no stem, no roots or trunk. It actually lives inside the tissues of another plant, one of the many vines which grow so abundantly in the rainforest canopy, as a net of tiny threads. How it actually gets inside these vines is not known. Thus the stem from which this truly monstrous flower grows does not belong to the flower, but to rafflesia's victim. Once the parasitic flower grows to its full size, the host vine has usually collapsed into the leaf litter of the forest, and is hidden from view.

The *corpse flower* is an excellent example of a predator which stalks its prey from within, biding its time in the body of the host vine, before leaping to new victims by means of its extraordinary flowers. To move through the forest, it enlists the help of tree shrews (small arboreal mammals) and insects. Viewed from the inside, the corpse flower is even more remarkable. Exuding a powerful aroma similar to a pile of fish after three days in the hot sun, it not surprisingly attracts flies. The flies crawl down to the sex organs, and there they are daubed with pollen which may then be carried to female flowers. The female flowers produce seeds which are then transported by small tree shrews to new vine hosts, elsewhere in the rainforest.

## Religions as Parasites

Rafflesia is particularly relevant to the "religion as a disease" debate because it represents a rare example of a large predator that inhabits the bodies of its prey. It allows us to begin to imagine how an organisation such as a religion could be seen as a life form which inhabits the bodies of people, in the same way that rafflesia inhabits the bodies of rainforest vines. Anatomically, the corpse flower and cultural life forms have much more in common than one might expect.

This is a representation of the anatomy of a religion and its material products, in this case the megalithic tombs of Neolithic western Europe (Cullen, 1994). The outer ring of structures represent megaliths, but they could just as well be churches, or mosques or cathedrals. The central structures of the religion are composed of people, which are represented by the figure in the bottom corner--a circle for a head and a monk's robe for the body. There is an outer ring of acolytes or novices, whose heads are represented by circles A, B, C, and D their bodies by a robe, and the actions they perform in service of the religion by arrows. Finally there is an inner circle of priests or clergy, whose heads are represented by circles 1, 2, 3, and 4, their bodies by robes, and their actions by arrows. The religion or cultural organism consists solely of the shaded black areas and their material products - the people involved in the organisation may perform many other roles in society, some of these roles might work against the religion. So this schematic diagram gives us some idea of what a cultural life form looks like.

Now, using the rules of parasite ecology discussed a little earlier, we can begin to predict which religions will evolve toward disease-like states, and which religions should evolve toward benign or symbiotic states. Vertically transmitted religions, like vertically transmitted viruses and other parasites, should evolve toward symbiotic or benign states. This immediately disqualifies most of the world's established religions from the "disease" scenario, because they are generally passed on from generation to generation within the family. Family dependent religious affiliation should therefore evolve toward benign or symbiotic configurations. Dawkins actually points out the vertically transmitted nature of religions, although he does not connect it to any prediction of symbiosis.

For the same reasons, we would expect the religions of small scale, kinship based societies, such as those of indigenous Australian cultures and their associated artistic traditions, to evolve toward symbiotic configurations. When a religion occupies of small community in which members of subsequent generations are of close genetic relationship to previous generations, a pathological religion would tend to plunge itself and its community into oblivion.

But this is not to say that there can be no such thing as a parasitic religion. However, if we are to find a truly pathological religion, we will have to look for examples where the pattern of transmission is horizontal, or essentially indiscriminate with respect to genetic relatedness of new recruits. Established world religions cannot be totally excluded from the picture, as while congregational membership may be essentially vertically transmitted, membership of the priesthood is not so rigid. Nor can the religions of kinship based communities be totally excluded, as kinship is often employed as a means of social categorisation for both strangers or newcomers and family alike. So called "fictive" kinship of this kind can and often does follow a

line of descent which does not conform to strict lines of genetic inheritance. So we might expect occasional parasitic forms of religion to arise in the established traditions.

But are there any more dramatic examples around? I think that we may not have to look too far. Religious cults would appear to be a good candidate. Chizuo Matsumoto, better known as Shoko Asahara, is the leader of the doomsday cult of Aum Shinrikyo or Way of the Divine Truth. The Way of the Divine Truth organization was apparently responsible for the recent nerve gas attack on the Tokyo subway system.

The Aum Shinrikyo doomsday cult undoubtedly follows a horizontal pattern of transmission, selecting its recruits more for what they can do for the cult than for their relatedness to existing members. The rate at which it acquired members, many of them outside Japan, is testimony to the non-vertical nature of its reproduction. In the space of a decade or so, the cult acquired some 30,000 members in Russia, presumably none of them related to Asahara himself. Within Japan the cult boasted a surprising number of members from all walks of life. Asahara carefully chose the most talented individuals for his priests, recruiting lawyers, scientists, chemists and engineers. With so many thousands of members in such a short time, it is clear that the cult did not rely on the fecundity of its members for its expansion. Clearly, then, the Aum Shinrikyo doomsday cult was a horizontally transmitted phenomenon.

Parasite ecology, then, would predict that it would exploit some or all of its members. Is there any evidence of this? It would appear that there was. One recruit had to provide a list of her assets and promise that they would be passed over to the cult on her death. The most privileged were allowed to drink Asahara's blood or semen, but actually had to pay an incredible 74,000 pounds for the pleasure. Cult members received very little in return for these immense financial contributions to the cult. Apart from the lucky few allowed to drink a little blood or semen, they received dormitory or boarding school style accommodation, and two meager vegetarian meals per day. The 53 children moved from the cult's commune proved to be suffering from severe malnutrition, while Asahara himself lived in luxury. An older recruit related how he was told to drink several gallons of water and vomit them as part of his training. So it would seem that the Aum Shinrikyo doomsday cult of Shoko Asahara does indeed conform to principles of parasite ecology.

In conclusion, then, it would appear that the blanket view of religion as a disease, as advocated by Dawkins (1993a,b), is not consistent with recent research into the nature of parasite evolution. Many religions are being vertically transmitted or family dependent, and we would therefore expect them to evolve toward symbiosis or at least benignness. As Dawkins has remarked, it is an extraordinary fact that if we adhere to a faith at all, it is overwhelmingly likely to be the same as that of our parents. This simple fact ought to ensure that if a religion which followed this pattern of transmission ruthlessly exploited its congregation, it would eventually plunge both itself and its people into extinction.

## References

Cullen B. (1993): The Darwinian Resurgence and the Cultural Virus Critique, Cambridge Archaeological Journal 3, p. 179-202.

Cullen B. (1994): When Megoliths Behave like a Bad Cold, British Archaeological News (19 Dec. 1994), p. 4.

Cullen B. (1995a): Living Artefact, Personal Ecosystem, Biocultural Schizophrenia: a novel synthesis, Proceedings of the Prehistoric Society 61, p. 69-90.

Cullen B. (1995b): Social Interaction and Viral Phenomena, in : Power, Sex and Tradition: the archaeology of human ancestry, Shennan S.J. & Steele J. (eds.), (Routledge, London), p. 420-432.

Cullen B. (1995c): Cultural Virus Theory and the Eusocial Pottery Assemblage, in: Darwinian Archaeologies, Maschner H. (ed.), (Plenum, New York).

Cullen B. (1996): Cultural Virus Theory: archaeology and the nature of construction (Cambridge University Press) (in press)

Dawkins R (1989) The selfish gene. New edition. Oxford university press.

Dawkins R (1993a) Viruses of the mind. In: Bo Dalbom (ed.) Dennett and His Critics: Demystifying Mind. Cambridge, Mass.: Blackwell

Dawkins R (1993b) Is religion just a disease? The Daily Telegraph, Wednesday December 15, p. 18.

Ewald P. (1990): Transmission models and the evolution of virulence. With special reference to cholera, influenza and AIDS. Human Nature. ; 2(1), p. 1-30.

Ewald P. (1994): Evolution of mutation rate and virulence among human retroviruses. Phil. Trans. R. Soc. London. ; 346, p. 333-343.

Ewald P. (1995) The evolution of virulence: A unifying link between parasitology and ecology. Journal of Parasitology 81(5) , p. 659-669.

Goodenough O. R. & R. Dawkins (1994): "The 'St Jude' mind virus", Nature 371 (September 1, 1994), p. 23-24.

Nesse R. M. & G. C. Williams (1995) Evolution and Healing: The new science of Darwinian medicine (Weidenfeld and Nicolson).

## CHAPTER 11 FROM “THE SELFISH GENE”

By  
Richard Dawkins

First published 1976; 1989 edition: Oxford University Press, ISBN 0-19-286092-5 (paperback): the best short introduction to, and the text that kicked off, the new science of **MEMETICS**, (and, also, the text where Dawkins coined the term ‘meme’).

The following key paragraph of this chapter may perhaps serve as an **abstract**:

*Examples of memes are tunes, ideas, catch-phrases, clothes fashions, ways of making pots or of building arches. Just as genes propagate themselves in the gene pool by leaping from body to body via sperms or eggs, so memes propagate themselves in the meme pool by leaping from brain to brain via a process which, in the broad sense, can be called imitation. If a scientist hears, or reads about, a good idea, he passes it on to his colleagues and students. He mentions it in his articles and his lectures. If the idea catches on, it can be said to propagate itself, spreading from brain to brain. As my colleague N.K. Humphrey neatly summed up an earlier draft of this chapter: ‘... memes should be regarded as living structures, not just metaphorically but technically.(3) When you plant a fertile meme in my mind you literally parasitize my brain, turning it into a vehicle for the meme’s propagation in just the way that a virus may parasitize the genetic mechanism of a host cell. And this isn’t just a way of talking -- the meme for, say, “belief in life after death” is actually realized physically, millions of times over, as a structure in the nervous systems of individual men the world over.’*

**The notes (1), (2) ... are from the 1989 edition. Highlights \*\* and text in square brackets are not original.**

### 11. Memes: the new replicators

So far, I have not talked much about man in particular, though I have not deliberately excluded him either. Part of the reason I have used the term ‘survival machine’ is that ‘animal’ would have left out plants and, in some people’s minds, humans. The arguments I have put forward should, *prima facie*, apply to any evolved being. If a species is to be excepted, it must be for good reasons. Are there any good reasons for supposing our own species to be unique? I believe the answer is yes.

Most of what is unusual about man can be summed up in one word: ‘culture’. I use the word not in its snobbish sense, but as a scientist uses it. Cultural transmission is analogous to genetic transmission in that, although basically conservative, it can give rise to a form of evolution. Geoffrey Chaucer could not hold a conversation with a modern Englishman, even though they are linked to each other by an unbroken chain of some twenty generations of Englishmen, each of whom could speak to his immediate neighbors in the chain as a son speaks to his father. Language seems to ‘evolve’ by non-genetic means, and at a rate which is orders of magnitude faster than genetic evolution.

Cultural transmission is not unique to man. The best non-human example that I know has recently been described by P.F. Jenkins in the song of a bird called the saddleback which lives on islands off New Zealand. On the island where he worked there was a total repertoire of about nine distinct songs. Any given male sang only one or a few of these songs. The males could be classified into dialect groups. For example, one group of eight males with neighboring territories sang a particular song called the CC song. Other dialect groups sang different songs. Sometimes the members of a dialect group shared more than one distinct song. By comparing the songs of fathers and sons, Jenkins showed that song patterns were not inherited genetically. Each young male was likely to adopt songs from his territorial neighbors by imitation, in an analogous way to human language. During most of the time Jenkins was there, there was a fixed number of songs on the island, a kind of 'song pool' from which each young male drew his own small repertoire. But occasionally Jenkins was privileged to witness the 'invention' of a new song, which occurred by a mistake in the imitation of an old one. He writes: 'New song forms have been shown to arise variously by change of notes and the combination of parts of other existing songs ... The appearance of the new form was an abrupt event and the product was quite stable over a period of years. Further, in a number of cases the variant was transmitted accurately in its new form to younger recruits so that a recognizably coherent group of like singers developed.' Jenkins refers to the origins of new songs as 'cultural mutations'.

Song in the saddleback truly evolves by non-genetic means. There are other examples of cultural evolution in birds and monkeys, but not these are just interesting oddities. It is our own species that really shows what cultural evolution can do. Language is one example out of many. Fashions in dress and diet, ceremonies and customs, art and architecture, engineering and technology, all evolve in historical time in a way that looks like highly speeded up genetic evolution, but has really nothing to do with genetic evolution. As in genetic evolution though, the change may be progressive. There is a sense in which modern science is actually better than ancient science. Not only does our understanding of the universe change as the centuries go by: it improves. Admittedly the current burst of improvement dates back to the Renaissance, which was preceded by a dismal period of stagnation, in which European scientific culture was frozen at the level achieved by the Greeks. But, as we saw in chapter 5, genetic evolution too may proceed as a series of brief spurts between stable plateau.

The analogy between cultural and genetic evolution has frequently been pointed out, sometimes in the context of quite unnecessary mystical overtones. The analogy between scientific progress and genetic evolution by natural selection has been illuminated especially by Sir Karl Popper. I want to go even further into directions which are also being explored by, for example, the geneticist L.L. Cavalli-Sforza, the anthropologist F.T. Cloak, and the ethologist J.M. Cullen.

As an enthusiastic Darwinian, I have been dissatisfied with explanations that my fellow-enthusiasts have offered for human behavior. They have tried to look for 'biological advantages' in various attributes of human civilization. For example, tribal religion has been seen as a mechanism for solidifying group identity, valuable for a pack-hunting species whose individuals rely on cooperation to catch large and fast prey. Frequently the evolutionary preconception in terms of which such theories are framed is implicitly group-selectionist, but it is possible to rephrase the theories in terms of orthodox gene selection. Man may well have spent large portions of the last several million years living in small kin groups. Kin selection and selection

in favor of reciprocal altruism may have acted on human genes to produce many of our basic psychological attributes and tendencies. These ideas are plausible as far as they go, but I find that they do not begin to square up to the formidable challenge of explaining culture, cultural evolution, and the immense differences between human cultures around the world, from the utter selfishness of the Ik of Uganda, as described by Colin Turnbull, to the gentle altruism of Margaret Mead's Arapesh. I think we have got to start again and go right back to first principles. The argument I shall advance, surprising as it may seem coming from the author of the earlier chapters, is that, for an understanding of the evolution of modern man, we must begin by throwing out the gene as the sole basis of our ideas on evolution. I am an enthusiastic Darwinian, but, I think Darwinism is too big a theory to be confined to the narrow context of the gene. The gene will enter my thesis as an analogy, nothing more.

What, after all, is so special about genes? The answer is that they are replicators. The laws of physics are supposed to be true all over the accessible universe. Are there any principles of biology that are likely to have similar universal validity? When astronauts voyage to distant planets and look for life, they can expect to find creatures too strange and unearthly for us to imagine. But is there anything that must be true of all life, wherever it is found, and whatever the basis of its chemistry? If forms of life exist whose chemistry is based on silicon rather than carbon, or ammonia rather than water, if creatures are discovered that boil to death at -100 degrees centigrade, if a form of life is found that is not based on chemistry at all but on electronic reverberating circuits, will there still be any general principle that is true of all life? Obviously I do not know but, if I had to bet, I would put my money on one fundamental principle. This is the law that all life evolves by the differential survival of replicating entities. (1) The gene, the DNA molecule, happens to be the replicating entity that prevails on our planet. There may be others. If there are, provided certain other conditions are met, they will almost inevitably tend to become the basis for an evolutionary process.

But do we have to go to distant worlds to find other kinds of replicator and other, consequent, kinds of evolution? I think that a new kind of replicator has recently emerged on this very planet. It is staring us in the face. It is still in its infancy, still drifting clumsily about in its primeval soup, but already it is achieving evolutionary change at a rate that leaves the old gene panting far behind.

The new soup is the soup of human culture. We need a name for the new replicator, a noun that conveys the idea of a unit of cultural transmission, or a unit of *imitation*. 'Mimeme' comes from a suitable Greek root, but I want a monosyllable that sounds a bit like 'gene'. I hope my classicist friends will forgive me if I abbreviate mimeme to *meme*.(2) If it is any consolation, it could alternatively be thought of as being related to 'memory', or to the French word *même*. It should be pronounced to rhyme with 'cream'.

Examples of memes are tunes, ideas, catch-phrases, clothes fashions, ways of making pots or of building arches. Just as genes propagate themselves in the gene pool by leaping from body to body via sperms or eggs, so memes propagate themselves in the meme pool by leaping from brain to brain via a process which, in the broad sense, can be called imitation. If a scientist hears, or reads about, a good idea, he passes it on to his colleagues and students. He mentions it in his articles and his lectures. If the idea catches on, it can be said to propagate itself, spreading

from brain to brain. As my colleague N.K. Humphrey neatly summed up an earlier draft of this chapter: '... memes should be regarded as living structures, not just metaphorically but technically.(3) When you plant a fertile meme in my mind you literally parasitize my brain, turning it into a vehicle for the meme's propagation in just the way that a virus may parasitize the genetic mechanism of a host cell. And this isn't just a way of talking -- the meme for, say, "belief in life after death" is actually realized physically, millions of times over, as a structure in the nervous systems of individual men the world over.'

Consider the idea of God. We do not know how it arose in the meme pool. Probably it originated many times by independent 'mutation'. In any case, it is very old indeed. How does it replicate itself? By the spoken and written word, aided by great music and great art. Why does it have such high survival value? Remember that 'survival value' here does not mean value for a gene in a gene pool, but value for a meme in a meme pool. The question really means: What is it about the idea of a god that gives it its stability and penetrance in the cultural environment? The survival value of the god meme in the meme pool results from its great psychological appeal. It provides a superficially plausible answer to deep and troubling questions about existence. It suggests that injustices in this world may be rectified in the next. The 'everlasting arms' hold out a cushion against our own inadequacies which, like a doctor's placebo, is none the less effective for being imaginary. These are some of the reasons why the idea of God is copied so readily by successive generations of individual brains. God exists, if only in the form of a meme with high survival value, or infective power, in the environment provided by human culture.

Some of my colleagues have suggested to me that this account of the survival value of the god meme begs the question. In the last analysis they wish always to go back to 'biological advantage'. To them it is not good enough to say that the idea of a god has 'great psychological appeal'. They want to know *why* it has great psychological appeal. Psychological appeal means appeal to brains, and brains are shaped by natural selection of genes in gene-pools. They want to find some way in which having a brain like that improves gene survival.

I have a lot of sympathy with this attitude, and I do not doubt that there are genetic advantages in our having brains of the kind we have. But nevertheless I think that these colleagues, if they look carefully at the fundamentals of their own assumptions, will find that they beggar just as many questions as I am. Fundamentally, the reason why it is good policy for us to try to explain biological phenomena in terms of gene advantage is that genes are replicators. As soon as the primeval soup provided conditions in which molecules could make copies of themselves, the replicators themselves took over. For more than three thousand million years, DNA has been the only replicator worth talking about in the world. But it does not necessarily hold these monopoly rights for all time. Whenever conditions arise in which a new kind of replicator *can* make copies of itself, the new replicators *will* tend to take over, and start a new kind of evolution of their own. Once this new evolution begins, it will in no necessary sense be subservient to the old. The old gene-selected evolution, by making brains, provided the 'soup' in which the first memes arose. Once self-copying memes had arisen, their own, much faster, kind of evolution took off. We biologists have assimilated the idea of genetic evolution so deeply that we tend to forget that it is only one of many possible kinds of evolution.

Imitation, in the broad sense, is how memes *can* replicate. But just as not all genes that can replicate do so successfully, so some memes are more successful in the meme-pool than others. This is the analogue of natural selection. I have mentioned particular examples of qualities that make for high survival value among memes. But in general they must be the same as those discussed for the replicators of Chapter 2: longevity, fecundity, and copying-fidelity. The longevity of any one copy of a meme is probably relatively unimportant, as it is for any one copy of a gene. The copy of the tune 'Auld Lang Syne' that exists in my brain will last only for the rest of my life. (4) The copy of the same tune that is printed in my volume of *The Scottish Student's Song Book* is unlikely to last much longer. But I expect there will be copies of the same tune on paper and in people's brains for centuries to come. As in the case of genes, fecundity is much more important than longevity of particular copies. If the meme is a scientific idea, its spread will depend on how acceptable it is to the population of individual scientists; a rough measure of its survival value could be obtained by counting the number of times it is referred to in successive years in scientific journals.(5) If it is a popular tune, its spread through the meme pool may be gauged by the number of people heard whistling it in the streets. If it is a style of women's shoe, the population memeticist may use sales statistics from shoe shops. Some memes, like some genes, achieve brilliant short-term success in spreading rapidly, but do not last long in the meme pool. Popular songs and stiletto heels are examples. Others, such as the Jewish religious laws, may continue to propagate themselves for thousands of years, usually because of the great potential permanence of written records.

This brings me to the third general quality of successful replicators: copying-fidelity. Here I must admit that I am on shaky ground. At first sight it looks as if memes are not high-fidelity replicators at all. Every time a scientist hears an idea and passes it on to somebody else, he is likely to change it somewhat. I have made no secret of my debt in the book to the ideas of R.L. Trivers. Yet I have not repeated them in his own words. I have twisted them round for my own purposes, changing the emphasis, blending them with ideas of my own and of other people. The memes are being passed on to you in altered form. This looks quite unlike the particulate, all-or-none quality of gene transmission. It looks as though meme transmission is subject to continuous mutation, and also to blending.

It is possible that this appearance of non-particulateness is illusory and that the analogy with genes does not break down. After all, if we look at the inheritance of many genetic characters such as human height or skin-coloring, it does not look like the work of indivisible and unbendable genes. If a black and a white person mate, their children do not come out either black or white: they are intermediate. This does not mean the genes concerned are not particulate. It is just that there are so many of them concerned with skin color, each one having such a small effect, that they *seem* to blend. So far I have talked of memes as though it was obvious what a single unit-meme consisted of. But of course that is far from obvious. I have said a tune is one meme, but what about a symphony: how many memes is that? Is each movement one meme, each recognizable phrase of melody, each bar, each chord, or what?

I appeal to the same verbal trick as I used in Chapter 3. There I divided the 'gene complex' into large and small genetic units, and units within units. The 'gene' was defined, not in a rigid all-or-none way, but as a unit of convenience, a length of chromosome with just sufficient copying-fidelity to serve as a viable unit of natural selection. If a single phrase of Beethoven's ninth

symphony is sufficiently distinctive and memorable to be abstracted from the context of the whole symphony, and used as the call-sign of a maddeningly intrusive European broadcasting station, then to that extent it deserves to be called one meme. It has, incidentally, materially diminished my capacity to enjoy the original symphony.

Similarly, when we say that all biologists nowadays believe in Darwin's theory, we do not mean that every biologist has graven in his brain, an identical copy of the exact words of Charles Darwin himself. Each individual has his own way of interpreting Darwin's ideas. He probably learned them not from Darwin's own writings, but from more recent authors. Much of what Darwin said is, in detail, wrong. Darwin if he read this book would scarcely recognize his own theory in it, though I hope he would like the way I put it. Yet, in spite of all this, there is something, some essence of Darwinism, which is present in the head of every individual who understands the theory. If this were not so, then almost any statement about two people agreeing with each other would be meaningless. An 'idea-meme' might be defined as an entity that is capable of being transmitted from one brain to another. The meme of Darwin's theory is therefore that essential basis of the idea which is held in common by all brains that understand the theory. The *differences* in the ways that people represent the theory are then, by definition, not part of the meme. If Darwin's theory can be subdivided into components, such that some people believe component *A* but not component *B*, while others believe *B* but not *A*, then *A* and *B* should be regarded as separate memes. If almost everybody who believes in *A* also believes in *B* -- if the memes are closely 'linked' to use the genetic term -- then it is convenient to lump them together as one meme.

Let us pursue the analogy between memes and genes further. Throughout this book, I have emphasized that we must not think of genes as conscious, purposeful agents. Blind natural selection, however, makes them behave rather \*as if\* they were purposeful, and it has been convenient, as a shorthand, to refer to genes in the language of purpose. For example, when we say 'genes are trying to increase their numbers in future gene pools', what we really mean is 'those genes that behave in such a way as to increase their numbers in future gene pools tend to be the genes whose effects we see in the world'. Just as we have found it convenient to think of genes as active agents, working purposefully for their own survival, perhaps it might be convenient to think of memes in the same way. In neither case must we get mystical about it. In both cases the idea of purpose is only a metaphor, but we have already seen what a fruitful metaphor it is in the case of genes. We have even used words like 'selfish' and 'ruthless' of genes, knowing full well it is only a figure of speech. Can we, in exactly the same spirit, look for selfish or ruthless memes?

There is a problem here concerning the nature of competition. Where there is sexual reproduction, each gene is competing particularly with its own alleles -- rivals for the same chromosomal slot. Memes seem to have nothing equivalent to alleles. I suppose there is a trivial sense in which many ideas can be said to have 'opposites'. But in general memes resemble the early replicating molecules, floating chaotically free in the primeval soup, rather than modern genes in their neatly paired, chromosomal regiments. In what sense then are memes competing with each other? Should we expect them to be 'selfish' or 'ruthless', if they have no alleles? The answer is that we might, because there is a sense in which they must indulge in a kind of competition with each other.

Any user of a digital computer knows how precious computer time and memory storage space are. At many large computer centers they are literally costed in money; or each user may be allotted a ration of time, measured in seconds, and a ration of space, measured in 'words'. The computers in which memes live are human brains. (6) Time is possibly a more important limiting factor than storage space, and it is the subject of heavy competition. The human brain, and the body that it controls, cannot do more than one or a few things at once. If a meme is to dominate the attention of a human brain, it must do so at the expense of 'rival' memes. Other commodities for which memes compete are radio and television time, billboard space, newspaper column-inches, and library shelf-space.

In the case of genes, we saw in Chapter 3 that co-adapted gene complexes may arise in the gene pool. A large set of genes concerned with mimicry in butterflies became tightly linked together on the same chromosome, so tightly that they can be treated as one gene. In Chapter 5 we met the more sophisticated idea of the evolutionarily stable set of genes. Mutually suitable teeth, claws, guts, and sense organs evolved in carnivore gene pools, while a different stable set of characteristics emerged from herbivore gene pools. Has the god meme, say, become associated with any other particular memes, and does this association assist the survival of each of the participating memes? Perhaps we could regard an organized church, with its architecture, rituals, laws, music, art, and written tradition, as a co-adapted set of mutually-assisting memes.

To take a particular example, an aspect of doctrine that has been very effective in enforcing religious observance is the threat of hell fire. Many children and even some adults believe that they will suffer ghastly torments after death if they do not obey the priestly rules. This is a peculiarly nasty technique of persuasion, causing great psychological anguish throughout the middle ages and even today. But it is highly effective. It might almost have been planned deliberately by a Machiavellian priesthood trained in deep psychological indoctrination techniques. However, I doubt if the priests were that clever. Much more probably, unconscious memes have ensured their own survival by virtue of those same qualities of pseudo-ruthlessness that successful genes display. The idea of hell fire is, quite simply, *self perpetuating*, because of its own deep psychological impact. It has become linked with the god meme because the two reinforce each other, and assist each other's survival in the meme pool.

Another member of the religious meme complex is called faith. It means blind trust, in the absence of evidence, even in the teeth of evidence. The story of Doubting Thomas is told, not so that we shall admire Thomas, but so that we can admire the other apostles in comparison. Thomas demanded evidence. Nothing is more lethal for certain kinds of meme than a tendency to look for evidence. The other apostles, whose faith was so strong that they did not need evidence, are held up to us as worthy of imitation. The meme for blind faith secures its own perpetuation by the simple unconscious expedient of discouraging rational inquiry.

Blind faith can justify anything. (7) If a man believes in a different god, or even if he uses a different ritual for worshipping the same god, blind faith can decree that he should die -- on the cross, at the stake, skewered on a Crusader's sword, shot in a Beirut street, or blown up in a bar in Belfast. Memes for blind faith have their own ruthless ways of propagating themselves. This is true of patriotic and political as well as religious blind faith.

Memes and genes may often reinforce each other, but they sometimes come into opposition. For example, the habit of celibacy is presumably not inherited genetically. A gene for celibacy is doomed to failure in the gene pool, except under very special circumstances such as we find in the social insects. But still, a *meme* for celibacy can be successful in the meme pool. For example, suppose the success of a meme depends critically on how much time people spend in actively transmitting it to other people. Any time spent in doing other things than attempting to transmit the meme may be regarded as time wasted from the meme's point of view. The meme for celibacy is transmitted by priests to young boys who have not yet decided what they want to do with their lives. The medium of transmission is human influence of various kinds, the spoken and written word, personal example, and so on. Suppose, for the sake of argument, it happened to be the case that marriage weakened the power of a priest to influence his flock, say because it occupied a large proportion of his time and attention. This has, indeed, been advanced as an official reason for the enforcement of celibacy among priests. If this were the case, it could follow that the meme for celibacy could have greater survival value than the meme for marriage. Of course, exactly the opposite would be true for a *gene* for celibacy. If a priest is a survival machine for memes, celibacy is a useful attribute to build into him. Celibacy is just a minor partner in a large complex of mutually-assisting religious memes.

I conjecture that co-adapted meme-complexes evolve in the same kind of way as co-adapted gene-complexes. Selection favors memes that exploit their cultural environment to their own advantage. This cultural environment consists of other memes which are also being selected. The meme pool therefore comes to have the attributes of an evolutionarily stable set, which new memes find it hard to invade.

I have been a bit negative about memes, but they have their cheerful side as well. When we die there are two things we can leave behind us: genes and memes. We were built as gene machines, created to pass on our genes. But that aspect of us will be forgotten in three generations. Your child, even your grandchild, may bear a resemblance to you, perhaps in facial features, in a talent for music, in the color of her hair. But as each generation passes, the contribution of your genes is halved. It does not take long to reach negligible proportions. Our genes may be immortal but the *collection* of genes that is any one of us is bound to crumble away. Elizabeth II is a direct descendant of William the Conqueror. Yet it is quite probable that she bears not a single one of the old king's genes. We should not seek immortality in reproduction.

But if you contribute to the world's culture, if you have a good idea, compose a tune, invent a sparking plug, write a poem, it may live on, intact, long after your genes have dissolved in the common pool. Socrates may or may not have a gene or two alive in the world today, as G.C. Williams has remarked, but who cares? The meme-complexes of Socrates, Leonardo, Copernicus and Marconi are still going strong.

However speculative my development of the theory of memes may be, there is one serious point which I would like to emphasize once again. This is that when we look at the evolution of cultural traits and at their survival value, we must be clear *whose* survival we are talking about. Biologists, as we have seen, are accustomed to looking for advantages at the gene level (or the individual, the group, or the species level according to taste). What we have not previously

considered is that a cultural trait may have evolved in the way that it has, simply because it is *advantageous to itself*.

We do not have to look for conventional biological survival values of traits like religion, music, and ritual dancing though these may also be present. Once the genes have provided their survival machines with brains that are capable of rapid imitation, the memes will automatically take over. We do not even have to posit a genetic advantage in imitation, though that would certainly help. All that is necessary is that the brain should be *capable* of imitation: memes will then evolve that exploit the capacity to the full.

I now close the topic of the new replicators, and end the chapter on a note of qualified hope. One unique feature of man, which may or may not have evolved memically, is his capacity for conscious foresight. Selfish genes (and, if you allow the speculation of this chapter, memes too) have no foresight. They are unconscious, blind, replicators. The fact that they replicate, together with certain further conditions means, willy nilly, that they will tend towards the evolution of qualities which, in the special sense of this book, can be called selfish. A simple replicator, whether gene or meme, cannot be expected to forgo short-term selfish advantage even if it would really pay it, in the long term, to do so. We saw this in the chapter on aggression. Even though a 'conspiracy of doves' would be better for *every single individual* than the evolutionarily stable strategy [=ESS], natural selection is bound to favor the ESS.

It is possible that yet another unique quality of man is a capacity for genuine, disinterested, true altruism. I hope so, but I am not going to argue the case one way or another, nor to speculate over its possible memic evolution. The point I am making now is that, even if we look on the dark side and assume that individual man is fundamentally selfish, our conscious foresight -- our capacity to simulate the future in imagination -- could save us from the worst selfish excesses of the blind replicators. We have at least the mental equipment to foster our long-term selfish interests rather than merely our short-term selfish interests. We can see the long-term benefits of participating in a 'conspiracy of doves', and we can sit down together to discuss ways of making the conspiracy work. We have the power to defy the selfish genes of our birth and, if necessary, the selfish memes of our indoctrination. We can even discuss ways of deliberately cultivating and nurturing pure, disinterested altruism -- something that has no place in nature, something that has never existed before in the whole history of the world. We are built as gene machines and cultured as meme machines, but we have the power to turn against our own creators. We, alone on earth, can rebel against the tyranny of the selfish replicators. (8)

## NOTES

- (1) *I would put my money on one fundamental principle ... all life evolves by the differential survival of replicating entities.*

My wager that all life, everywhere in the universe, would turn out to have evolved by Darwinian means has now been spelled out and justified more fully in my paper 'Universal Darwinism' and in the last chapter of *The Blind Watchmaker*. I show that all the alternatives to Darwinism that have ever been suggested are in principle incapable of doing the job of explaining the organized complexity of life. The argument is a general one, not based upon particular facts about life as

we know it. As such it has been criticized by scientists pedestrian enough to think that slaving over a hot test tube (or cold muddy boot) is the only method of discovery in science. One critic complained that my argument was 'philosophical', as though that was sufficient condemnation. Philosophical or not, the fact is that neither he nor anybody else has found any flaw in what I said. And 'in principle' arguments such as mine, far from being irrelevant to the real world, can be *more* powerful than arguments based on particular factual research. My reasoning, if it is correct, tells us something important about life everywhere in the universe. Laboratory and field research can tell us only about life as we have sampled it here.

## (2) *Meme*

The word meme seems to be turning out to be a good meme. It is now quite widely used and in 1988 it joined the official list of words being considered for future editions of Oxford English Dictionaries. This makes me the more anxious to repeat that my designs on human culture were modest almost to vanishing point. My true ambitions -- and they are admittedly large -- lead in another direction entirely. I want to claim almost limitless power for slightly inaccurate self-replicating entities, once they arise anywhere in the universe. This is because they tend to become the basis for Darwinian selection which, given enough generations, cumulatively builds systems of great complexity. I believe that, given the right conditions, replicators automatically band together to create systems, or machines, that carry them around and work to favour their continued replication. The first ten chapters of *The Selfish Gene* had concentrated exclusively on one kind of replicator, the gene. In discussing memes in the final chapter I was trying to make the case for replicators in general, and to show that genes were not the only members of that important class. Whether the milieu of human culture really does have what it takes to get a form of Darwinism going, I am not sure. But in any case that question is subsidiary to my concern. Chapter 11 will have succeeded if the reader closes the book with the feeling that DNA molecules are not the only entities that might form the basis for Darwinian evolution. My purpose was to cut the gene down to size, rather than to sculpt a grand theory of human culture.

## (3) ... *memes should be regarded as living structures, not just metaphorically but technically*

DNA is a self-replicating piece of hardware. Each piece has a particular structure, which is different from rival pieces of DNA. If memes in brains are analogous to genes they must be self-replicating brain structures, actual patterns of neurological wiring-up that reconstitute themselves in one brain after another. I had always felt uneasy spelling this out aloud, because we know far less about brains than about genes, and are therefore necessarily vague about what such a brain structure might actually be. So I was relieved to receive very recently a very interesting paper by Juan Delius of the University of Konstanz in Germany. Unlike me, Delius doesn't have to feel apologetic, because he is a distinguished brain scientist whereas I am not a brain scientist at all. I am delighted, therefore, that he is bold enough to ram home the point by actually publishing a detailed picture of what the neuronal hardware of a meme might look like. Among the other interesting things he does is to explore, far more searchingly than I had done, the analogy of memes with parasites; to be more precise, with the spectrum of which malignant parasites are one extreme, benign 'symbionts' the other extreme. I am particularly keen on this approach because of my own interest in 'extended phenotypic' effects of parasitic genes on host behavior (see Chapter 13 of this book and in particular chapter 12 of *The Extended Phenotype*). Delius,

by the way, emphasizes the clear separation between memes and their ('phenotypic') effects. And he reiterates the importance of coadapted meme-complexes, in which memes are selected for their mutual compatibility.

#### (4) 'Auld Lang Syne'

'Auld Lang Syne' was, unwittingly, a revealingly fortunate example for me to have chosen. This is because, almost universally, it is rendered with an error, a mutation. The refrain is, essentially always nowadays, sung as 'For the sake of auld lang syne', whereas Burns actually wrote 'For auld lang syne'. A memically minded Darwinian immediately wonders what has been the 'survival value' of the interpolated phrase, 'the sake of'. Remember that we are not looking for ways in which *people* might have survived better through singing the song in altered form. We are looking for ways in which the alteration *itself* might have been good at surviving in the meme pool. Everybody learns the song in childhood, not through reading Burns but through hearing it sung on New Year's Eve. Once upon a time, presumably, everybody sang the correct words. 'For the sake of' must have arisen as a rare mutation. Our question is, why has the initially rare mutation spread so insidiously that it has become the norm in the meme pool?

I don't think the answer is far to seek. The sibilant 's' is notoriously obtrusive. Church choirs are drilled to pronounce 's' sounds as lightly as possible, otherwise the whole church echoes with hissing. A murmuring priest at the altar of a great cathedral can sometimes be heard, from the back of the nave, only as a sporadic sussuration of 's's. The other consonant in 'sake', 'k', is almost as penetrating. Imagine that nineteen people are correctly singing 'For auld lang syne', and one person, somewhere in the room, slips in the erroneous 'For the sake of auld lang syne'. A child, hearing the song for the first time, is eager to join in but uncertain of the words. Although almost everybody is singing 'For auld lang syne', the hiss of an 's' and the cut of a 'k' force their way into the child's ears, and when the refrain comes round again he too sings 'For the sake of auld lang syne'. The mutant meme has taken over another vehicle. If there are any other children there, or adults unconfident of the words, they will be more likely to switch to the mutant form next time the refrain comes round. It is not that they 'prefer' the mutant form. They genuinely don't know the words and are honestly eager to learn. Even if those who know better indignantly bellow 'For auld lang syne' at the top of their voice (as I do!), the correct words happen to have no conspicuous consonants, and the mutant form, even if quietly and diffidently sung, is far easier to hear.

A similar case is 'Rule Britannia'. The correct second line of the chorus is 'Britannia, rule the waves'. It is frequently, though not quite universally, sung as 'Britannia rules the waves'. Here the insistently hissing 's' of the meme is aided by an additional factor. The intended meaning of the poet (James Thompson) was presumably imperative (Britannia, go out and rule the waves!) or possibly subjunctive (let Britannia rule the waves). But it is superficially easier to misunderstand the sentence as indicative (Britannia, as a matter of fact, does rule the waves). This mutant meme, then, has two separate survival values over the original form that it replaced: it sounds more conspicuous and it is easier to understand.

The final test of a hypothesis should be experimental. It should be possible to inject the hissing meme, deliberately, into the meme pool at a very low frequency, and then watch it spread

because of its own survival value. What if just a few of us were to start singing 'God saves our gracious Queen'?

(5) *If the meme is a scientific idea, its spread will depend on how acceptable it is to the population of individual scientists; a rough measure of its survival value could be obtained by counting the number of times it is referred to in successive years in scientific journals.*

[Sorry, I left this note out. It's rather long, and contains 3 figures (relatively hard to copy and put into an HTML page) that unfortunately *are* important to the note's text -- and anyway, the note is probably of interest only to settled bureaucratic scientists concerned mainly with the # of times their own publications are quoted in papers by others. :-):-) But !, since you have read so far, I think you are pretty interested in this stuff -- please consider buying the book ! I think it really would be a worthwhile investment in yourself.]

(6) *The computers in which memes live are human brains.*

It was obviously predictable that manufactured electronic computers, too, would eventually play host to self-replicating patterns of information -- memes. Computers are increasingly tied together in intricate networks of shared information. Many of them are literally wired up together in electronic mail exchange. Others share information when their owners pass floppy disks around. It is a perfect milieu for self-replicating programs to flourish and spread. When I wrote the first edition of this book I was naïve enough to suppose that an undesirable computer meme would have to arise by a spontaneous error in the copying of a legitimate program. Alas, that was a time of innocence. Epidemics of 'viruses' and 'worms', deliberately released by malicious programmers, are now familiar hazards to computer-users all over the world. [Un-original paragraph break]

My own hard disc has to my knowledge been infected in two different virus epidemics during the past year, and that is a fairly typical experience among heavy computer users. I shall not mention the names of particular viruses for fear of giving any nasty little satisfaction to their nasty little perpetrators. I say 'nasty', because their behavior seems to me morally indistinguishable from that of a technician in a microbiology laboratory, who deliberately infects the drinking water and seeds epidemics in order to snigger at people getting ill. I say 'little', because these people are mentally little. There is nothing clever about designing a computer virus. Any half-way competent programmer could do it, and half-way competent programmers are two-a-penny in the modern world. I'm one myself. I shan't even bother to explain how computer viruses work. It's too obvious.

[ Hear, hear ! .... So even Dawkins is not immune to burst off in 'flames' and in useless gratuitous morality and ethics. :-):-) Still, nevertheless, this note (bar the moralisms) does contain some interesting stuff. ]

What is less easy is how to combat them. Unfortunately some very expert programmers have had to waste their valuable time writing virus-detector programs, immunization programs and so on (the analogy with medical vaccination, by the way, is astonishingly close, even down to the injection of a 'weakened strain' of the virus). The danger is that an arms race will develop, with

each advance in virus-prevention being matched by counter-advances in new virus programs. So far, most anti-virus programs are written by altruists and supplied free of charge as a service. But I foresee the growth of a whole new profession -- splitting into lucrative specialisms just like any other profession -- of 'software doctors' on call with black bags full of diagnostic and curative floppy disks. I use the name 'doctors', but real doctors are solving natural problems that are not deliberately engineered by human malice. My software doctors, on the other hand, will be, like lawyers, solving man-made problems that should never have existed in the first place. In so far as virus-makers have any discernible motive, they presumably feel vaguely anarchistic. I appeal to them: do you really want to pave the way for a new cat-profession? If not, stop playing at silly memes, and put your modest programming talents to better use.

(7) *Blind faith can justify anything.*

I have had the predictable spate of letters from faith's victims, protesting about my criticisms of it. Faith is such a successful brainwasher in its own favor, especially a brainwasher of children, that it is hard to break its hold. But what, after all, is faith? It is a state of mind that leads people to believe something -- it doesn't matter what -- in the total absence of supporting evidence. If there were good supporting evidence then faith would be superfluous, for the evidence would compel us to believe it anyway. It is this that makes the often-parroted claim that 'evolution is a matter of faith' so silly. People believe in evolution not because they arbitrarily want to believe it but because of overwhelming, publicly available evidence.

I said 'it doesn't matter what' the faithful believe, which suggests that people have faith in entirely daft, arbitrary things, like the electric monk in Douglas Adam's delightful *Dirk Gently's Holistic Detective Agency*. He was purpose-built to do your believing for you, and very successful at it. On the day that we meet him he unshakingly believes, against all the evidence, that everything in the world is pink. I don't want to argue that things in which a particular individual has faith are necessarily daft. They may or may not be. The point is that there is no way of deciding whether they are, and no way of preferring one article of faith over another, because evidence is explicitly eschewed. Indeed the fact that true faith doesn't need evidence is held up as its greatest virtue; this was the point of my quoting the story of Doubting Thomas, the only really admirable member of the apostles.

Faith cannot move mountains (though generations of children are solemnly told the contrary and believe it). But it is capable of driving people to such dangerous folly that faith seems to me to qualify as a kind of mental illness. It leads people to believe in whatever it is so strongly that in extreme cases they are prepared to kill and die for it without the need for further justification. Keith Henson has coined the name 'memeoids' for 'victims that have been taken over by a meme to the extent that their own survival becomes inconsequential ... You see lots of these people on the evening news from such places as Belfast or Beirut'. Faith is powerful enough to immunize people against all appeals to pity, to forgiveness, to decent human feelings. It even immunizes them against fear, if they honestly believe that a martyr's death will send them straight to heaven. What a weapon! Religious faith deserves a chapter to itself in the annals of war technology, on an even footing with the longbow, the warhorse, the tank, and the hydrogen bomb.

(8) *We, alone on earth, can rebel against the tyranny of the selfish replicators.*

The optimistic tone of my conclusion has provoked skepticism among critics who feel that it is inconsistent with the rest of the book. In some cases the criticism comes from doctrinaire sociobiologists jealously protective of the importance of genetic influence. In other cases the criticism comes from a paradoxically opposite quarter, high priests of the left jealously protective of a favorite demonological icon! Rose, Kamin, and Lewontin in *Not in Our Games* have a private bogey called 'reductionism'; and all the best reductionists are also supposed to be 'determinists', preferably 'genetic determinists'.

Brains, for reductionists, are determinate biological objects whose properties produce the behaviors we observe and all the states of thought or intention we infer from that behavior & bnbsp;... Such a position is, or ought to be, completely in accord with the principles of sociobiology offered by Wilson and Dawkins. However, to adopt it would involve them in the dilemma of first arguing the innateness of much human behavior that, being liberal men, they clearly find unattractive (spite, indoctrination, etc.) and then to become entangled in liberal ethical concerns about responsibility for criminal acts, if these, like all other acts, are biologically determined. To avoid the problem, Wilson and Dawkins invoke a free will that enables us to go against the dictates of our genes if we so wish ... This is essentially a return to unabashed Cartesianism, a dualistic *deus ex machina*.

I think that Rose and his colleagues are accusing us of eating our cake and having it. Either we must be 'genetic determinists' or we believe in 'free will'; we cannot have it both ways. But -- and here I presume to speak for Professor Wilson as well as for myself -- it is only in the eyes of Rose and his colleagues that we are 'genetic determinists'. What they don't understand (apparently, though it is hard to credit) is that it is perfectly possible to hold that genes exert a statistical influence on human behavior while at the same time believing that this influence can be modified, overridden or reversed by other influences. Genes must exert a statistical influence on any behavior pattern that evolves by natural selection. Presumably Rose and his colleagues agree that human sexual desire has evolved by natural selection, in the same sense that anything ever evolves by natural selection. They therefore must agree that there have been genes influencing their sexual desires -- in the same sense as genes ever influence anything. Yet they presumably have no trouble with curbing their sexual desires when it is socially necessary to do so. What is dualist about that? Obviously nothing. And no more is it dualist for me to advocate rebelling 'against the tyranny of the selfish replicators'. We, that is our brains, are separate and independent enough from our genes to rebel against them. As already noted, we do so in a small way every time we use contraception. There is no reason why we should not rebel in a large way, too.

## KEY TO BIBLIOGRAPHY

The numbers in brackets refer to the numbered references in the bibliography. References for the body of Chapter 11 are preceded by a -, those for the notes by a >.

- Arapesh tribe (133)
- blending inheritance (69)

- Cavalli-Sforza, L.L. (32, 33)
- Cloak, F.T. (37)
- cultural evolution (20, 32, 33, 37, 62, 128)
- Darwin, C.R. (41)
- > Delius, J.D. (58)
- > determinism (47, 51, 154)
- faith (94)
- > Henson, H.K. (94)
- Humphrey, N.K. (99)
- Ik tribe (175)
- Jenkins, P.F. (101)
- > Kamin, L.J. (154)
- > Lewontin, R.C. (110, 154)
- Mead, M. (133)
- > meme (20, 58)
- > parasites (47, 89, 90, 160)
- particulate inheritance (69, 129, 153)
- Popper, K. (150, 151)
- primeval soup (144)
- > reductionism (154)
- religion (94)
- replicator (47, 48)
- > Rose, S. (154)
- saddleback (101)
- Trivers, R.L. (170, 171, 172, 173, 174)
- Turnbull, C. (175)
- > universal Darwinism (49, 50)
- Williams, G.C. (181, 183)
- > Wilson, E.O. (185)

## BIBLIOGRAPHY

20. Bonner, J.T. (1980) *The Evolution of Culture in Animals*. Princeton: Princeton University Press.
32. Cavalli-Sforza, L.L. (1971) Similarities and dissimilarities of sociocultural and biological evolution. In *Mathematics in the Archaeological and Historical Sciences* (eds. F.R. Hodson, D.G. Kendall, and P. Tautu). Edinburgh: Edinburgh University Press. pp. 553-41.
33. Cavalli-Sforza, L.L. and Feldman, M.W. (1981) *Cultural Transmission and Evolution: A Quantitative Approach*. Princeton: Princeton University Press.
37. Cloak, F.T. (1975) Is a cultural ethology possible? *Human Ecology* 3, 161-82.
41. Darwin, C.R. (1859) *The Origin of Species*. London: John Murray.
47. Dawkins, R. (1982) *The Extended Phenotype*. Oxford: W.H. Freeman.
48. Dawkins, R. (1982) Replicators and vehicles. in *Current Problems in Sociobiology* (eds. King's College Sociobiology Group). Cambridge: Cambridge University Press. pp. 45-64.
49. Dawkins, R. (1983) Universal Darwinism. In *Evolution from Molecules to Men* (ed. D.S. Bendall). Cambridge: Cambridge University Press. pp. 403-25.

50. Dawkins, R. (1986) *The Blind Watchmaker*. Harlow: Longman.
51. Dawkins, R. (1986) Sociobiology: the new storm in a teacup. In *Science and Beyond* (eds. S. Rose and L. Appignanesi). Oxford: Basil Blackwell. pp. 61-78.
58. Delius, J.D. (in press [in 1989]) Of mind memes and brain bugs: a natural history of culture. In *The Nature of Culture* (ed. W.A. Koch). Bochum: Studienlag Brockmeyer.
62. Dobzhansky, T. (1962) *Mankind Evolving*. New Haven: Yale University Press.
69. Fisher, R.A. (1930) *The Genetical Theory of Natural Selection*. Oxford: Clarendon Press.
89. Hamilton, W.D. (1998) Sex versus non-sex versus parasite. *Oikos* **35**, 282-90.
90. Hamilton, W.D. and Zuk, M. (1982) Heritable true fitness and bright birds: a role for parasites? *Science* **218**, 384-7.
94. Henson, H.K. (1985) Memes, L5 and the religion of the space colonies. *L5 News*, September 1985, pp. 5-8.
99. Humphrey, N. (1986) *The Inner Eye*. London: Faber and Faber.
101. Jenkins, P.F. (1978) Cultural transmission of song patterns and dialect development in a free-living bird population. *Animal Behaviour* **26**, 50-78.
110. Lewontin, R.C. (1983) The organism as the subject and object of evolution. *Scientia* **118**, 65-82.
128. Maynard Smith, J. (1988) *Games, Sex and Evolution*. New York: Harvester Wheatsheaf.
129. Maynard Smith, J. (1988) *Evolutionary Genetics*. Oxford: Oxford University Press.
133. Mead, M. (1950) *Male and Female*. London: Gollancz.
144. Orgel, L.E. (1973) *The Origins of Life*. London: Chapman and Hall.
150. Popper, K. (1974) The rationality of scientific revolutions. In *Problems of Scientific Revolution* (ed. R. Harré). Oxford: Clarendon Press. pp. 72-101.
151. Popper, K. (1974) Natural selection and the emergence of mind. *Dialectica* **32**, 339-55.
153. Ridley, M. (1985) *The Problems of Evolution*. Oxford: Oxford University Press.
154. Rose, S., Kamin, L.J., and Lewontin, R.C. (1984) *Not In Our Genes*. Oxford: Oxford University Press.
160. Seger, J. and Hamilton, W.D. (1988) Parasites and sex. In *The Evolution of Sex* (eds. R.E. Michod and B.R. Levin). Sunderland, Massachusetts: Sinauer. pp. 176-93.
170. Trivers, R.L. (1971) The evolution of reciprocal altruism. *Quarterly Review of Biology* **46**, 35-57.
171. Trivers, R.L. (1972) Parental investment and sexual selection. In *Sexual Selection and the Descent of Man* (ed. B. Campbell). Chicago: Aldine. pp 136-79.
172. Trivers, R.L. (1974) Parent-offspring conflict. *American Zoologist* **14**, 249-64.
173. Trivers, R.L. (1985) *Social Evolution*. Menlo Park: Benjamin/Cummings.
174. Trivers, R.L. and Hare, H. (1976) Haplodiploidy and the evolution of the social insects. *Science* **191**, 249-63.
175. Turnbull, C. (1972) *The Mountain People*. London: Jonathan Cape.
181. Williams, G.C. (1975) *Sex and Evolution*. Princeton: Princeton University Press.
183. Williams, G.C. (1975) A defense of reductionism in evolutionary biology. In *Oxford Surveys in Biology* (eds. R. Dawkins and M. Ridley), **2**, pp. 1-27.
185. Wilson, E.O. (1975) *Sociobiology: The New Synthesis*. Cambridge, Massachusetts: Harvard University Press.

# **SURVIVAL OF THE INSTITUTIONALLY FITTEST CONCEPTS**

**By**  
**Martin de Jong**

Faculty of Technology, Policy and Management  
Delft University of Technology  
Jaffalaan 5, 2628 BX Delft, The Netherlands  
Tel. +31 15 278 3433 Fax. +31 15 278 6439  
martinj@sepa.tudelft.nl

## **Abstract**

Certain arguments generated by political and administrative actors find their way to tangible policy actions, others do not. Some information is embraced by actors in institutional systems, whereas other arguments and facts can be ignored with impunity. Apparently, institutional structures constitute a persistent tendency to favour particular arguments at the cost of others. In decision making processes, i.e. processes during which a selection is to be made among various alternative policy options, institutional structures, consisting of existing decision rules and practices, operate as an information filter creating a conceptual bias.

This article spots the issue of political decision making from an evolutionary and memetics perspective, employing terms such as variation and selection, mutation and replication, information transmission and fit concepts. With the aid of the evolutionary theoretical framework, the mechanism that decides why and when certain concepts are deemed fruitful and others die is pinpointed. Examples from the field of investments in transport infrastructure in England are used to clarify the line of thought. At the end, the evolutionary perspective derived from biology is compared to well-known authors in political science to see if complementary ground can be found.

## **1 Introduction**

Certain arguments generated by political and administrative actors find their way to tangible policy actions, others do not. Some information is embraced by actors in institutional systems, whereas other arguments and facts can be ignored with impunity. Apparently, existing decision rules and practices constitute a persistent tendency to favor particular arguments at the cost of others. Schattschneider (1965), a well-known political scientist, dubbed the implicit or explicit favoring of certain actors and subjects and the waiving of others mobilization of bias; a bias built into the institutional structure addressing the finite attention of powerful actors to certain policy-matters at the expense of others.

This article introduces an additional element to mobilization of bias: concepts. Decisions are often made with the help of conceptual models indicating which information is relevant and to be searched and which is not to be paid any attention to. The political agenda not only favors specific actors and policy-issues, but also the way they are addressed. Argumentative practices

follow the bias built into these conceptual decision models and thereby determine the criteria on which policy-decisions are based. The models frame both importance and interpretation of concepts, conceptions, reasonings and data. Some arguments are generated because the institutional structure evokes them; others never even see the light of day. Some arguments are selected in and transformed into actual decisions; others are weeded out before even reaching the finishing post. Argumentative practices that have been institutionalized are part of the 'dominant institutional complex' and guide the way to the gathering of decisionally relevant information. Reasonings and data along this line are successful because they fit into the conceptual frame embraced by actors possessing the prevalent policy instruments (funds, properties, legal competences). Arguments which are either not produced, which remain latent or which are discussed but ignored when it comes to decision-making do not comply with the criteria defined by those 'in power' in the institutional system. They are relegated to the 'recessive complex', a discourse remaining mainly under the surface. Apparently, we are dealing with a selective mechanism for concepts here: the present institutional structure consists of existing political rules and argumentative practices that act as an information filter for incoming arguments and data.

In this article, the mobilization process described above is reframed in evolutionary terminology. I claim that current theories using an institutional framework can be enriched by observing them as processes in which a certain variation of concepts is created and further on selected by the institutional structure, which acts as a selective environment, as a filter concepts have to pass before having any material consequences. If we consider individual concepts as single memes (Price & Shaw 1996), conceptual frameworks as memplexes and information filters consisting of selection criteria as memetic filters as described by Dennett (1995), we may bring political science and memetics, the conceptual version of biological evolutionary theory, much closer to each other. Observing institutional development, exchange of argumentation and actor behavior as processes of evolutionary change adds an element of dynamics to political analysis that in my view is often overlooked. Evolutionary theory is by its nature engaged in describing and explaining mechanisms of mutation, change and development and is therefore a bride bringing in a very attractive dowry: a fruitful analogy. [Note 1]

Memetics, on the other hand, being a theoretically rich and promising framework badly needs empirical filling. This article aims to provide this type of evidence by linking memetics to conceptual change in political decision making.

In section 2, 'Institutional structures as information filters', I will first present the idea of information filters as it was coined by Dennett (1995). In a situation of information abundance, a framework of rules evolves to base selections on. This very same mechanism works in both the selection of good poetry and the choice between policy alternatives, as will be shown with an example derived from English transport policy.

As information filters have the form of institutional structures when policy-making is concerned, we will then dive into the definition and description of institutions. Section 3, 'Institutional structures, behavioral bias and conceptual bias', will highlight the distinction between behavioral and conceptual institutions and their interplay in the institutional structure as a whole. This look at institutional structures synthesizes current insights. From the literature we know that formal and informal power structures can exclude certain actors from the policy-process. We also know

that the use of a dominant terminology may prevent certain ideas or certain notions from being aired.

Section 3 combines these two interdependent aspects and shows how their interplay influences the gathering and selection of information and alternatives.

Section 4, 'Evolution of conceptual bias in institutional structures', gives some examples of institutional change. Particular concepts which first were sieved out by the filter later on came to be embraced by powerful actors and incorporated in the newly adapted institutional complex. This can happen because initially recessive concepts sometimes keep lingering around and 'wait' for some later opportunity. Dominant actors usually have self-interested motives for such a sudden volte face.

'Institutional selection and fit concepts', the fifth section, recapitulates the evolutionary insights in this article and tries to identify the essentials of a 'fit concept', a concept that finds its way through the filter and is thus selected in. What mechanisms decide whether a concept becomes dominant in an institutional structure?

The concluding section 6 investigates how the approach propagated here relates to other political theories about learning as developed by Sabatier (1987), Sabatier & Jenkins-Smith (1993), Hajer (1995) and Hall (1993).

## **2 Institutional structures as information filters**

In modern evolutionary biology organisms' genes or geneplexes are considered the unities which are replicated from generation to generation. Organisms are built up following the construction plan contained in their genes and after procreation copy this genetic material to their offspring. They are therefore the carriers of genetic information. Their role is not restricted to that of being only passive vehicles however. Among different organisms or populations of organisms survival and procreation rates differ. They live in a natural environment that allows certain exemplars to survive and prosper, while others fade away or starve. This natural environment leads some (populations of) organisms to survive and spread where others do not. It is a selection mechanism in the sense that only the genetic material of the successful is replicated unto following generations. The differential genetic replication rates between successful and unsuccessful lineages of carriers result in genetic evolution across generations.

Recently, this line of evolutionary thought has been enlarged to the social world as well. Just as we can observe a replication process of genetic information in the world of biology, there is a replication process of memetic information in the world of thought and learning. Ideas ('memes') living in the one actor's brain can by means of communication be replicated to other brains, and the more this occurs to a certain idea the higher is its replication rate and the greater is its success. A very intriguing thought. But if, according to this analogy, memetic information replaces genetic information and if we see an actor's brains as the carriers, what can we say about the natural selection environment? What ruthless selective filter sieves out the unsuccessful ideas? Dennett (1993) gives an impression of what a memetic filter looks like. Here is an

extensive quote telling how 'good poems' in an electronic network may be traced in a vast pool of poetic dump produced by both brilliant and mediocre poets.

*John McCarthy, one of the founders of Artificial Intelligence (...) once suggested to a humanist audience that electronic-mail networks could revolutionize the ecology of the poet. Only a handful of poets can make their living by selling poems, McCarthy noted, because poetry books are slender, expensive volumes purchased by very few individuals and libraries. But imagine what would happen if poets could put their poems on an international network, where anybody could read them or copy them for a penny, electronically transferred to the poet's royalty account. This could provide a steady source of income for many poets, he surmised. Quite independently of any aesthetic objections poets and poetry-lovers might have to poems embodied in electronic media, the obvious counter hypothesis arises from population memetics. If such a network were established, no poetry-lover would be willing to wade through thousands of electronic files filled with doggerel, looking for the good poems; there would be a niche created for various memes for poetry filters. One could subscribe, for a few pennies, to an editorial service that scanned the infosphere for good poems. Different services, with different critical standards, would flourish, as would services that screened, collected, formatted and presented the works of the best poets in slender electronic volumes which only a few would purchase. In other words, the memes for editing and criticism will find niches in any environment in the infosphere; they flourish because of the short supply and limited capacity of minds, whatever the transmission media between minds. Do you doubt this prediction? If so, I'd like to discuss framing a suitable wager with you. Here once again, as we have seen so often in evolutionary thinking, explanation proceeds by an assumption that the processes -whatever their media, and whatever the contingent zigs and zags of their particular trajectories- will home in on the forced moves and other Good Tricks in the relevant space.*

*The structure of filters is complex and quick to respond to new challenges, but of course it doesn't always 'work'. The competition among memes to break through the filters leads to an 'arms race' of ploy and counterploy, with ever more elaborate 'advertising' raised against ever more layers of selective filters (Dennett 1993: 350).*

The quality selection devices Dennett speaks of consist of a framework of search rules or heuristics. Poems meeting certain standards in terms of poet's name and reputation, employed vocabulary, length, subject, rhyme or melody and conformity to style characteristics will flow through, while lousy amateur near-sonnets or unconventional absurdist products by unknown artists won't find their way out of the electronic maze. In the end, the brain of the actor using the search device will only copy the ideas contained in the selected poems: the filter did its work.

Now suppose civil servants working at the English Department of Transport have to decide on the future of the national transport network. They are responsible for the funding of projects, programs or bids other actors submit to them. They do not have the financial means to approve of them all, so they have to make a warranted selection among them. This leaves them no other choice but to establish criteria by which they can determine what information is relevant. If they need to justify the quality of their selection they cannot escape developing rules telling which projects can be considered valuable and which cannot.

To cope with information overload the Department of Transport (**DoT**) in London has developed a cost-benefit analysis (**COBA**) to evaluate proposals submitted by local governments. As local governments can hardly levy any financial means by themselves, they rely on central funds for the realization of infrastructure projects. Well aware of this situation, DoT conscientiously applies this COBA model and has promulgated a paper explaining the functioning of the priority-setting method enumerating criteria by which the desirability of infrastructure projects can be determined. It pays specific attention to the following aspects (DoT 1989, May 1991, DoT 1994, Steer 1995, De Jong 1999):

- financial profitability
- at least 50% financial contributions from the private sector and as many private participants as possible
- improvement of run down inner city areas
- the increase or decrease in externalities caused by the project

Just like the poetry filters select poetic products that meet certain standards, COBA's criteria weed out project bids with low scores on the abovementioned criteria. Local governments' financial claims should be substantiated by a summing up of information on all of these aspects and preferably expressing the total value of projects in the monetary unit: profitability in hard pounds. There clearly is a bias in COBA. It explicitly excludes user benefits from consideration, claiming that they will be discounted in the market price. On the other hand, the rise in market value in certain areas generated by new infrastructure should be included in the calculation. These arguments set public transport at a significant disadvantage (May 1991). Benefits and costs are only calculated for criteria summed up in the method of appraisal and their outcome ratios reflect the conceptual bias.

To put it differently, DoT desires rather detailed information from local governments before they can get the money they need. To make them understand what kind of data they should deliver, it proposes an informational format stipulating the aspects they should pay attention to and how they are expected to add them all up. Using the format is not compulsory; if they wish to calculate profitability in their own way, they are free to do so. But when they submit their proposal in London, at least they are abreast of the criteria they will be judged on. Dennett's poets are in exactly the same position, they are still allowed to put experimental or amateur work on the web, but their chances for selection are minimized by the format of the filter. They may still avoid treating love and sex as the primary issues in their work, but they risk missing out lots of readers.

Rigorous and uncompromising application places dissident local authorities in a difficult position. Information they possibly attach value to will not make it through the COBA filter. As a result, their pet projects will not receive the required funding.

How do local governments generally react in a competitive environment where their local brothers are not friends but foes? They adapt to London standards or die. Authorities that have their bids sanctioned underscore the profitability of their projects, introduce an impressive list of well-known and less well-known banks, project developers and consultancy firms whose actual contribution is sometimes doubtful, ask for new roads and by-passes around the city-centre and

invest only in 'shinier kinds' of public transport such as tramways leading to the central areas. Furthermore, as DoT mentions in its 'S56-circular', it is highly irrational to extensively calculate environmental benefits which are not or not easily quantifiable. On the other hand, it is highly rational to have good contacts with the Government Offices, deconcentrated national offices in the regions. They know all the ins and outs of London decision-making and can give good advice on how and when to plug bids in and what terminology is appreciated in the capital. It is precisely this compliance which begets both success for the money-claimants and materialization of the COBA-worldview for the money-provider, a perhaps even bigger success. As one can easily understand, the effect a poetry filter has on individual poets, at least those desiring to be read, will be no different. They too will cunningly market their products and start treating each other as rivals in the infospace. [note 2]

To conclude, the complex of rules indicating what criteria or standards poems or transport investment bids should meet to be in the selection and what actor(s) have a say on this matter in policy-making is called 'institutional structure'.

### **3 Institutional structures, behavioral bias and conceptual bias**

An enormous number of volumes from different authors and various scientific disciplines have appeared on 'institutions', but little common ground has been found. Here, we follow Simon (1982, 1992), North (1981, 1990), March and Olsen (1976, 1989) and Hall (1993).

*Institutions are rules which enable actors to cope with uncertainty when making decisions. [Note 3]*

One can say that all complex choices actors need to make require an amount of information not directly available. It is not always feasible to collect this information; it is either too expensive, too time-consuming or too annoying. Nor is it always necessary; knowledge gathered in the past has often been recorded in terms of usual practices, norms of conduct, traditions or standard operating procedures. The decision rules implied in these institutions, 'correct' or not, are readily available and may save actors hugely in terms of transactions costs. If an actor has a decision model based on experience, it does not make sense to put it in doubt every time it is used: rules are very often functional and reassuring. We can clearly see the parallel with the filter mentioned in preceding section.

For the purpose of this article, showing how mobilization of conceptual bias works, the above mentioned definition of institutions is still too general. I would therefore like to make a double distinction resulting in a two by two matrix of types of institutions.

Institutions can be either formal or informal. Formal institutions are legal-verbal procedures reflecting the official 'rules of the game'. Informal institutions are the non-verbal, non-official current practices developed by players during the game.

Institutions can be either conceptual or behavioral. Conceptual institutions are rules indicating how actors should think about an impending choice; they delimit the interpretative freedom.

Behavioral institutions are rules that decide how actors should or should not act; they restrict the interactive freedom. [Note 4]

### 3.1 Four types of institutions

<i>Type of institution</i>	<i>Formal</i>	<i>Informal</i>
<b><i>Behavioral; rules for appropriate interaction</i></b>	Procedures, arrangements	Behavioral codes, Norms of conduct
<b><i>Conceptual; rules for appropriate interpretation</i></b>	Decision rules, standards, criteria	Approaches, visions, conceptions, frames

The ensemble of these various rules in a specific (policy) area is called the institutional complex or institutional structure. Arguments brought in by different actors are 'filtered' by this complex, and only those fitting in with the dominant complex embraced by the currently most powerful actors acquire influence on the 'real' decision making. It seems that concepts are 'fit' when they 'fit in'. Though this is true to large extent, it is also highly tautological. If concepts did not survive, they apparently didn't fit in. Fortunately, this is not totally true. Some concepts do not conform to the institutional standards at first, but do reach institutional hegemony after fierce conceptual struggle. Other concepts should be fit considering their possession of the properties demanded by the institutional environment, and yet don't make it if competition is too strong. [Note 5]

One important point still begs clarification. I have mentioned that institutional structures have a bias in favor of certain arguments and facts at the expense of others. I have also written that institutional complexes function as a selective environment for concepts introduced by actors. Does this mean that 'structures' are the same as 'complexes' and that 'arguments and facts' equal 'concepts'?

The answer to the first question is yes, more or less. The only slight difference is that the word 'complex' sounds more evolutionary and more open to the conceptual side of choice-making than 'structure', which has a static ring to it. Nevertheless, the terms are interchangeable here.

The answer to the second question is less straightforward. So far I have used the terms concepts, conceptions, arguments, facts and criteria rather loosely, as is unfortunately quite often also the case in how the term 'meme' is generally used. This conceptual informality will end from now on. Concepts are singular words having a particular significance. They can be seen as the smallest thinkable entity which can be replicated. Conceptions or conceptual systems are a totality, a combination of concepts forming a meaningful frame. Arguments are assumed causal connections between concepts, which provide insight into a specific question or problem. Giving an argument actually means picking out two or a few concepts from a conceptual system which have a relation with each other to address a particular issue. If a specific argument is widely accepted as being the correct way to address a specific issue at all times, it becomes a criterion: the argument is formalized. Last but not least, facts are events or data that are shown to prove

particular arguments right or wrong. They can be either qualitative or quantitative, but are only meaningful in the context of the argument for which they are used.

Toulmin (1972) wrote a wise phrase saying that concepts are micro-institutions, and institutions are macro-concepts. Disregarding the fact that institutional complexes not only consist of conceptual elements, but of behavioral elements as well, it helps to understand how institutional complexes composed of conceptual notions generally select like-minded and related concepts and occasionally incorporate new concepts within this system, which then become micro-institutions. As a result, institutional change occurs. Which micro-institutions will be part of the macro-complex and when they will obtain their place depends mainly on the relations contained in the behavioral institutions. Dominant actors have their interests and are supportive of concepts that promote these interests.

## **4 Evolution of conceptual bias in institutional structures**

The preceding sections have demonstrated that institutional structures work as filters that select certain incoming concepts in and weed others out. But the relationship between concepts and institutions can also be approached from a more dynamic angle: how previously ignored concepts become institutionalized themselves, succeeding thereby in changing argumentative links in the till then dominant institutional complex. This does not happen all that often, however; it requires political momentum. Big changes only come to be when the rigid existing complex can no longer deal with societal problems, such as abnormal congestion or exuberant pollution (Krasner 1984, 1988). Such situations put the dominant complex under strain and can lead to institutional evolution via two different paths which I will describe underneath:

### **4.1 The complete replacement of the conceptual framework by another (paradigm shift)**

In this case, the dominant actor(s) is/are confronted with increased resistance from recessive actors that no longer accept its/their conceptual and behavioral hegemony and challenge the existing the institutional structure as a whole. Following the increased behavioral strength of the recessive actors, a fierce conceptual struggle begins in which the dominant actor(s) also loses the conceptual initiative and finally has to give way.

In the South East of England counties and districts were confronted with DoT's fairly centralist attitude to infrastructure planning, as described in the section 2. Feeling frustrated and noting they were continually by-passed and competed by each other while London actively exploited their weak positions, they decided to get in touch with one another more intensely. Some local governments realized they had good roads and mediocre public transport and for others the situation was the other way around. Acknowledging their interdependence and regrouping together they were able to restructure the negotiations with London along regional lines (DoE/SERO 1994, SERPLAN 1994, Steer 1995). Consequently, the behavioral institutions started to change. After an 'initiation process' in which they had to get used to cooperating with their former rivals, they established 'package bids' with proposals for funds aimed at the whole region. They challenged London's hegemony by speaking with one voice. Besides they found active conceptual support from the London and South East Regional Planning Conference, a successful advisory organization funded by all local governments in the region that catalyzed and

facilitated mutual adaptation. It had developed knowledge in spatial planning and environmental costs and has usually been ahead in conceptual innovation. As DoT sensed that vital information was no longer spontaneously provided from local authorities to itself, but only to other regional authorities, it accepted the London and South East Regional Planning Conference as a regular institutional body. In this way the package bid approach replaced the application of COBA.

The coherence of thought implied by a package of measures does not lend itself well to a scheme-by-scheme appraisal, which is the established basis on which funding decisions are made. Funding agencies will no doubt continue to ask for a cost-benefit assessment of individual projects, but these appraisals do not answer the question inevitably posed by the package approach: what contribution to the overall strategy does this particular scheme or policy make? The aims of the overall strategy are not capable of expression in terms of quantified benefit-cost ratio targets. It follows that at the scheme, as well as strategy, level attempts will have to be made to develop new appraisal systems (Steer 1995: 204).

The basis for this investment approach was the 'Integrated Transport Studies' in which traffic trends from the past were not simply extrapolated to the future, but various effects of traffic growth were correlated with dynamic patterns of spatial evolution. In this approach the environmental and spatial effect in the region came much more to the fore, as well as a more holistic view of traffic streams. The change in the behavioral institutions, the stronger position of regional governments and their private representative, led to the institutionalization of a wholly new conceptual framework. As a consequence the old institutional complex was replaced by a new approach. DoT did try and continue to judge the package bids with the help of COBA (in effect the two 'filters' temporarily co-existed), but it gradually had to give way. In the end, the project bids for the South East were submitted to a completely changed filter.

#### **4.2 A partial change in the conceptual framework (slight mutation)**

In this less revolutionary case, dominant actors anticipate the weakening of their behavioral position and prevent it by allowing recessive actors to have some more influence by accepting a few of their recessive concepts. They integrate them into the dominant institutional structure to placate the recessive actors without having to give up the dominant paradigm. In this case, dominant actors are more forward-looking and do not have to go through the painful process of giving up funds, political positions or legal competences. They only have to consent to some conceptual adaptations in the institutional complex to accommodate and placate dissident actors.

DoT in London has for some decades applied a variant of COBA for the assessment of trunk roads and motorways. Until recently, the model utilized a purely demand based philosophy: if road congestion exceeded a certain threshold, more asphalt had to be constructed. The Transport Research Laboratory, a research organization outside the centre of the decision making network, had been criticizing the demand based line of thought since the beginning of the 1980s. They promoted the idea of so-called 'induced traffic': a higher supply of good infrastructure creates its own increased demand for transport. For this reason, when DoT had more roads built, automatically new and till then latent car drivers would suddenly appear on the road network. The conventional COBA model did not take such phenomena into account and therefore arranged for the accommodation of traffic it continually generated itself. Even though many

observers outside the Transport Research Laboratory were aware of this phenomenon, at that time the automobile lobby was too strong to push through any new concepts. The Special Advisory Commission on Trunk Roads Assessment (**SACTRA**), a high status advisory board for counseling on new motorway projects agreed, but felt the time was not yet right for a fundamental change in orientation. DoT also had many investment programs running, based on the existing model.

But by the mid nineties, popular support for the construction of more roads had declined tremendously. Environmental groups antagonized more strongly than ever against new inroads on British scenery. The conservative government decided to drastically curtail the budget for infrastructure planning and the accompanying taxes in order to increase its chances to win the elections. DoT was forced to adapt its ways and SACTRA issued an innovative report telling the nation that every new supply of infrastructure did indeed generate its own demand. DoT now voluntarily accepted the ideas in this official advice and redesigned COBA in line with the new more sober philosophy and had all running road projects recalculated and reweighed by SACTRA. Forty-nine formerly approved investments could be scrapped and cancelled. In this second case, the dominant actor was able to prevent a dramatic change in the institutional structure by allowing some conceptual mutation, but clearly within the existing conceptual framework (DoT 1994, CBI 1995).

## **5 Institutional selection and fit concepts**

Considering the exchange of arguments as concepts waging a struggle for survival in institutional complexes gives us some valuable insights in the dynamics of decision making. But there is a philosophical problem here. Do concepts, mere words that is, actually struggle? Aren't concepts only puppets on strings, used as manipulative tools by humans or actors to serve their interests? They do not have a will of their own, do they?

In biology, there is an equivalent discussion concerning the 'levels of selection' (Brandon 1988). Leading theorists in evolutionary theory do not agree among themselves about these levels of selection. Some say that selfish genes really fight for their own subsistence and that they are the ones that live through generations. In that case individual organisms are no more than helpless vehicles that replicate invisible information codes vital to evolution (Dawkins 1976). In the social science analogy, this would imply that actors are only the means through which concepts are transferred and that real selection is 'executed' only on concepts. Others claim that organisms are the ones that matter when it comes to the weeding out process, not genes. As a result, the transplantation of genetical material depends on the survival and reproduction of their carriers. Organisms are not simple vehicles, but acting and interacting entities that during their lives 'determine' whether the genes according to which codes they are built deserve a future (Sober 1984).

For the process of conceptual evolution, this would mean that the success of actor fitness and behavior make a difference. They are active carriers of ideas, not passive vehicles, so in the end the survival of ideas depends on them. The last group of theorists focuses on whole populations of organisms. Considering a much larger time span and the success of species, subspecies or geographically dispersed groups within a species, these theorists generally conclude that

populations that are bigger or have a greater variation in their total 'geneplex' are also fitter in the long run. Besides, individual organisms may have a strong physical constitution but may lack the advantages of a protective social group (Brandon 1988). [Note 6]

As all lines of thought have a point, the decision on what the level of selection is cannot be definitively answered. Hull (1988) found a way out of this dilemma by introducing the difference between interactors and replicators. He defines them as follows:

- *Replicator: an entity that passes on its structure largely intact in successive replications.*
- *Interactor: an entity that interacts as a cohesive whole with its environment in such a way that this interaction causes replication to be differential.*

*With the aid of these terms selection can be characterized succinctly as follows:*

*Selection: a process in which the differential extinction and proliferation of interactors cause the differential perpetuation of the relevant replicators.*

*Replicators and interactors are the entities that function in selection processes. Some general term is also needed for the entities that result from successive replications (Hull 1988: 408-409).*

Now jumping back to the world of concepts and actors, we can see that actors interact with their environment causing concepts and conceptual systems to be replicated differentially. Concepts only tied to 'unfit', i.e. politically weak, actors are also unfit and have a greater chance of being weeded out. For concepts used by fit actors the opposite goes: they are often selected in. The differential success of actors to spread their concepts produces an outcome known as conceptual evolution.

At the outset of this article, we saw that concepts went through an institutional structure or 'filter' before being selected or not. According to Hull's definition, (inter)actors operate in an environment that causes (conceptual) replication to be differential. In the case of biological evolution it is the natural environment or ecology that causes the replication of genetic information to be differential. When speaking of conceptual evolution, it is the institutional structure that makes the procreation of some concepts more successful than the spreading of others.

If, following Toulmin (1972), we make a distinction between the creation of a variation of concepts (A) and their selection process (B), we acquire a full picture of the mechanisms behind conceptual evolution.

## **5.1 The creation of a variation of concepts**

In the first place, actors must produce and utilize concepts before they are taken to their selective arena or filter. Concepts spread through replication from one actor to another. When concepts are replicated, small mutations or recombinations take place once in a while: their meaning changes, they acquire an extra connotation or they are applied to a different domain than previously. In other cases, whole new words are coined or invented. Conceptual mutation creates a greater variation of ideas to arise. In itself the replication of a concept from one actor to another only

generates conceptual change in the 'brain' of the receiving actor. But it may also provoke changes in this actor's perceptions and preferences and lead to different judgments of policies. This latter phenomenon leads them to act differently, carrying the germ of behavioral change.

## 5.2 The selection of concepts

The institutional complex reflects the dominant behavioral and conceptual practices developed from the past unto now and serves as a filter to the new incoming variation of concepts brought forward by actors. The behavioral institutions structure who is/are allowed to come in and perform what function with what means, the conceptual institutions are a conceptual system structuring what concepts and arguments are deemed meaningful by dominant actors. It contains the criteria applied to the selection of newly introduced concepts at a particular moment, but this too may evolve. If a specific concept has been replicated sufficiently to affect the wishes of dominant actors within the selective arena, it may institutionalize and become part of this filter. [Note 7]

If the existence of institutional structures can be established fairly easily by describing a system's written and oral rules of decision making, circumscribing what makes for the fitness of a concept is less self-evident. In the preceding sections, we saw that concepts in conformity with the criteria in the dominant complex are rather more successful than dissident ideas. Therefore, if dominant actors have developed and endorsed criteria for investment proposals demanding strong private participation and approve of schemes to promote renewal of run down inner city areas, investment proposals framed in those terms clearly have a higher probability to be selected than otherwise. But fitness is a stochastic and plural quality. It is stochastic in the sense that potentially viable concepts can fail for other reasons. They may enter the selection arena when other even stronger concepts happen to arrive as well or the actor presenting them may be out of grace. Fitness is plural or multiple in the sense that it has many sides to it. It is related to conceptual criteria as the ones just mentioned, but it is also related to behavioral criteria such as consulting friendly representatives from dominant actors for some good advice or allying with friend actors to reinforce one's position. However, generally it is the concepts' usefulness in the eyes of actors that makes them viable. Hull (1988) spoke of conceptual inclusive fitness: concepts are mainly replicated because of their ability to make individual actors understand and solve problems. In that sense they should be (1) applicable to a wide range of phenomena and (2) provide deep insight in these problems.

Also, when concepts have a structuring effect on other concepts, they will obtain a vital role in the web of a conceptual system or paradigm and attract or catch many other concepts in their webs. Concepts like 'the market', 'flexibility' or 'paradigm' seem to have these characteristics. If a sufficiently large number of actors share the opinion that a concept is needed or useful, one can say that it is 'fit'. This situation occurs in the two types of situations mentioned in the preceding section: namely (1) when a concept has already been institutionalized, or (2) when the current institutional structure is threatened or in crisis and actors recognize that the till then recessively lingering concept fills the problem-solving gap all concepts in the old complex were unable to fill.

As we saw in section 2, in some cases dependent actors have to adapt to the concepts imposed by more dominant ones. And in other cases, the ones in section 4, a whole group of not too strong actors may once in a while succeed in forcing dominant ones to retreat and by and by push in a wholly new conceptual system, realizing a 'paradigm shift' so to say. Often the combination of many fitness aspects together augments probability of survival, but some kind of 'chance' may always get in the way.

But increasing probabilities to be successful is not the same as predicting with certainty.

The evolutionary perspective may be helpful in understanding the way conceptual bias is mobilized, but framing decision making in terms of conceptual fitness does not help us to automatically forecast the success of investment proposals, even though we know they are verbally framed. That would have been too beautiful anyway. In the concluding section, the evolutionary framework presented in this article will be compared to the lines of thought by leading political theorists to see if some common or complementary ground can be found.

## 6 Conceptual learning in political theory

Even though 'conceptual evolution' is as far as I know a new term in the field of political science, this does not mean that attention to processes of policy learning has been absent. On the contrary, theorists such as Hall (1993), Sabatier (1988, 1993) and Hager (1995) have tried to describe social or political learning in terms of 'social learning' and 'paradigm shifts' (Hall), 'advocacy coalitions' and 'belief systems' (Sabatier) and 'discourse coalitions' and 'story lines' (Hager). After a short description of their central ideas, I will point out their similarities and differences with the line of thought in this article. [note 8]

### 6.1 Hall's work: Social learning and its three orders

Hall describes the process of policy learning as the acquisition of new information on the working of current policy measures and goals and the resulting changes in these measures and goals due to this new experience. He then disaggregates learning into three different types or 'orders':

First order learning concerns only adaptations in precise settings by which policy instruments and techniques are put to use.

Second order learning involves the types of policy instruments which are employed in the policy field and how possible combinations of these instruments are made.

Third order learning leads to a complete change in the overarching hierarchy of goals that guide policy in a particular field, in effect being a 'Gestalt change' or 'paradigm shift'.

Hall has rather specified ideas about this Gestalt or paradigm guiding public policymaking, and they bear a strong resemblance to the framework in this article:

*(...) policies are made within some system of ideas and standards which is comprehensible and plausible to the actors involved. More precisely, policy makers customarily work within a framework of ideas and standards that specifies not only the goals of policy and the kind of instruments that can be used to attain them, but also the very nature of the problems they are meant to be addressing. Like a Gestalt, this framework is embedded in the very terminology through which policymakers communicate about their work, and it is influential precisely because so much of it is taken for granted and not amenable to scrutiny as a whole. I am going to call this interpretive framework a policy paradigm (Hall 1993: 279).*

Hall pays attention to different levels of learning, his terminology is slightly different and in this article the emphasis on the process of conceptual evolution is more important, but there clearly seems to be some common ground.

## **6.2 Sabatier's work: advocacy coalitions holding incompatible belief systems**

Sabatier claims that at the heart of certain policy debates, between groups of actors there is not just a clash of ad hoc arguments, but an incompatibility of deeper lying belief systems consisting of diverging images of man, society and administration. He explicitly bases his term 'belief systems' on Kuhn's scientific 'paradigms'. The prime example of a belief system struggle for which it is impossible to find definitive compromises, is the debate between economy and ecology. The materialist, efficiency assumptions on which the one, and the post-materialist sustainability assumptions on which the second, thrives cannot be unisoned at an abstract level. For this reason, both will basically remain distinct belief systems that conflicting groups of actors align around.

Incompatibility in the abstract, however, does not imply that at the level of more specific policy measures, concessions from both sides are impossible. The only manner to make decisions robust enough to be effective is either to push through one of both approaches by forming a coalition of actors strong enough to take control of all necessary policy instruments or to build compromises between both coalitions. The gap between a belief system and concrete measures can be bridged by translating the central assumptions to more peripheral elaborations. In this way, choices that boost the economy may be prevented from harming the environment or decisions can be made that reduce pernicious emissions while at the same time benefiting production processes. Pollution prevention may pay, as ecological modernism has it. By incorporating peripheral ideas from the enemy belief system in their own framework, an advocacy coalition can learn.

To conclude, agreement in the abstract about belief systems will for ever remain impossible, but ephemeral junctions between them are not. As a result, the latter aspect keeps the policy process going and learning, while the first keeps on producing struggles.

## **6.3 Hajer's work: discourse coalitions producing potentially compatible story lines**

Hajer finds much of value in Sabatier's approach, but he emphatically rejects the idea that any core belief systems exist. Following a more post-modernist line, he explains how policies consist of narratives composed of competing story lines. Different actors employ different narratives and

once they have come to a textual agreement, one may say that a new, possibly innovative discourse has institutionalized. As there are no core assumptions coalitions of actors can be pinned down on, there might very well be diverging discourses, but they are not by necessity incompatible. Newly created and creative story lines introduced to the ongoing debate can displace old ones if they have a high potential of being picked up by others. However, as in Sabatier's case, some actors are more equal than others. When this line of reasoning is translated to discourses, it appears that some discourse coalitions are more potent than others. Hager defines discourse coalitions as 'assemblages of ideas, concepts and categories through which meaning is given to phenomena'. Those 'in power' having their say on what should financially or physically happen determine largely which discourses overpower others and in what manner.

To conclude, as story lines instead of core assumptions decide on the orientation of political debates there is no essential incompatibility between different approaches; they are not immobile, but evolve. No essence can be established in any narrative. Still, some discourses win, while others lose, depending on institutional structures. By inventing new attractive story lines, losers might recover their lost ground.

#### **6.4 Similarities, differences and surplus value**

Apart from certain evident conflicting points between the various approaches to policy learning, there are two remarkable points of agreement.

Firstly, they all implicitly or explicitly refer to lines of thought developed in the philosophy of science, mainly Kuhn's paradigms (1962). Hager and this author are more hesitant in following the paradigm approach, as the incompatibility of (policy) paradigms is not self-evident or probably even wrong. Words or concepts may have no essential meaning, but adapt their significance when being referred to other words or concepts, as Toulmin (1972) and Hull (1988) have clearly shown. But still, the theories of political learning all originate to an important extent in the philosophy of science.

Secondly, all authors consider both aspects of political power and the exchange of ideas as crucial to policy learning. In fact, it is precisely their intertwinement which lies at its basis. Or, as Hall beautifully summarizes:

*The important point here is that 'powering' and 'puzzling' often go together. Both are dimensions of the process whereby policy changes, especially in democratic polities, whose institutions tend to combine the two endeavors. Politicians compete for office precisely by propounding new solutions to collective problems which appeal to the electorate. Officials advance their own fortunes within the bureaucracy partly by devising new approaches to old dilemmas. The institutional arrangements designed to marry the public interest to private interests in a democracy rarely work perfectly, but they do operate so as to militate against a rigid distinction between power-based and ideas-based models of politics. The competition for power can itself be a vehicle of social learning (Hall 1993: 289).*

Nevertheless, the approach in this article has been different from or complementary to the others in two respects. It has highlighted the role of the existing institutional structure at a particular

moment in time in filtering the incoming new information. It has also focused on the survival chances of new information by taking notice of 'fitness aspects'. Both elements shed some light on the mechanisms behind conceptual evolution and may provide additional insights as compared to those already existing. On the other hand, a theory looking mainly at concepts and much less on policy instruments and techniques may be too strongly verbally oriented. In that sense, Hall's framework can add to the insights developed here.

## Acknowledgements

This paper was written after intense and intensive discussion with my colleague and friend Hans Cees Speel. I mainly owe him for his explanations of the working of 'selective environments' in biology and the 'science of memetics'. Frans van Waarden, Joop Koppenjan and Bert Enserink also gave valuable comments.

## Notes

Even though most of the current evolutionary terminology can be traced back to great masters in biology, it finds increasing application in the social sciences. Philosophical biologists or philosophers of biology such as Toulmin (1972), Dawkins (1976), Sober (1984), Hull (1988) and Dennett (1993) have deftly paved this way by developing a generic theoretical framework which in principle can be freed from its traditional biological subject matter. More recently, Price (1995) and Price & Shaw (1996) have adapted evolutionary theory to cover 'organizational memetics'. Some readers will be struck by the apparent similarities between books written by students of technology dynamics on co-evolution (Dosi 1984, Nelson & Winter 1982) and the framework embraced in this article. Others who know work on the population ecology of organizations (Hannan & Freeman 1977, 1989) will not fail to see use of a common metaphor either. In some cases, the type of analysis performed by these theoreticians may be quite similar to the view taken here. Though it would be interesting to explore the analogies between these related bodies of theory, it is beyond the scope of this article. The objective here is not to study the evolution of technology or changes in the numbers of various groups of organizations.

In the DoT example, a conceptual monoculture centered around financial profitability and market forces developed. But not in all institutional structures is just one actor able to impose its conceptual frame on all the others. In structures where funds and competences are more evenly spread across the various participants, especially lower tiers of government, other concepts and arguments play a role in discussions and negotiations and are integrated into the applied conceptual systems. In federal Germany, where subnational authorities are undeniable forces in infrastructure planning, the Standardisierte Bewertung (German equivalent of COBA) is a social cost-benefit analysis covering a full range of economic, ecological, scenery, urbanistic and political motives. Lots of less ponderable criteria such as 'spatial quality', 'transport network effects', 'intermodal connections' and 'development of economically weak regions' are explicitly mentioned and valued even when they cannot be quantified.

This definition of 'institutions' is remarkably close to how Dosi and Nelson & Winter view 'technologies'.

From the two distinctions made, the first is quite common: it can be found in what legal theorists and what sociologists view as 'rules' and in organization theory it is also a crucial duo. The second distinction between conceptual and behavioral is much less spread. Simon (1982) mentions the existence of both 'cognitive and institutional rules', whereas Campbell (NIG 1995b) discerns 'interpretative and interactive institutions'. I can very well identify with both.

Sober (1984) dedicates a full chapter of his book to the philosophical meaning of fitness. If being fit means remaining alive after the natural selection, then those who survived turned out to be fit. Such a definition denies the possibility to explain or predict fitness before the selection, leaving it no independent meaning. To be fit is to survive. However, if fitness is defined in terms of physical strength (for living beings) or semantic aptness (for words) increasing the probability of survival, it is no longer tautological. One can have a constitution well-adapted to the environment and still die, one can be a weak exemplar and yet happen to survive. To be fitter is only having a bigger probability to survive.

Theoretically speaking, we could even discern a fourth line of thought (related to the first) claiming that selection happens on populations of genes or populations of concepts (conceptual systems or paradigms). Some genes are more equal than others in the sense that they prestructure the operations of other genes. Equally we could say that some concepts within conceptual systems occupy more central places than others, because they determine the functioning these others have within the system. But whether natural or institutional selection occurs on genes/concepts as such or on geneplexes/conceptual systems was impossible to decide even for the gifted philosophers Toulmin (1972) and Hull (1973).

Institutional evolution takes more time than conceptual evolution. Mutants have to be spread and incorporated into 'brains' and then be translated to new standard and decision rules. New thinking must be transformed into new acting, which puts the power relations in existing interactive institutions at stake: innovators will always find resistance.

It is also promising to relate the concepts embraced here to the work by Hannan & Freeman (1977, 1989), Morgan (1986) and Price (1995) and Price & Shaw (1996) on the evolution of organizations. This will be done in a later article.

## References

- Brandon, R.N. (1988): The levels of selection: A hierarchy of interactors, in: Plotkin, H.C. (ed.): *The role of behavior in evolution*, MIT Press, Boston.
- Confederation of British Industry (1995): Missing links; settling national transport priorities; a CBI discussion document, London.
- Dawkins, R. (1976): *The selfish gene*, Oxford University Press, Oxford.
- Dennett, D. (1995): *Darwin's dangerous idea*, Simon and Schuster, New York.

Department of the Environment South East Regional Office (1994): Regional Planning Guidance for the South East, HMSO, London.

Department of Transport/Office of Passenger Rail Franchising/Office of the Rail Regulator (1995): Transport Report; the Government's Expenditure Plans 1995-96 to 1997-98, HMSO, London.

Department of Transport (1994): Trunk Roads in England 1994 Review, HMSO, London.

Dosi, G. (1984): *Technical change and industrial transformation*, Macmillan, London.

Hajer, M.A. (1995): *The politics of environmental discourse; ecological modernization and the policy process*, Clarendon Press, London.

Hall, P.A. (1993): Policy paradigms, social learning and the state; The case of economic policy-making in Britain, in: *Comparative politics*, April 1993, pp. 275-296.

Hannan, M.T. and J. Freeman (1977): The population ecology of organizations, *American Journal of Sociology*, pp. 929-964, Volume 82 nr 5.

Hannan, M.T. and J. Freeman (1989): *Organizational ecology*, Harvard University Press, Cambridge Massachusetts/London England.

Hull, D.L. (1988): *Science as a process; an evolutionary account of the social and conceptual development of science*, University of Chicago Press, Chicago/London.

Jong, W.M. de (1999): *Institutional transplantation; how to adopt good transport infrastructure decision making ideas from other countries*, Eburon, Delft.

Krasner, S.D. (1984): Approaches to the state; alternative conceptions and historical analysis, in: *Comparative politics*, Vol 16 no 1, pp. 223-246.

Krasner, S.D. (1988): Sovereignty. An institutional perspective, in: *Comparative political studies*, Vol 21 no 1, pp. 66-94.

Kuhn, T.S. (1962): *The structure of scientific revolutions*, University of Chicago Press.

Lakatos, I. (1978): *The methodology of scientific research programmes, Philosophical papers Volume 1*, edited by J. Worrall and G. Currie, Cambridge University Press, Cambridge/London/New York/Melbourne.

London and South East Regional Planning Conference (SERPLAN) (1994): Transport Investment Expenditure in the South East; Transport Monitoring Sub Group, A regional Profile of the South East 1993, London.

March, J. en J. Olsen (1976): *Ambiguity and choice in organizations*, Bergen Universitetsforlaget, Bergen.

March, J. en J. Olsen (1989): *Rediscovering institutions, the organizational basis of politics*, New York Free Press.

May, A.D. (1991): Integrated transport strategies: a new approach to urban transport policy formulation in the UK, in: *Transport reviews*, 1991, Vol. 11, No 3, pp. 223-247.

Morgan, G. (1986): *Images of organizations*, Sage Publications, London.

Nelson, R.R. and S.G. Winter (1982): *An evolutionary theory of economic change*, Belknap Press of Harvard University, Cambridge Massachusetts.

Netherlands Institute of Government (1995): Neo-institutionalism in government and policy, papers for the NIG-Conference, March 8-10 1995, Oosterbeek.

Netherlands Institute of Government (1995): The Role of Institutions in the Public Sector: Constraints or Resources for Policymaking, papers for the NIG-Conference, November 9-10 1995, Tilburg.

North, D.C. (1981): *Structure and change in economic history*, Norton, New York/London.

North, D.C. (1990): *Institutions, institutional change and economic performance*, Sage, London.

Price, I. (1995): Organisational memetics?: Organisational learning as a selection process, in: *Management learning* Vol. 26, pp. 299-318.

Price, I. and R. Shaw (1996): The learning organisation meme: Emergence of a management replicator; or Parrots, patterns and performance, in: Campbell, T. and V. Duperrêt-Than: Proceedings of the third ECLO Conference 1996, Copenhagen.

Sabatier, P.A. (1987): Policy oriented learning and policy change: An advocacy coalition framework, in: *Knowledge* 8, pp. 649-92.

Sabatier, P.A. and H.C. Jenkins-Smith (eds) (1993): *An advocacy coalition approach*, Boulder, Westview Press, New York.

Schattschneider, J. (1965): *The semi-sovereign people*, Dryden Press, Hindsdale.

Simon, H.A. (1982): From substantive to procedural rationality, in: H.A. Simon: *Models of bounded rationality, Volume 2: Behavioral economics and business organization*, MIT-Press, Cambridge Massachusetts.

Sober, E. (1984): *The nature of selection; evolutionary theory in philosophical focus*, University of Chicago Press, Chicago and London.

Speel, H.C.A.M. and J. Wilkins (1999): Is cultural evolution a Lamarckian process?; article to be submitted to: Journal of Biology and Philosophy.

Steer, G.M. (1995): Towards a regional transport policy; the South East's experience, in; Proceedings of the institution of Civil engineers: *Transport*, August 1995, pp. 186-197.

Toulmin, S. (1972): *Human understanding*, Clarendon Press, Oxford.

# **MEMES AND THE EXPLOITATION OF IMAGINATION**

By  
**Daniel Dennett**

*Journal of Aesthetics and Art Criticism*, 48, 127-35, Spring 1990.  
The David and Miriam Mandel Lecture  
American Society for Aesthetics, 47th Annual Meeting, October 27, 1989  
New York, NY

The general issue to be addressed in a Mandel Lecture is how (or whether) art promotes human evolution or development. I shall understand the term "art" in its broadest connotations--perhaps broader than the American Society for Aesthetics would normally recognize: I shall understand art to include all artifice, all human invention. What I shall say will *a fortiori* include art in the narrower sense, but I don't intend to draw particular attention to the way my thesis applies to it.

There are few ideas more hackneyed than the idea of the evolution of ideas. It is often said that schools of thought evolve into their successors; in the struggle for attention, the best ideas win, according to the principle of the survival of the fittest, which ruthlessly winnows out the banal, the unimaginative, and the false. Few ideas are more hackneyed--or more abused; almost no one writing about the evolution of ideas or cultural evolution treats the underlying Darwinian ideas with the care they deserve. I propose to begin to remedy that.

The outlines of the theory of evolution by natural selection are now clear: evolution occurs whenever the following conditions exist:

- variation: a continuing abundance of different elements
- heredity or replication: the elements have the capacity to create copies or replicas of themselves
- differential "fitness": the number of copies of an element that are created in a given time varies, depending on interactions between the features of that element (whatever it is that makes it different from other elements) and features of the environment in which it persists. Endnote 1

Notice that this definition, drawn from biology, says nothing specific about organic molecules, nutrition, or even life. It is a more general and abstract characterization of evolution by natural selection. As the zoologist Richard Dawkins has pointed out, the fundamental principle is "that all life evolves by the differential survival of replicating entities" Endnote 2.

The gene, the DNA molecule, happens to be the replicating entity which prevails on our own planet. There may be others. If there are, provided certain other conditions are met, they will almost inevitably tend to become the basis for an evolutionary process.

But do we have to go to distant worlds to find other kinds of replication and other, consequent, kinds of evolution? I think that a new kind of replicator has recently emerged on this very planet. It is staring us in the face. It is still in its infancy, still drifting clumsily about in its primeval

soup, but already it is achieving evolutionary change at a rate which leaves the old gene panting far behind. Endnote 3

These newfangled replicators are, roughly, ideas. Not the "simple ideas" of Locke and Hume (the idea of red, or the idea of round or hot or cold), but the sort of complex ideas that form themselves into distinct memorable units--such as the ideas of

- arch
- wheel
- wearing clothes
- vendetta
- right triangle
- alphabet
- calendar
- the Odyssey
- calculus
- chess
- perspective drawing
- evolution by natural selection
- impressionism
- Greensleeves
- "read my lips"
- deconstructionism

Intuitively these are more or less identifiable cultural units, but we can say something more precise about how we draw the boundaries--about why *D-F#-A* isn't a unit, and the theme from the slow movement of Beethoven's Seventh Symphony is: the units are the smallest elements that replicate themselves with reliability and fecundity. Dawkins coins a term for such units: *memes*--

a unit of cultural transmission, or a unit of *imitation*. 'Mimeme' comes from a suitable Greek root, but I want a monosyllable that sounds a bit like 'gene' . . . it could alternatively be thought of as being related to 'memory' or to the French word *même*.

Examples of memes are tunes, ideas, catch-phrases, clothes fashions, and ways of making pots or of building arches. Just as genes propagate themselves in the gene pool by leaping from body to body via sperm or eggs, so memes propagate themselves in the meme pool by leaping from brain to brain via a process which, in the broad sense, can be called imitation. If a scientist hears, or reads about, a good idea, he passes it on to his colleagues and students. He mentions it in his articles and his lectures. If the idea catches on, it can be said to propagate itself, spreading from brain to brain. Endnote 4

So far, no doubt, this seems to be just a crisp reworking of the standard fare about the evolution and spread of ideas, but in *The Selfish Gene*, Dawkins urges us to take the idea of meme evolution literally. Meme evolution is not just analogous to biological or genetic evolution, not just a process that can be metaphorically described in these evolutionary idioms, but a phenomenon that obeys the laws of natural selection exactly. The theory of evolution by natural

selection is neutral regarding the differences between memes and genes; these are just different kinds of replicators evolving in different media at different rates. And just as the genes for animals could not come into existence on this planet until the evolution of plants had paved the way (creating the oxygen-rich atmosphere and ready supply of convertible nutrients), so the evolution of memes could not get started until the evolution of animals had paved the way by creating a species--*homo sapiens*--with brains that could provide shelter, and habits of communication that could provide transmission media, for memes.

This is a new way of thinking about ideas. It is also, I hope to show, a good way, but at the outset the perspective it provides is distinctly unsettling, even appalling. We can sum it up with a slogan:

A scholar is just a library's way of making another library.

I don't know about you, but I am not initially attracted by the idea of my brain as a sort of dung heap in which the larvae of other people's ideas renew themselves, before sending out copies of themselves in an informational Diaspora. It does seem to rob my mind of its importance as both author and critic. Who is in charge, according to this vision--we or our memes?

There is, of course, no simple answer. We would like to think of ourselves as godlike creators of ideas, manipulating and controlling them as our whim dictates, and judging them from an independent, Olympian standpoint. But even if this is our ideal, we know that it is seldom if ever the reality, even with the most masterful and creative minds. As Mozart famously observed of his own brainchildren:

When I feel well and in a good humor, or when I am taking a drive or walking after a good meal, or in the night when I cannot sleep, thoughts crowd into my mind as easily as you would wish. Whence and how do they come? I do not know and *I have nothing to do with it.*[emphasis added] Those which please me I keep in my head and hum them; at least others have told me that I do so. Endnote 5

Mozart is in good company. Rare is the novelist who *doesn't* claim characters who "take on a life of their own"; artists are rather fond of confessing that their paintings take over and paint themselves, and poets humbly submit that they are the servants or even slaves to the ideas that teem in their heads, not the bosses. And we all can cite cases of memes that persist unbidden and unappreciated in our own minds.

The other day I was embarrassed--dismayed--to catch myself walking along humming a melody to myself: not a theme of Haydn or Brahms or Charlie Parker or even Bob Dylan: I was energetically humming: "It Takes Two to Tango"--a perfectly dismal and entirely unredeemed bit of chewing gum for the ears that was unaccountably popular sometime in the 50's. I am sure I have never in my life chosen this melody, esteemed this melody, or in any way judged it to be better than silence, but there it was, a horrible musical virus, at least as robust in my meme pool as any melody I actually esteem. And now, to make matters worse, I have resurrected the virus in many of you, who will no doubt curse me in days to come when you find yourself humming, for the first time in thirty years, that boring tune.

The first rules of memes, as it is for genes, is that replication is not necessarily for the good of anything; replicators flourish that are good at . . . replicating! -- for whatever reason. As Dawkins has put it,

A meme that made its bodies run over cliffs would have a fate like that of a gene for making bodies run over cliffs. It would tend to be eliminated from the meme-pool. . . . But this does not mean that the ultimate criterion for success in meme selection is gene survival. . . . Obviously a meme that causes individuals bearing it to kill themselves has a grave disadvantage, but not necessarily a fatal one. . . . a suicidal meme can spread, as when a dramatic and well-publicized martyrdom inspires others to die for a deeply loved cause, and this in turn inspires others to die, and so on. Endnote 6

The important point is that there is no *necessary* connection between a meme's replicative power, its "fitness" from *its* point of view, and its contribution to *our* fitness (by whatever standard we judge that). The situation is not totally desperate. While some memes definitely manipulate us into collaborating on their replication *in spite of* our judging them useless or ugly or even dangerous to our health and welfare, many--most, if we are lucky--of the memes that replicate themselves do so not just *with* our blessings, but *because of* our esteem for them. I think there can be little controversy that the following memes are, all things considered, good *from our perspective*, and not just from their own perspective as selfish self-replicators:

Such very general memes as:

- cooperation
- music
- writing
- calendars
- education
- environmental awareness
- arms-reduction
- and such particular memes as:
- The Prisoner's Dilemma
- *The Marriage of Figaro*
- *Moby Dick*
- long weekends
- returnable bottles
- the SALT Treaties
- undergraduate major

Other memes are more controversial; we can see why they spread, and why, all things considered, we should tolerate them, in spite of the problems they cause for us:

- colorization of classic films
- teaching assistants
- grade point averages

- advertising on television
- *Hustler* magazine

Still others are unquestionably pernicious, but extremely hard to eradicate:

- anti-Semitism
- hijacking airliners
- computer viruses
- spray-can graffiti

Genes are invisible; they are carried by gene-vehicles (organisms) in which they tend to produce characteristic effects ("phenotypic" effects) by which their fates are, in the long run, determined. Memes are also invisible, and are carried by meme-vehicles--pictures, books, sayings (in particular languages, oral or written, on paper or magnetically encoded, etc.) A meme's existence depends on a physical embodiment in some medium; if all such physical embodiments are destroyed, that meme is extinguished. It may, of course, make a subsequent independent reappearance--just as dinosaur genes could, in principle, get together again in some distant future--but the dinosaurs they created and inhabited would not be descendants of the original dinosaurs--or at least not any more directly than we are. The fate of memes--whether copies and copies of copies of them persist and multiply--depends on the selective forces that act directly on the physical vehicles that embody them.

Meme vehicles inhabit our world alongside all the fauna and flora, large and small. By and large they are "visible" only to the human species, however. Consider the environment of the average New York City pigeon, whose eyes and ears are assaulted every day by approximately as many words, pictures, and other signs and symbols as assault each human New Yorker. These physical meme-vehicles may impinge importantly on the pigeon's welfare, but not in virtue of the memes they carry--it is nothing to the pigeon that it is under a page of *The National Inquirer*, not *The New York Times*, that it finds a crumb.

To human beings, on the other hand, each meme-vehicle is a potential friend or foe, bearing a gift that will enhance our powers or a gift horse that will distract us, burden our memories, derange our judgment. We might compare these airborne invaders of our eyes and ears to the parasites that enter our bodies by other routes: there are the beneficial parasites such as the bacteria in our digestive systems without which we could not digest our food, the tolerable parasites, not worth the trouble of eliminating (all the denizens of our skin and scalps, for instance), and the pernicious invaders that are hard to eradicate (the AIDS virus, for instance).

So far, the meme's eye perspective may still appear to be simply a graphic way of organizing very familiar observations about the way items in our cultures affect us, and affect each other. But Dawkins suggests that in our explanations we tend to overlook the fundamental fact that "a cultural trait may have evolved in the way it has simply because it is *advantageous to itself*."  
 [Endnote 7] This is the key to answering the question of whether or not the meme meme is one we should exploit and replicate. There is an unmistakable tension between the meme's-eye view and our normal perspective on the transmission of ideas, and it is time to clarify it.

The normal view is also a normative view: it embodies a canon or ideal about which ideas we ought to "accept" or admire or approve of. (It concentrates on acceptance, rather than transmission and replication; it tends to be individualistic, not communitarian. It is epistemology and aesthetics, not communication theory.) In brief, we ought to accept the true and the beautiful.

In the normal view, the fact that an idea is deemed true or beautiful is sufficient to explain why it is accepted, and the fact that it is deemed false or ugly is sufficient to explain its rejection. These norms are *constitutive*. We require particular explanations of deviations from these norms; their status grounds the air of paradox in such aberrations as "The Metropolitan Museum of Banalities" or "The Encyclopedia of Falsehoods." There is a nice parallel to be found in physics. Aristotelian physics supposed that an object's continuing to move in a straight line required explanation, in terms of something like forces continuing to act on it. Central to Newton's great perspective shift was the idea that such rectilinear motion did *not* require explanation; only deviations from it did--accelerations. We can discern a similar difference in what requires explanation in the two views of ideas. According to the normal view, the following are virtually tautological:

- Idea *X* was believed by the people because *X* was deemed true.
- People approved of *X* because people found *X* to be beautiful.

What requires special explanation are the cases in which, in spite of the truth or beauty of an idea, it is *not* accepted, or in spite of its ugliness or falsehood it *is*. The meme's-eye view purports to be a general alternative perspective from which these deviations can be explained: what is tautological for *it* is

Meme *X* spread among the people because *X* was a good replicator.

Now there is a non-random correlation between the two; it is no accident. We would not survive unless we had a better than chance habit of choosing the memes that help us. Our meme-immunological systems are not foolproof, but not hopeless either. We can rely, as a general, crude, rule of thumb, on the coincidence of the two perspectives: by and large, the good memes are the ones that are also the good replicators.

The theory becomes interesting only when we look at the exceptions, the circumstances under which there is a pulling apart of the two perspectives; only if meme theory permits us better to understand the deviations from the normal scheme will it have any warrant for being accepted. (Note that in its own terms, whether or not the meme meme replicates successfully is strictly independent of its epistemological virtue; it might spread in spite of its perniciousness, or go extinct in spite of its virtue.)

I need not dwell on the importance of the founding memes for language, and much later, for writing, in creating the infosphere. These are the underlying technologies of transmission and replication analogous to the technologies of DNA and RNA in the biosphere. Nor shall I bother reviewing the familiar facts about the explosive proliferation of these media via the memes for movable type, radio and television, xerography, computers, fax machines, and electronic mail.

Suffice it to say that we are all well aware that we live, today, awash in a sea of paper-borne memes, breathing in an atmosphere of electronically-borne memes.

Memes now spread around the world at the speed of light, and replicate at rates that make even fruit flies and yeast cells look glacial in comparison. They leap promiscuously from vehicle to vehicle, and from medium to medium, and are proving to be virtually unquarantinable. Memes, like genes, are *potentially* immortal, but, like genes, they depend on the existence of a continuous chain of physical vehicles, persisting in the face of the Second Law of Thermodynamics. Books are relatively permanent, and inscriptions on monuments even more permanent, but unless these are under the protection of human conservators, they tend to dissolve in time. As with genes, immortality is more a matter of replication than of the longevity of individual vehicles. The preservation of the Platonic memes, via a series of copies of copies, is a particularly striking case of this. Although some papyrus fragments of Plato's texts roughly contemporaneous with him have been recently discovered, the survival of the memes owes almost nothing to such long-range persistence. Today's libraries contain thousands if not millions of physical copies (and translations) of the *Meno*, and the key ancestors in the transmission of this text turned to dust centuries ago.

Brute physical replication of vehicles is not enough to ensure meme longevity. A few thousand hard-bound copies of a new book can disappear with scarcely a trace in a few years, and who knows how many brilliant letters to the editor, reproduced in hundreds of thousands of copies, disappear into landfills and incinerators every day? The day may come when non-human meme-evaluators suffice to select and arrange for the preservation of particular memes, but for the time being, memes still depend at least indirectly on one or more of their vehicles spending at least a brief, pupal stage in a remarkable sort of meme-nest: a human mind.

Minds are in limited supply, and each mind has a limited capacity for memes, and hence there is a considerable competition among memes for entry into as many minds as possible. This competition is the major selective force in the infosphere, and, just as in the biosphere, the challenge has been met with great ingenuity.

For instance, whatever virtues (from our perspective) the following memes have, they have in common the property of having phenotypic expressions that tend to make their own replication more likely by disabling or pre-empting the environmental forces that would tend to extinguish them: the meme for *faith*, which discourages the exercise of the sort of critical judgment that might decide that the idea of faith was all things considered a dangerous idea [Endnote 8]; the meme for *tolerance* or *free speech*, the meme of including in a chain letter a warning about the terrible fates of those who have broken the chain in the past, the *conspiracy theory* meme, which has a built-in response to the objection that there is no good evidence of the conspiracy: "Of course not--that's how powerful the conspiracy is!" Some of these memes are "good" perhaps and others "bad"; what they have in common is a phenotypic effect that systematically tends to disable the selective forces arrayed against them. Other things being equal, population memetics predicts that conspiracy theory memes will persist quite independently of their truth, and the meme for faith is apt to secure its own survival, and that of the religious memes that ride piggyback on it, in even the most rationalistic environments. Indeed, the meme for faith exhibits *frequency-dependent fitness*: it flourishes particularly in the company of rationalistic memes.

Other concepts from population genetics also transfer smoothly: here is a case of what a geneticist would call *linked loci*: two memes that happen to be physically tied together so that they tend always to replicate together, a fact that affects their chances. There is a magnificent ceremonial march, familiar to us all, and one that would be much used for commencements, weddings, and other festive occasions, perhaps driving "Pomp and Circumstance" and the Wedding March from "Lohengrin" to near extinction, were it not for the fact that its musical meme is too tightly linked to its title meme, which we all tend to think of as soon as we hear the music: Sir Arthur Sullivan's unusable masterpiece, "Behold the Lord High Executioner."

This is actually just a vivid case of one of the most important phenomena in the infosphere: the mis-filtering of memes due to such linkages. We all have filters of the following sort:

Ignore everything that appears in *X*

For some people, *X* is *The National Inquirer* or *Pravda*; for others it is *The New York Review of Books*; we all take our chances, counting on the "good" ideas to make it eventually through the stacks of filters of others into the limelight of our attention.

This structure of filters is itself a meme construction of considerable robustness. John McCarthy, the founder of Artificial Intelligence (or in any event, the coiner of its name, a meme with its own, independent base in the infosphere) once suggested to a humanist audience that electronic mail networks could revolutionize the ecology of the poet. Only a handful of poets can make their living by selling their poems, McCarthy noted, because poetry books are slender, expensive volumes purchased by very few individuals and libraries. But imagine what would happen if poets could put their poems on an international network, where anybody could read them or copy them for a penny, electronically transferred to the poet's royalty account. This could provide a steady source of income for many poets, he surmised. Quite independently of any aesthetic objections poets and poetry lovers might have to poems embodied in electronic media (more to the point: poems displayed in patterns of excited phosphor dots on computer screens), the obvious counter-hypothesis arises from population memetics. If such a network were established, no poetry lover would be willing to wade through thousands of electronic files filled with doggerel, looking for the good poems; there would be a niche created for various memes for poetry-filters. One could subscribe, for a few pennies, to an editorial service that scanned the infosphere for good poems, and different services, with different critical standards, would flourish, as would services for reviewing all the different services--and services that screened, collected, formatted, and presented the works of the best poets in slender electronic volumes which only a few would purchase. The memes for editing and criticism will find niches in any environment in the infosphere; they flourish because of the short supply and limited capacity of minds, whatever the transmission media between minds.

The structure of filters is complex and quick to respond to new challenges, but of course it doesn't always "work". The competition among memes to break through the filters leads to an "arms race" of ploy and counterploy, with ever more elaborate "advertising" raised against ever more layers of selective filters.

Whether this is a good thing or a bad thing depends on your point of view. The huge arrays of garish signs that compete for our attention along commercial strips in every region of the country are the exact counterpart, in the infosphere, of the magnificent redwood forests of the biosphere; if only those redwoods could get together and agree on some sensible zoning restrictions and stop competing with each other for sunlight, they could save all the trouble of building those ridiculous and expensive trunks, stay low and thrifty shrubs, and get just as much sunlight as before! Endnote 9 In the more dignified ecology of academia, the same arms race is manifested in department letterheads, "blind refereeing", the proliferation of specialized journals, book reviews, reviews of book reviews, and anthologies of "classic works".

These filters are not even always intended to preserve the best. Philosophers might care to ask themselves, for instance, how often they are accomplices in increasing the audience for a second-rate article simply because their introductory course needs a simple-minded version of a bad idea that even the freshmen can refute. Some of the most often reprinted articles in twentieth century philosophy are famous precisely because *nobody* believes them; *everybody* can see what is wrong with them. Endnote 10

A related phenomenon in the competition of memes for our attention is *positive feedback*. In biology, this is manifested in such phenomena as the "runaway sexual selection" that explains the long and cumbersome tail of the bird of paradise or the peacock. Dawkins provides an example from the world of publishing: "Best-seller lists of books are published weekly, and it is undoubtedly true that as soon as a book sells enough copies to appear in one of these lists, its sales increase even more, simply by virtue of that fact. Publishers speak of a book "taking off", and those publishers with some knowledge of science even speak of a 'critical mass for take-off'. Endnote 11

The haven all memes depend on reaching is the human mind, but a human mind is itself an artifact created when memes restructure a human brain in order to make it a better habitat for memes. The avenues for entry and departure are modified to suit local conditions, and strengthened by various artificial devices that enhance fidelity and prolixity of replication: native Chinese minds differ dramatically from native French minds, and literate minds differ from illiterate minds. What memes provide in return to the organisms in which they reside is an incalculable store of advantages--with some Trojan horses thrown in for good measure, no doubt. Normal human brains are not all alike; they vary considerably in size, shape, and in the myriad details of connection on which their prowess depends. But the most striking differences in human prowess depend on micro-structural differences (still inscrutable to neuroscience) induced by the various memes that have entered them and taken up residence. The memes enhance each other's opportunities: the meme for education, for instance, is a meme that reinforces the very process of meme-implantation.

But if it is true that human minds are themselves to a very great degree the creations of memes, then we cannot sustain the polarity of vision with which we started; it cannot be "memes versus us" because earlier infestations of memes have already played a major role in determining *who or what we are*. The "independent" mind struggling to protect itself from alien and dangerous memes is a myth; there is, in the basement, a persisting tension between the biological imperative of the genes and the imperatives of the memes, but we would be foolish to "side with"

our genes--that is to commit the most egregious error of pop sociobiology. What foundation, then, can we stand on as we struggle to keep our feet in the memestorm in which we are engulfed? If replicative might does not make right, what is to be the eternal ideal relative to which "we" will judge the value of memes? We should note that the memes for normative concepts--for *ought* and *good* and *truth* and *beauty* are among the most entrenched denizens of our minds, and that among the memes that constitute us, they play a central role. Our existence as us, as what we as thinkers are--not as what we as organisms are--is not independent of these memes.

Dawkins ends *The Selfish Gene* with a passage that many of his critics must not have read:

We have the power to defy the selfish genes of our birth and, if necessary, the selfish memes of our indoctrination. . . . We are built as gene machines and cultured as meme machines, but we have the power to turn against our creators. We, alone on earth, can rebel against the tyranny of the selfish replicators. (p.215.)

In distancing himself thus forcefully from the oversimplifications of pop sociobiology, he somewhat overstates his case. This "we" that transcends not only its genetic creators but also its memetic creators is, I have just claimed, a myth. Dawkins himself seems to acknowledge this in his later work. In *The Extended Phenotype*, Dawkins argues for the biological perspective that recognizes the beaver's dam, the spider's web, the bird's nest as not merely *products* of the phenotype--the individual organism considered as a functional whole--but *parts* of the phenotype, on a par with the beaver's teeth, the spider's legs, the bird's wing. From this perspective, the vast protective networks of memes that we spin is as integral to our phenotypes--to explaining our competences, our chances, our vicissitudes--as anything in our more narrowly biological endowment. Endnote 12 There is no radical discontinuity; one can be a mammal, a father, a citizen, scholar, Democrat, and an associate professor with tenure. Just as man-made barns are an integral part of the barn swallow's ecology, so cathedrals and universities--and factories and prisons--are an integral part of our ecology, as are the memes without which we could not live in these environments.

*Homo sapiens* have been around for half a million years. The first serious invasion of memes began with spoken language only tens of thousands of years ago, and the second great wave, riding on the meme for writing, is considerably less than ten thousand years in progress--a brief moment in biological time. Since memetic evolution occurs on a time scale thousands of times faster than genetic evolution, however, in the period since there have been memes--only tens of thousands of years--the contributing effects of meme-structures on our constitution--on human phenotypes--vastly outweigh the effects of genetic evolution during that period. So we can answer the defining question of the Mandel Lecture with a rousing affirmative. Does art (in the broad sense) contribute to human evolution? It certainly does, in the most literal sense. In fact, since art appeared on the scene, it has virtually supplanted all other contributions to human evolution. [Endnote 13]

I would like to close with some observations on the history of the meme meme itself, and how its spread was temporarily curtailed. When Dawkins introduced memes in 1976, he described his innovation as a literal extension of the classical Darwinian Theory and so I have treated it here.

Dawkins himself, however, has since drawn in his horns slightly. In *The Blind Watchmaker* (1988), he speaks of an analogy "which I find inspiring but which can be taken too far if we are not careful." (p.196). Later in the same chapter, he says "Cultural 'evolution' is not really evolution at all if we are being fussy and purist about our use of words, but there may be enough in common between them to justify some comparison of principles." (p.216) Why did he retreat like this? Why, indeed, is the meme meme so little discussed thirteen years after *The Selfish Gene* appeared?

In *The Extended Phenotype*, Dawkins replies forcefully to the storm of criticism from sociobiologists, while conceding some interesting but inessential disanalogies between genes and memes--

memes are not strung out along linear chromosomes, and it is not clear that they occupy and compete for discrete 'loci', or that they have identifiable 'alleles'... The copying process is probably much less precise than in the case of genes . . . memes may partially blend with each other in a way that genes do not. (p.112)

But then he retreats further, apparently in the face of unnamed and unquoted adversaries:

My own feeling is that its [the meme meme's] main value may lie not so much in helping us to understand human culture as in sharpening our perception of genetic natural selection. This is the only reason I am presumptuous enough to discuss it, for I do not know enough about the existing literature on human culture to make an authoritative contribution to it. (p.112)

I submit that the meme's-eye view of what happened to the meme meme is quite obvious: "humanist" minds have set up a particularly aggressive set of filters against memes coming from "sociobiology," and once Dawkins was identified as a sociobiologist, this almost guaranteed rejection of whatever this interloper had to say about culture--not for good reasons, but just in a sort of immunological rejection. Endnote 14

But look how the meme meme has now infiltrated itself into another, less alien vehicle, a clearly identified, card-carrying academic humanist, a philosopher. In this guise--clothed in a philosopher's sort of words--will it find better chances of replication? I hope so.

My chosen role in this Mandel Lecture has been a humble one, a mere *vector*, a transmitter, with just a few embellishments and mutations, of a meme that has come to play a large role in my mind--large enough, for instance, to determine the content of this lecture. My purpose, after all, has been to create in your minds robust, aggressive copies of various memes that inhabit my mind. I am fairly confident that I have succeeded in that modest goal, and I hope, moreover, that you will forgive me for reviving "It Takes Two to Tango" and be grateful to me for passing on the meme meme.

## Endnotes

1. See, for instance, Richard Lewontin, "Adaptation," *The Encyclopedia Einaudi*, 1980, Milan; Robert Brandon, "Adaptation and Evolutionary Theory," *Studies in the History and Philosophy*

*of Science*, 1978, 9, pp. 181-206, both reprinted in E. Sober, ed., *Conceptual Issues in Evolutionary Biology*, 1984, Cambridge, MA: MIT Press.

2. Richard Dawkins, *The Selfish Gene*, Oxford: Oxford University Press, 1976, p.206.

3. ibid

4. ibid

5. Peter Kivy informed me after the Mandel Lecture that this oft-quoted passage is counterfeit--not Mozart at all. I found it in Jacques Hadamard's classic study, *The Psychology of Inventing in the Mathematical Field*, Princeton Univ. Press., 1949, p. 16, and first quoted it myself in "Why the Law of Effect Will Not Go Away," *Journal of the Theory of Social Behaviour*, V, 1975, pp.169-87, reprinted in my book, *Brainstorms*, Cambridge, MA: MIT Press/A Bradford Book, 1978. I persist in quoting it here, in spite of Kivy's correction, because it not only expresses but exemplifies the thesis that memes, once they exist, are independent of authors and critics alike. Historical accuracy is important (which is why I have written this footnote), but the passage so well suits my purposes that I am choosing to ignore its pedigree. I might not have persisted in this, had I not encountered a supporting meme the day after Kivy informed me: I overheard a guide at the Metropolitan Museum of Art, commenting on the Gilbert Stuart portrait of George Washington: "This may not be what George Washington looked like then, but this is what he looks like now."

6. Richard Dawkins, *The Extended Phenotype*, Oxford: W.H. Freeman, 1982, p.110-11.

7. *The Selfish Gene*, p.214.

8. *The Selfish Gene*, p. 212.

9. This, the "tragedy of the commons," deserves a more careful treatment than I can offer on this occasion. Note too that I am submerging a complication that properly should bring our discussion full circle, back to the ideas of Adam Smith about *economic* competition that first inspired Darwin. The competition of the billboards is competition for our attention, but the ulterior goal of acquiring our attention is the seller's goal of acquiring our money, not the meme's goal of replicating itself. The academic examples are not independent of economics, of course, but economics plays a less dominant role, as was ironically acknowledged on a T-shirt worn by a member of the audience at the Mandel Lecture: "Philosophy: I'm in it for the money."

10. The confirmation of this claim is left as an exercise for the reader. Among the memes that structure the infosphere and hence affect the transmission of other memes are the laws of libel.

11. *The Blind Watchmaker*, London: Longman Scientific, 1986, p.219. Dawkins' discussion of these complex phenomena, in the chapter "Explosions and Spirals" (pp. 195-220), is a *tour de force* of explanatory clarity and vividness.

12. In several recent essays I have expanded on the claim that the very structure of our minds is more a product of culture than of the neuroanatomy we are born with: "Julian Jaynes' Software

Archeology," in *Canadian Psychology*, **27**, 1986, pp.149-54; "The Self as the Center of Narrative Gravity," (originally published as "Why we are all novelists," *Times Literary Supplement*, Sept. 16-22, 1988, p.1029), forthcoming in F. Kessel, P. Cole, D. Johnson, eds., *Self and Consciousness: Multiple Perspectives*, Hillsdale, NJ: Erlbaum; "The Evolution of Consciousness," forthcoming in *The Reality Club*, volume 3; and "The Origins of Selves," forthcoming in *Cogito*. See also Nicholas Humphrey and Daniel Dennett, "Speaking For Our Selves: An Assessment of Multiple Personality Disorder," *Raritan*, **9**, 1989, pp.68-98.

13. Those who are familiar with the Baldwin Effect will recognize that art contributes not merely to the fixing of phenotypic plasticity, but can *thereby* change the selective environment and hence hasten the pace of genetic evolution. See my discussion in "The Evolution of Consciousness," *oc.cit.*, and Jonathan Schull, "Are Species Intelligent?" forthcoming in *Behavioral and Brain Sciences*.

14. A striking example of the vituperative and uncomprehending dismissal of Dawkins by a humanist who identifies him as a sociobiologist is found in Mary Midgley, "Gene Juggling," *Philosophy*, **54**, 1979, pp.439-58, an attack so wide of the mark that it should not be read without its antidote: Dawkins' response, "In Defence of Selfish Genes," *Philosophy*, **56**, 1981, pp. 556-73.

# **MEMES: MYTHS, MISUNDERSTANDINGS, AND MISGIVINGS**

By  
**Daniel C. Dennett**

DRAFT for Chapel Hill, October 1998

Richard Dawkins introduced the concept of memes in his 1976 book, *The Selfish Gene*, and the reception for many years was chilly. Then, recently, thanks in part to some energetic campaigning by me on behalf of the meme meme (mainly in *Consciousness Explained* and *Darwin's Dangerous Idea*), the friends of memes began to come out of the woodwork, and a number of books and articles about memes (of varying quality) have begun attracting a second look. Currently the internet blooms with dozens of websites proclaiming the birth of the new science of memetics. Most of this is simply awful, but that should not surprise us. As Sturgeon's Law reminds us, 95% of everything is crap. The hard part--especially during these early days of proto-memetics--is to identify the 5% that is actually good. Sturgeon's Law also suggests, of course, that 95% of the *criticism* of memes and memetics is also crap, so we needn't waste our time rebutting every silly, anxiety-driven objection. My talk will be an attempt to focus on some of the more attention-worthy issues.

## **1. Perspectives on Cultural Evolution<sup>(1)</sup>**

When one says that cultures evolve, this can be taken as a truism, or as asserting one or another controversial, speculative, unconfirmed theory. Consider a cultural inventory at time *t*: it includes all the languages, practices, ceremonies, edifices, methods, tools, myths, music, art, and so forth that compose a culture. Over time, the inventory changes. Some items disappear, some multiply, some merge, some change. (When I say some change, I mean to be neutral at this point about whether this amounts to their being replaced by similar items, or their undergoing a transformation.) A verbatim record of this history would not be science; it would be a data base. That is the truism: cultures evolve over time. Now the question remains: how are we to explain the patterns found in that data base? Are there any good theories or models of cultural evolution?

The traditional model to be found in most accounts by historians and anthropologists treats culture as composed of goods, possessions of the people, who husband them in various ways, wisely or foolishly. They carefully preserve their traditions of fire-lighting, house-building, speaking, counting, justice, etc. They trade cultural items as they trade other goods. And of course some cultural items (wagons, pasta, recipes for chocolate cake, etc.) are definitely goods, and we can plot their trajectories using the tools of economics. The people, on this model, are seen as having an autonomous or independent rationality; deprive a person of his goods, and he stands there, naked but rational and full of informed desires. When he clothes himself and arms himself and equips himself with goods, he increases his powers, complicates his desires, etc.

On this way of thinking, the relative "replicative" power of various cultural goods is measured in the marketplace of cost-benefit calculations performed by the people. If Coca Cola bottles proliferate around the world, it is because more and more people prefer to buy a Coke. Advertising may fool them. But then we look to the advertisers, or those who have hired them, to

find the relevant loci of values for our calculations. *Cui bono?* Who benefits? The purveyors of the goods, and those they hire to help them.

Biologists, too, can often make sense of the evolution (in the neutral sense) of features by treating them as goods: one's food, one's nest, one's burrow, one's territory, one's mate[s], one's time and energy. Cost-benefit analyses shed light on the husbandry engaged in by the members of the different species inhabiting some shared environment.<sup>(2)</sup> Not every "possession" is considered a good, however; one's accompanying flies and fleas, the dirt and grime that accumulates on one's body, are of no value, or of negative value, for instance. One's symbionts are not normally considered as goods by biologists, except when the benefits derived from them (by whom?) are manifest.

This perspective is not uniformly illuminating, nor is it obligatory. I would like to suggest that both biologists and economists (and other social scientists) can benefit from adopting a different vantage point on some of these phenomena, one which quite properly gives pride of place to the *Cui bono* question, which can provide alternative answers that are often overlooked. This is Dawkins' meme's-eye point of view, which recognizes--and takes seriously--the possibility that cultural entities may evolve according to selectional regimes that make sense only when the answer to the *Cui bono* question is that it is the cultural items themselves that benefit from the adaptations they exhibit.<sup>(3)</sup>

Dawkins' theory of memes, as briefly sketched in a single chapter of *The Selfish Gene* (1976, but see also Dawkins, 1993), is hardly a theory at all, especially compared to the models of cultural evolution developed by other biologists, such as Cavalli-Sforza and Feldman (1981), Lumsden and Wilson (1981), and Boyd and Richerson (1985). Unlike these others, Dawkins offers no formal development, no mathematical models, no quantitative predictions, no systematic survey of relevant empirical findings. But Dawkins does present an idea that is overlooked by all the others, and it is, I think, a most important idea. It is the key to understanding how we can be not just guardians and transmitters of culture, but cultural entities ourselves--all the way in.

Whenever costs and benefits are the issue we need to ask *Cui bono?* A benefit by itself is not explanatory; a benefit in a vacuum is indeed a sort of mystery; until it can be shown how the benefit actually redounds to enhance the replicative power of a replicator, it just sits there, alluring, perhaps, but incapable of explaining anything.

We see an ant laboriously climbing up a stalk of grass. Why is it doing that? Why is that adaptive? What good accrues to the ant by doing that? That is the wrong question to ask. No good accrues to the ant; its brain has been invaded by a fluke (*Dicrocoelium dendriticum*), one of a gang of tiny parasites that need to get themselves into the intestines of a sheep in order to reproduce (Ridley, 1995, p258). (Salmon swim up stream, these parasitic worms drive ants up grass stalks, to improve their chances of being ingested by a passing sheep.) The benefit is not to the reproductive prospects of the ant but the reproductive prospects of the fluke.<sup>(4)</sup>

Dawkins points out that we can think of cultural items, memes, as parasites, too. Actually, they are more like a simple virus than a worm. Memes are supposed to be analogous to genes, the replicating entities of the cultural media, but they also have vehicles, or phenotypes; they are like

not-so-naked genes. They are like viruses (Dawkins, 1993). As with viruses, there is a phenotype/genotype distinction, but just barely. Basically, a virus is just a string of DNA (or RNA) with attitude. And similarly, a meme is an information-packet (the information, not the vehicle) with attitude--with some phenotypic clothing that has differential effects in the world that thereby influence its chances of getting replicated.

And in the domain of memes, the ultimate beneficiary, the beneficiary in terms of which the final cost-benefit calculations must apply is: the meme itself, not its carriers. This is not to be read as itself a bold empirical claim, ruling out (for instance) the role of individual human agents in devising, appreciating and securing the spread and prolongation of cultural items. It is rather a proposal that we adopt a perspective or point of view, from which a wide variety of different empirical claims can be compared, and the evidence for them considered in a neutral setting, a setting that does not prejudge these hot-button questions.

In the analogy with the fluke, we are invited to consider a meme as like a parasite which commandeers an organism for its own replicative benefit, but we should remember that symbionts can be classified into three fundamental categories:

- *Parasites*, whose presence lowers the fitness of their host
- *Commensals*, whose presence is neutral (though, as the etymology reminds us, they "share the same table"); and
- *Mutualists*, whose presence enhances the fitness of both host and guest

Since these varieties are arrayed along a continuum, the boundaries between them need not be too finely drawn; just where benefit drops to zero or turns to harm is not something to be directly measured by any practical test, though we can explore the consequences of these turning points in models.

The main point to note is that we should expect memes to come in all three varieties, too. This means, for instance, that it is a mistake to assume that the "cultural selection" of a cultural trait is always "for cause"--always because of some perceived (or even misperceived) benefit it provides to the host. We can always ask if the hosts, the human agents that are the vectors, perceive some benefit and (for that reason, good or bad) assist in the preservation and replication of the cultural item in question, but we must be prepared to entertain the answer that they do not. In other words, we must consider as a real possibility the hypothesis that the human hosts are, individually or as a group, either oblivious to, or agnostic about, or even positively dead set against, some cultural item, which nevertheless is able to exploit its hosts as vectors.

The most familiar cases of cultural transmission and evolution discussed are innovations that are obviously of some direct or indirect benefit to the Darwinian--that is, genetic--fitness of the host. A better fishhook catches more fish, feeds more bellies, makes for more surviving grandchildren, etc. The only difference between stronger arms and a better fishhook in the (imagined) calculation of impact on fitness is that the stronger arms might be--might be--passed on quite directly through the germ line, while the fishhook definitely must be culturally transmitted. (The stronger arms could be culturally transmitted as well, of course. A tradition of body-building, for instance, could explain why there was very low [genetic] heritability for strong adult arms, and

yet a very high rate of strong adult arms in a population.) But however it might be that strong arms or fishhooks are transmitted, they are typically supposed to be a good bargain from the perspective of genetic fitness. The bargain might, however, be myopic--only good in the short run. After all, even agriculture, in the long run, may be a dubious bargain if what you are taking as your *summum bonum* is Darwinian fitness (see Diamond, 1997, for fascinating reflections on the uncertain benefits of abandoning the hunter-gatherer lifestyle). What alternatives are there?

First, we need to note that in the short run (evolutionarily speaking--that is, from the perspective of a few centuries or even millennia) something might flourish independently of whether it was of actual benefit to genetic fitness, but strongly linked to whether it was of *apparent* benefit to genetic fitness. Even if you think that Darwinian fitness enhancement is the principle driving engine of cultural evolution, you have to posit some swifter, more immediate mechanism of retention and transmission. It's not hard to find one. Cultural items may exploit machinery that had earned its keep in the past by embodying a fitness-enhancing set of preferences. We are genetically endowed with a quality space in which some things feel good and some things don't, and we tend to live by the rule: *if it feels good, keep it.* This rough and ready rule can be tricked, of course. The sweet tooth is a standard example. The explosion of cultural items--artifacts, practices, recipes, patterns of agriculture, trade routes--that depend quite directly on the exploitation of the sweet tooth has probably had a considerable net negative effect on human genetic fitness. Notice that explaining the emergence of these cultural items by citing their "apparent" benefit to genetic fitness does not in any way commit us to the (preposterous) claim that people think (mistakenly) that they are enhancing their genetic fitness by acquiring and consuming sugar. The rationale is not theirs, but Mother Nature's. They just go with what they like.

Still, given what they like, they choose rationally, and indeed ingeniously and often with impressive foresight, how to obtain what they like. This is still the traditional model of cultural evolution, with agents husbanding their goods in order to maximize what they prefer--and getting their preferences quite directly from their genetic heritage. A more interesting possibility is acquiring new preferences that are themselves culturally transmitted symbionts of one sort or another. Each will have to bootstrap itself into the memosphere by exploiting some pre-established preference, but this recursive process, which can proceed at breakneck speed relative to the glacial pace of genetic evolution, can transform human agents indefinitely far away from their genetic beginnings. In an oft-quoted passage, E. O. Wilson claimed otherwise: The genes hold culture on a leash. The leash is very long, but inevitably values will be constrained in accordance with their effects on the human gene pool. (Wilson, 1978, p167)

This leash, I am claiming, is indefinitely long, in the sense that the constraints Wilson speaks of can be so co-opted, exploited, and obtunded in a recursive cascade of cultural products and meta-products that it is not clear that there are any points in imaginable cultural design space that could not, in principle, be occupied by some product that could ultimately be traced back, via Wilson's leash of historical processes, to the genes. Many of these imaginable points would no doubt be genetic culs-de-sac (*H. sapiens* would sooner or later go extinct as a result of occupying those points), but this is no barrier to their evolving in the swift time of cultural history.<sup>(5)</sup>

Not only can we acquire tastes; we can acquire meta-tastes. That is, we can discover in the

culture, and thereupon adopt, a taste for "cultivating" further acquired tastes, and so forth. At each stage we can anticipate finding parasites, commensals and mutualists--but we can classify these only by asking the *Cui bono?* question against a new background and making one local determination or another. One person's scholarly connoisseurship is another person's addiction to trash. Meta-memes for "traveling" or "being a collector" or "having a hobby" or "educating oneself" can themselves be viewed as either exploiters or enhancers of the pre-established personal (no longer genetic) preferences. It is interesting that in common parlance we often call our preferences "weaknesses,"--as in "I have a weakness for strong cheese (or puns or redheads)"--deftly implying a standard to which in the same breath we deny any personal allegiance.

And this, then, is the main point I wanted to emphasize in Dawkins' vision. The memes that proliferate will be the memes that replicate by hook or by crook. Think of them as entering the brains of culture members, making phenotypic alterations thereupon, and then submitting themselves to the great selection tournament--not the Darwinian genetic fitness tournament (life is too short for that) but the Dawkinsian meme-fitness tournament. It is their fitness as memes that is on the line, not their host's genetic fitness, and the environments that embody the selective pressures that determine their fitness are composed in large measure of other memes.

Why do their hosts put up with this? Why should the overhead costs of establishing a whole new system of differential reproduction be borne by members of *H. sapiens*? Note that the question to be asked and answered here is parallel to the question we ask about any symbiont-host relationship: why do the hosts put up with it? And the short answer is that it is too costly to eradicate, but this just means that the benefits accruing to the machinery that is being exploited by the parasites are so great that keeping the machinery and tolerating the parasites (to the extent that they are tolerated) has so far been the best deal available. And whether or not in the long run (millions of years) this infestation will be viewed as mutualism or commensalism or parasitism, in the short run (the last few millennia) the results have been spectacular: the creation of a new biological type of entity: a person.

I like to compare this development to the arrival of the eukaryotes more than a billion years ago. Relatively simple prokaryotes got invaded by some of their neighbors, and the resulting endosymbiotic teams were more fit, and prospered, enabling a biological revolution. The eukaryotes, living alongside their prokaryotic cousins, but enormously more complex, versatile and competent, opened up the design space of multi-cellular organisms. Similarly, the emergence of culture-infected hominids has opened up yet another region of hitherto unoccupied and untraversable design space. We live alongside our animal cousins, but we are enormously more complex, versatile and competent. And by joining forces with our memes, we create new candidates for the locus of benefit, new answers to *Cui bono?*

## 2. Two all too standard Objections

One often hears it said that the ways in which cultural entities evolve are profoundly un-Darwinian. Two claims in particular, are often presented as if they carried the day: cultural evolution, unlike Darwinian evolution, is "Lamarckian," and cultural evolution, unlike Darwinian evolution, is replete with "horizontal transmission"--that is to say, design elements

can hop freely from lineage to lineage, not bound by the requirements of heredity. Once reptiles and mammals have gone their separate ways, reptile innovations cannot jump to mammals, but only to descendant reptiles, but this restriction does not exist in cultural evolution. I have sometimes wondered why we don't hear more about a third disanalogy: cultural ideas don't reproduce sexually--mama and papa ideas getting it on to make little baby ideas of both genders. Probably we don't hear it because it would wear its disingenuousness on its sleeve--a lazy (or desperate) stab at something that would excuse one from having to think further about the prospects of a Darwinian account of culture. *Sexual* reproduction is not, after all, an obligatory element of Darwinian evolution; surely 99% of all the Darwinian evolution that has ever occurred on this planet was among asexually reproducing replicators, and however large sexuality looms now, it is itself an evolved feature, not a precondition for Darwinian evolution. So the absence of sexual reproduction in the memosphere is no challenge to neo-Darwinian explanation. But exactly the same point can be made about the purported disanalogies of Lamarckianism and horizontal transmission or anastomosis (lineage-joining).

Let's consider Lamarckianism first. Neo-Darwinian orthodoxy, since Weissman, declares that characteristics acquired through use cannot be transmitted genetically to one's progeny. Darwin himself, notoriously, was quite happy to countenance this feature of Lamarckianism, but he has long been deemed in error. Weissman's distinction between germ line--roughly, eggs and sperm-- and somatic line cells--all the rest--has proven itself over and over, and the doctrine that there are no avenues by which somatic line innovations could enter the germ line is indeed a textbook verity, although various exotic possibilities have been seriously discussed in the literature, and arguably exist in some restricted quarters. But notice that this, the orthodox, way of identifying Lamarckian phenomena (as things that don't happen) applies crisply only to multi-cellular organisms. What counts as a Lamarckian phenomenon in the world of bacteria, archaea, or in the world of viruses? In the case of a virus, which I have described as just a string of DNA with attitude, the line between soma and germ line is non-existent. Something that changes the structure of an individual virus string can be called a genotypic change--a mutation--if it is passed on in replication, and otherwise a mere phenotypic change. It is not that such a line can't be drawn, but it becomes a line that prohibits nothing. The claim that Lamarckianism has been vindicated in the world of viral evolution would thus be Pickwickian. And since memes are no more multicellular than they are sexual, the fact that there is no clear way---no "principled" way, as they used to say at MIT--of distinguishing mutations from phenotypic acquisitions hardly shows that they are disqualified from a neo-Darwinian treatment.<sup>(6)</sup> Most--much more than 99%-- of the life forms on this planet have evolved under just such a regime, and neo-Darwinism certainly covers their evolution handily.

And the same verdict applies to anastomosis, although this is a recent and ill-appreciated discovery: there is lots of horizontal transmission in protist and bacterial evolution--a fact that plays hob with attempts to define separate bacterial lineages in a "principled" way--and once again, the bulk of the evolution on the planet has been amongst just such tiny bits. Once we shift our focus away from our own multicellular, sexually reproducing lineages to the more numerous lineages on the planet, these standard objections lose much if not all their force. Memes are indeed not very much like elephant genomes, but so what?

### **3. But what about human reason--and creativity? <sup>(7)</sup>**

A confusion that misdirects the imagination of theorists in another direction derives, I suspect, from a subtle misreading of Darwin's original use of artificial selection (deliberate animal breeding) and "unconscious" selection (the unwitting promotion of favored offspring of domesticated animals) as bridges to his concept of natural selection. While it is true that Darwin wished to contrast the utter lack of foresight or intention in natural selection with the deliberate goal-seeking of the artificial selectors, in order to show how the natural process could in principle proceed without any mentality at all, he did not thereby establish (as many seem to have supposed) that deliberate, goal-directed, intentional selection is not a sub variety of natural selection! The short legs of dachshunds, and the huge udders of Holsteins are just as much products of natural selection as the wings of the eagle; they just evolved in an environment that included a particularly well-focused selective pressure consisting of human agents. These phenotypes fall under the same laws of transmission genetics, the same replicator dynamics, as any others--as special and extreme cases in which the default "randomness" or noisiness of selective pressure has been greatly reduced.

Applied to cultural evolution, the implication is this: There is no conflict between the claim that artifacts (including abstract artifacts--memes) are the products of natural selection, and the claim that they are (often) the foreseen, designed products of intentional human activity. It appears that some thinkers in the newly emerging school of *evolutionary archeology* have made this mistake. According to a critique by Boone and Smith (1998), at least some evolutionary archeologists think that the only way to be hardheaded and scientific about the Darwinian evolution of culture is to deny all intention, all rationality, on the part of human culture-makers. They opt for "selection rather than decision-making" [p11]. That is simply a mistake, for the same reason it would be a mistake to say that the fancy plumage of prize pigeons is the result of decision-making rather than selection. But Boone and Smith fall in the same trap, in their discussion of the interesting phenomenon of the spread of snowmobiles among the Cree in northern Canada. They are surely right that the adoption of snowmobiles by the Cree cannot be accounted for in terms of the differential biological replication of the snowmobile users, but they misread the more interesting meme's-eye view perspective. They say:

The alternative that 'snowmobile memes' were transmitted more effectively than 'snowshoe memes' to non-descendant Cree (as well as offspring), while plausible, is not natural selection [emphasis added]; more significantly, it requires precisely the kind of adaptive decision-making that EA [evolutionary archeology] is dedicated to eliminating from archeological explanation. [Ms P12]

On the contrary, if you adopt the meme's-eye perspective, in which the snowmobile meme is seen as the replicator, with its own fitness, then cultural evolution can be seen to be due to "adaptive decision-making" while also a variety of natural selection. Consider the fitness of the domesticated horses that spread so quickly among the Native Americans after their introduction, but then more recently, after the advent of the automobile, have dwindled sharply. These fluctuations in genetic fitness have been due to changes in the selective forces arrayed in the various environments in which the horses have existed, of course, and the fact that conscious,

foresighted human agents form the key component in those selective environments does nothing to remove the phenomena from the domain of standard genetic evolution by natural selection.

Among those who have overlooked this fact is Steven Pinker, who dismisses models of cultural evolution in a brief passage in *How the Mind Works* (1997):

Stop being so literal-minded! respond the fans of cultural evolution. Of course cultural evolution is not an exact replica of the Darwinian version. In cultural evolution, the mutations are directed and the acquired characteristics are inherited. Lamarck, while being wrong about biological evolution, turned out to be right about cultural evolution. To say that cultural evolution is Lamarckian is to confess that one has no idea how it works. The striking features of cultural products, namely their ingenuity, beauty, and truth (analogous to organisms' complex adaptive design), come from the mental computations that "direct"--that is, invent--the "mutations," and that "acquire"--that is, understand--the "characteristics." (1997, p209)

Pinker has imputed the wrong parallel; it is not Lamarck's model, but Darwin's model of artificial selection (as a special case of natural selection) that accommodates the phenomena he draws to our attention in this passage. And it is ironic that Pinker overlooks this, since the cultural phenomena he himself has highlighted as examples of evolution-designed systems, linguistic phenomena, are almost certainly not the products of foresighted, ingenious, deliberate human invention. Some designed features of human languages are no doubt genetically transmitted, but many others--such as changes in pronunciation, for instance--are surely culturally transmitted, and hence products of cultural, not genetic, evolution.

Some memes are like domesticated animals; they are prized for their benefits, and their replication is closely fostered and relatively well understood by their human owners. Some memes are more like rats; they thrive in the human environment in spite of being positively selected against--ineffectually--by their unwilling hosts. And some are more like bacteria or other viruses, commandeering aspects of human behavior (provoking sneezing, for instance) in their "efforts" to propagate from host to host. There is artificial selection of "good" memes--like the memes of arithmetic and writing, which are carefully taught to each new generation. And there is unconscious selection of memes of all sorts--like the subtle mutations in pronunciation that spread through linguistic groups, presumably with some efficiency advantage, but perhaps just hitchhiking on some quirk of human preference. And there is unconscious selection of memes that are positively a menace, but which prey on flaws in the human decision-making apparatus, as provided for in the genome and enhanced and adjusted by other cultural innovations--such as the abducted-by-aliens meme, which makes perfect sense when its own fitness as a cultural replicator is considered. Only the meme's-eye perspective unites all these possibilities under one view.<sup>(8)</sup>

Finally, one of the most persistent sources of discomfort about memes is the dread suspicion that an account of human minds in terms of brains being parasitized by memes will undermine the precious traditions of human creativity. On the contrary, I think it is clear that *only* an account of creativity in terms of memes has much of a chance of giving us any way to *identify with* the products of our own minds. We human beings extrude other products, on a daily basis, but after infancy, at any rate, we don't tend to view our feces with the pride of an author or artist. These

are mere biological byproducts, and although they have their own modest individuality and idiosyncrasy, it is not anything we cherish. How could we justify viewing the secretions of our poor infected brains with any more pride? Because we *identify with* some subset of the memes we harbor. Why? Because among the memes we harbor are those that put a premium on identifying with just such a subset of memes! Lacking that meme-borne attitude, we would be mere *loci* of interaction, but we have such memes--that is who we are.

## 4. Conclusion

This spectrum of possibilities, from the unwitting, unconscious hosting of culture-borne viruses (of all "attitudes") to the foresighted design and promulgation of inventions and creations that intelligently and artfully draw upon well-understood cultural resources, must be viewable under a single, unifying perspective. It is only from such a perspective that we can make sense of the trajectories that have taken us--and only us--beyond the horizons of our selfish genes, by creating new environments of selection--persons and their projects--that in turn create utterly unprecedented answers to the *Cui bono?* question. Such a view of cultural evolution doesn't deny the possibility of moving to what might be called a *mind's-eye* perspective of evaluation; it is precisely what makes such a transition--without any help from skyhooks--possible.

## References

- Boone, James L, and Eric Alden Smith, 1998, "A Critique of Evolutionary Archeology," *Current Anthropology*. [special issue, supplement, June, 1998]
- Boyd, Robert, and Richerson, Peter J., 1985, Culture and the Evolutionary Process, Chicago and London: Chicago University Press.
- Boyd, Robert, and Richerson, Peter J., 1992, "Punishment Allows the Evolution of Cooperation (or Anything Else) in Sizable Groups," *Ethology and Sociobiology*, 13, pp171-95.
- Cavalli-Sforza, Luca, and Feldman, Marcus, 1981, *Cultural Transmission and Evolution: A Quantitative Approach*, Princeton, NJ: Princeton University Press.
- Dawkins, Richard, 1976, *The Selfish Gene*, (2nd edition, 1989), Oxford: Oxford University Press.
- 1993, "Viruses of the Mind," in Bo Dahlbom, ed., *Dennett and his Critics*, Oxford: Blackwell.
- Dennett, Daniel C. 1998, "Snowmobiles, horses, rats, and memes," (a comment on "A Critique of Evolutionary Archeology," by James L. Boone and Eric Alden Smith) *Current Anthropology*, [special issue, supplement, June 1998]
- forthcoming, "The Evolution of Evaluators" in a volume of the Siena workshop on evolutionary economics.

- Diamond, Jared, 1997, *Guns, Germs and Steel*,
- Lumsden, C. and Wilson, E. O. 1981, *Genes, Mind and Culture*, Cambridge, MA: Harvard University Press.
- Pinker, Steven, 1997, *How the Mind Works*, New York: Norton.
- Ridley, Mark, 1995, *Animal Behaviour* (2nd edn), Boston: Blackwell Science.
- Sober, Elliott and Wilson, David Sloan, 1998, *Unto Others: The Evolution of Unselfish Behavior*, Harvard University Press.
- Wilson, E. O., 1978, *On Human Nature*, Cambridge, MA: Harvard University Press.
1. Parts of this section are adapted from Dennett, forthcoming.
  2. Such organisms need not be deemed to be making conscious decisions, of course, but the rationality, such as it is, of the "decisions" they make is typically anchored to the expected benefit to the individual organism. See Sober and Wilson (1998) for important discussions of gene, individual, and group benefits of such decision-making.
  3. Sober and Wilson (1998) note that there is a gap in their model of cultural evolution: "We can say that functionless [relative to human individual and group fitness] behavior should be more common in humans than other species, but we cannot explain why a particular functionless behavior has evolved in a particular culture. That kind of understanding probably requires detailed historical knowledge of the culture, and it may turn out that some behaviors evolved mainly by chance." p171.
  4. Strictly speaking, to the reproductive prospects of the fluke's genes (or the fluke's "group"'s genes), for as Sober and Wilson (1998) point out (p18) in their use of *D. dendriticum* as an example of altruistic behavior, the fluke that actually does the driving in the brain is a sort of kamikaze pilot, who dies without any chance of passing on its own genes, benefiting its [asexually reproduced] near-clones in other parts of the ant.
  5. Boyd and Richerson (1992) show that "Virtually *any* behavior can become stable within a social group if it is sufficiently buttressed by social norms." (Sober and Wilson, 1998, p.152)
  6. In fact in most of the brief, shallow discussions of the Lamarckian nature of cultural evolution that I have seen, it is never made clear which entities were deemed capable of transmitting acquired characteristics. Sometimes, I suspect, the objector had dimly in mind the strictly irrelevant fact that human *hosts* can transmit to other human hosts cultural items that they themselves had acquired during their lifetimes. That is not Lamarckianism at all.
  7. This section is adapted from Dennett, 1998.

8. The meme's-eye perspective offers many other points of theoretical leverage, but those are topics for another occasion.

# THE EVOLUTION OF CULTURE

By  
**Daniel C. Dennett**

The Charles Simonyi Lecture  
Oxford University, Feb 17, 1999

*The philosopher Daniel C. Dennett is interested in consciousness, and his view of it, similar to that of Minsky's, is as high-level, abstract thinking. He is known as the leading proponent of the computational model of the mind; he has clashed with philosophers such as John Searle who maintain that the most important aspects of consciousness — intentionality and subjective quality — can never be computed. He is the philosopher of choice of the AI community. In his more recent work, he has turned to what he calls "Darwin's dangerous idea"; he is squarely in the ultra-Darwinist camp of George C. Williams and Richard Dawkins, and he has with great energy mustered a serious critique of the scientific ideas of Stephen Jay Gould. "Dan Dennett is our best current philosopher," says Marvin Minsky. "He is the next Bertrand Russell. Unlike traditional philosophers, Dan is a student of neuroscience, linguistics, artificial intelligence, computer science, and psychology. He's redefining and reforming the role of the philosopher." — John Brockman*

Cultures evolve. In one sense, this is a truism; in other senses, it asserts one or another controversial, speculative, unconfirmed theory of culture. Consider a cultural inventory of some culture at some time--say 1900AD. It should include all the languages, practices, ceremonies, edifices, methods, tools, myths, music, art, and so forth, that compose that culture. Over time, that inventory changes. Today, a hundred years later, some items will have disappeared, some multiplied, some merged, some changed, and many new elements will appear for the first time. A verbatim record of this changing inventory through history would not be science; it would be a data base. That is the truism: cultures evolve over time. Everybody agrees about that. Now let's turn to the controversial question: how are we to explain the patterns to be found in that data base? Are there any good theories or models of cultural evolution?

## 1. Science or Narrative?

One possibility is that the only patterns to be found in cultural evolution defy scientific explanation. They are, some might want to say, *narrative* patterns, not scientific patterns. There is clearly something to this, but it won't do as it stands, for many scientific patterns are also historical patterns, and hence are revealed and explained in narratives--of sorts. Cosmology, geology, and biology are all historical sciences. The great biologist D'Arcy Thompson once said:

Everything is the way it is because it got that way.

If he is right--if *everything* is the way it is because it got that way--then every science must be, in part, a historical science. But not all history--all recounting of events in temporal sequence--is *narrative*, some might want to say. Human history is unique in that the patterns it exhibits require a different *form* of understanding: *hermeneutical* understanding or *Verstehen*, or--you can

count on the Germans to have lots of words for claims like this--*Geisteswissenschaft* (approximately: spiritual science). I think this too is partly right; there *is* a particular sort of understanding that is used to make sense of narratives about human agents. It is also true that the mark of a *good* story is that its episodes unfold *not* as the predicted consequences of general laws and initial conditions, but in delightfully surprising ways. These important facts do not show, however, that cultural evolution escapes the clutches of science and must be addressed in some other realm of inquiry. Quite the contrary; the humanistic comprehension of narratives and the scientific explanation of life processes, for all their differences of style and emphasis, have the same logical backbone. We can see this by examining the special form of understanding we use when following--and creating--good narratives.

Mediocre narratives are either a pointless series of episodes in temporal order--just "one damn thing after another"--or else so utterly predictable as to be boring. Between randomness and routine lie the good stories, whose surprising moments make sense in retrospect, in the framework provided by the unsurprising moments. The perspective from which we can understand these narratives is what I have called *the intentional stance*: the strategy of analyzing the flux of events into *agents* and their (rational) *actions* and *reactions*. Such agents--people, in this case--*do things for reasons*, and can be predicted--up to a point--by cataloguing their reasons, their beliefs and desires, and calculating what, given those reasons, the most rational course of action for each agent would be. Sometimes the most rational course is flat obvious, so while the narrative is predictive (or true), it is uninteresting and unenlightening. To take a usefully simple case, a particular game of chess is interesting to the extent that we are surprised by either the brilliant moves that outstrip our own calculations of what it would be rational to do, or the blunders, which we thought too sub-optimal to predict.

In the wider world of human activity, the same holds true. We don't find the tale of Jane going to the supermarket on her way home from work interesting precisely because it all unfolds so predictably from the intentional stance; today she never encountered any interesting options, given her circumstances. Other times, however, the most rational thing for an agent to do is far from obvious, and maybe practically incalculable. When we encounter these narratives, we are surprised (and sometimes delighted, sometimes appalled) by the actual outcome. It makes sense in retrospect, but who'd have guessed that she'd decide to do *that*? The vast mass of routinely rational human behavior doesn't make good novels, but it is just such humdrum rational narrative that provides the background pattern that permits us to make sense, retrospectively, of the intriguing vagaries we encounter, and to anticipate the complications that will arise when the trains of events they put in motion collide.

The traditional model used by historians and anthropologists to try to explain cultural evolution uses the intentional stance as its explanatory framework. These theorists treat culture as composed of goods, possessions of the people, who husband them in various ways, wisely or foolishly. People carefully preserve their traditions of fire-lighting, house-building, speaking, counting, justice, etc. They trade cultural items as they trade other goods. And of course some cultural items (wagons, pasta, recipes for chocolate cake, etc.) are definitely goods, and so we can plot their trajectories using the tools of economics. It is clear from this perspective that highly prized cultural entities will be protected at the expense of less favored cultural entities, and there will be a competitive market where agents both "buy" and "sell" cultural wares. If a

new method of house-building or farming or a new style of music sweeps through the culture, it will be because people perceive advantages to these novelties.

The people, on this model, are seen as having an autonomous rationality: deprive a person of his goods, and he stands there, naked but rational and full of informed desires. When he clothes himself and arms himself and equips himself with goods, he increases his powers, complicates his desires. If Coca Cola bottles proliferate around the world, it is because more and more people prefer to buy a Coke. Advertising may fool them. But then we look to the advertisers, or those who have hired them, to find the relevant agents whose desires fix the values for our cost-benefit calculations. *Cui bono?* Who benefits? The purveyors of the goods, and those they hire to help them. etc. On this way of thinking, then, the relative "replicative" power of various cultural goods--whether Coke bottles, building styles or religious creeds--is measured in the marketplace of cost-benefit calculations performed by the people.

Biologists , too, can often make sense of the evolution (in the neutral sense) of features of the natural world by treating them as goods belonging to various members of various species: one's food, one's nest, one's burrow, one's territory, one's mate[s], one's time and energy. Cost-benefit analyses shed light on the husbandry engaged in by the members of the different species inhabiting some shared environment.[1] Not every "possession" is considered a good, however. The dirt and grime that accumulates on one's body, to say nothing of the accompanying flies and fleas, are of no value, or of negative value, for instance. These hitchhikers are not normally considered as goods by biologists, except when the benefits derived from them (by *whom?*) are manifest.

This traditional perspective can obviously explain many features of cultural and biological evolution, but it is not uniformly illuminating, nor is it obligatory. I want to show how theorists of culture--historians, anthropologists, economists, psychologists, and others--can benefit from adopting a different vantage point on these phenomena. It is a different application of the intentional stance, one which still quite properly gives pride of place to the *Cui bono* question, but which can provide alternative answers that are often overlooked. The perspective I am talking about is Richard Dawkins' *meme's-eye point of view*, which recognizes--and takes seriously--the possibility that cultural entities may evolve according to selectional regimes that make sense only when the answer to the *Cui bono* question is that it is the cultural items *themselves* that benefit from the adaptations they exhibit. [2]

## 2. Memes as Cultural Viruses

Whenever costs and benefits are the issue we need to ask *Cui bono?* A benefit by itself is not explanatory; a benefit in a vacuum is indeed a sort of mystery; until it can be shown how the benefit actually redounds to enhance the replicative power of a replicator, it just sits there, alluring, perhaps, but incapable of explaining anything.

We see an ant laboriously climbing up a stalk of grass. Why is it doing that? Why is that adaptive? What good accrues to the ant by doing that? That is the wrong question to ask. No good at all accrues to the ant. Is it just a fluke, then? In fact, that's exactly what it is: a fluke! Its brain has been invaded by a fluke (*Dicrocoelium dendriticum*), one of a gang of tiny parasites

that need to get themselves into the intestines of a sheep in order to reproduce (Ridley, 1995, p258). (Salmon swim up stream, these parasitic worms drive ants up grass stalks, to improve their chances of being ingested by a passing sheep.) The benefit is not to the reproductive prospects of the ant but the reproductive prospects of the fluke. [3]

Dawkins points out that we can think of cultural items, memes, as parasites, too. Actually, they are more like a simple virus than a worm. Memes are supposed to be analogous to genes, the replicating entities of the cultural media, but they also have vehicles, or phenotypes; they are like not-so-naked genes. They are like viruses (Dawkins, 1993). Basically, a virus is just a string of nucleic acid with attitude--and a protein overcoat. A viroid is an even more naked gene. And similarly, a meme is an information-packet with attitude--with some phenotypic clothing that has differential effects in the world that thereby influences its chances of getting replicated. (What is a meme *made of*? It is made of information, which can be carried in *any* physical medium. More on this later.)

And in the domain of memes, the ultimate beneficiary, the beneficiary in terms of which the final cost-benefit calculations must apply is: the meme itself, not its carriers. This is not to be heard as a bold empirical claim, ruling out (for instance) the role of individual human agents in devising, appreciating and securing the spread and prolongation of cultural items. As I have already noted, the traditional perspective on cultural evolution handsomely explains many of the patterns to be observed. My proposal is rather that we adopt a perspective or point of view from which a wide variety of different empirical claims can be compared, *including the traditional claims*, and the evidence for them considered in a neutral setting, a setting that does not prejudge these hot-button questions.

In the analogy with the fluke, we are invited to consider a meme to be like a parasite which commandeers an organism for its own replicative benefit, but we should remember that such hitchhikers or *symbionts* can be classified into three fundamental categories:

- *Parasites*, whose presence lowers the fitness of their host;
- *Commensals*, whose presence is neutral (though, as the etymology reminds us, they "share the same table"); and
- *Mutualists*, whose presence enhances the fitness of both host and guest.

Since these varieties are arrayed along a continuum, the boundaries between them need not be too finely drawn; just where benefit drops to zero or turns to harm is not something to be directly measured by any practical test, though we can explore the consequences of these turning points in models.

We should expect memes to come in all three varieties, too. This means, for instance, that it is a mistake to assume that the "cultural selection" of a cultural trait is always "for cause"--always because of some perceived (or even misperceived) benefit it provides to the host. We can always ask if the hosts, the human agents that are the *vectors*, perceive some benefit and (for that reason, good or bad) assist in the preservation and replication of the cultural item in question, but we must be prepared to entertain the answer that they do not. In other words, we must consider as a real possibility the hypothesis that the human hosts are, individually or as a group, either

oblivious to, or agnostic about, or even positively dead set against, some cultural item, which nevertheless is able to exploit its hosts as vectors.

The most familiar cases of cultural transmission and evolution--the cases that tend to be in the spotlight--are innovations that are obviously of some direct or indirect benefit to the genetic fitness of the host. A better fishhook catches more fish, feeds more bellies, makes for more surviving grandchildren, etc. The only difference between stronger arms and a better fishhook in the (imagined) calculation of impact on fitness is that the stronger arms might be passed on quite directly through the germ line, while the fishhook definitely must be culturally transmitted. (The stronger arms *could* be culturally transmitted as well. A tradition of body-building, for instance, could explain why there was very low [genetic] heritability for strong adult arms, and yet a very high rate of strong adult arms in a population.) But however it might be that strong arms or fishhooks are transmitted, they are typically supposed to be a good bargain from the perspective of genetic fitness. The bargain might, however, be myopic--only good in the short run. After all, even agriculture, in the long run, may be a dubious bargain if what you are taking as your *summum bonum* is Darwinian fitness (*see Diamond, 1997, for fascinating reflections on the uncertain benefits of abandoning the hunter-gatherer lifestyle*). What alternatives are there?

First, we need to note that in the short run (evolutionarily speaking--that is, from the perspective of a few centuries or even millennia) something might flourish in a culture independently of whether it was of actual benefit to genetic fitness, but strongly linked to whether it was of *apparent* benefit to genetic fitness. Even if you think that Darwinian fitness enhancement is the principle driving engine of cultural evolution, you have to posit some swifter, more immediate mechanism of retention and transmission. It's not hard to find one. We are genetically endowed with a biased quality space: some things feel good and some things don't. We tend to live by the rule: *if it feels good, keep it*. This rough and ready rule can be tricked, of course. The sweet tooth is a standard example. The explosion of cultural items--artifacts, practices, recipes, patterns of agriculture, trade routes--that depend quite directly on the exploitation of the sweet tooth has probably had a considerable net *negative* effect on human genetic fitness. Notice that explaining the emergence of these cultural items by citing their "apparent" benefit to genetic fitness does not in any way commit us to the claim that people think that they are enhancing their genetic fitness by acquiring and consuming sugar. The rationale is not theirs, but Mother Nature's. They just go with what they like.

Still, given what people innately like, they go on to figure out, ingeniously and often with impressive foresight, how to obtain what they like. This is still the traditional model of cultural evolution, with people husbanding their goods in order to maximize what they prefer--and getting their preferences quite directly from their genetic heritage. But this very process of rational calculation can lead to more interesting possibilities. As such an agent complicates her life, she will almost certainly acquire new preferences that are themselves culturally transmitted symbionts of one sort or another. Her sweet tooth may lead her to buy a cookbook, which inspires her to enroll in a culinary arts program, which turns out to be so poorly organized that she starts a student protest movement, in which she is so successful that she is invited to head an educational reform movement, for which a law degree would be a useful credential, and so on. Each new goal will have to bootstrap itself into the memosphere by exploiting some pre-established preference, but this recursive process, which can proceed at breakneck speed relative

to the glacial pace of genetic evolution, can transform human agents indefinitely far away from their genetic beginnings. In an oft-quoted passage, E. O. Wilson claimed otherwise: The genes hold culture on a leash. The leash is very long, but inevitably values will be constrained in accordance with their effects on the human gene pool. (Wilson, 1978, p167)

But Wilson's leash is indefinitely long and elastic. Consider the huge space of *imaginable* cultural entities, practices, values. Is there any point in that Vast space that is utterly unreachable? Not that I can see. The constraints Wilson speaks of can be so co-opted, exploited, and blunted in a recursive cascade of cultural products and meta-products that there may well be traversable paths to every point in that space of imaginable possibilities. I am suggesting, that is, that cultural possibility is less constrained than genetic possibility. We can articulate persuasive biological arguments to the effect that certain imaginable species are unlikely in the extreme--flying horses, unicorns, talking trees, carnivorous cows, spiders the size of whales--but neither Wilson nor anybody else to my knowledge has yet offered parallel grounds for believing that there are similar obstacles to trajectories in imaginable cultural design space. Many of these imaginable points in design space would no doubt be genetic cul-de-sacs, in the sense that any lineage of *H. sapiens* that ever occupied them would eventually go extinct as a result, but this dire prospect is no barrier to the evolution and adoption of such memes in the swift time of cultural history. [4] To combat Wilson's metaphor with one of my own: the genes provide not a leash but a launching pad, from which you can get almost anywhere, by one devious route or another. It is precisely in order to explain the patterns in cultural evolution that are *not* strongly constrained by genetic forces that we need the memetic approach.

The memes that proliferate will be the memes that replicate one way or another--by hook or by crook. Think of them as entering the brains of culture members, making phenotypic alterations thereupon, and then submitting themselves to the great selection tournament--not the Darwinian genetic fitness tournament (life is too short for that) but the Dawkinsian meme-fitness tournament. It is their fitness as memes that is on the line, not their host's genetic fitness. And the environments that embody the selective pressures that determine their fitness are composed in large measure of other memes.

Why do their hosts put up with this? Why should the overhead costs of establishing a whole new system of differential reproduction be borne by members of *H. sapiens*? Note that the question to be asked and answered here is parallel to the question we ask about any symbiont-host relationship: why do the hosts put up with it? And the short answer is that it is too costly to eradicate, but this just means that the benefits accruing to the machinery that is being exploited by the parasites are so great that keeping the machinery and tolerating the parasites (to the extent that they are tolerated) has so far been the best deal available. And whether or not in the long run (millions of years) this infestation will be viewed as mutualism or commensalism or parasitism, in the short run (the last few millennia) the results have been spectacular: the creation of a new biological type of entity: a person.

I like to compare this development to the revolution that happened among the bacteria roughly a billion years ago. Relatively simple *prokaryotes* got invaded by some of their neighbors, and the resulting *endosymbiotic* teams were more fit than their uninfected cousins, and prospered. These *eukaryotes*, living alongside their prokaryotic cousins, but enormously more complex, versatile

and competent thanks to their hitchhikers, opened up the design space of multi-cellular organisms. Similarly, the emergence of culture-infected hominids has opened up yet another region of hitherto unoccupied and untraversable design space. We live alongside our animal cousins, but we are enormously more complex, versatile and competent. Our brains are bigger, to be sure, but it is mainly due to their infestation by memes that they gain their powers. Joining forces with our own memes, we create new candidates for the locus of benefit, new answers to *Cui bono?*

### 3. Darwin's Path to Memetic Engineering

The meme's-eye view doesn't just open up new vistas for the understanding of patterns in culture; it also provides the foundation for answering a question left dangling by the traditional model of cultural evolution. The traditional view presupposes rational self-interested agents, intent on buying and selling, and improving their lot. *Where did they come from?* The standard background assumption is that they are just animals, whose *Cui bono?* question is to be dealt with in terms of the impact on genetic fitness, as we have seen. But when people acquire other interests, including interests directly opposed to their genetic interests, they enter a new space of possibilities--something no salmon or fruitfly or bear can do. How could this great river of novelty get started?

Here I think we can get help from Darwin's opening exposition of the theory of natural selection. In the first chapter of *Origin of Species*, Darwin introduces his great idea of natural selection by an ingenious expository device, an instance of the very gradualism that he was about to discuss. He begins not with natural selection--his destination--but what he calls *methodical selection*: the deliberate, foresighted, intended "improvement of the breed" by animal and plant breeders. He begins, in short, with familiar and uncontroversial ground that he can expect his readers to share with him.

We cannot suppose that all the breeds were suddenly produced as perfect and as useful as we now see them; indeed, in several cases, we know that this has not been their history. The key is man's power of accumulative selection: nature gives successive variations; man adds them up in certain directions useful to him. [p30, Harvard facsimile edn]

But, he goes on to note, in addition to such methodical selection, there is another process, which lacks the foresight and intention, which he calls *unconscious selection*:

At the present time, eminent breeders try by methodical selection, with a distinct object in view, to make a new strain or sub-breed, superior to anything existing in the country. But, for our purpose, a kind of Selection, which may be called Unconscious, and which results from every one trying to possess and breed from the best individual animals, is more important. Thus, a man who intends keeping pointers naturally tries to get as good dogs as he can, and afterwards breeds from his own best dogs, but he has no wish or expectation of permanently altering the breed. [p34].

Long before there was deliberate breeding, unconscious selection was the process that created and refined all our domesticated species, and even at the present time, unconscious selection

continues. Darwin gives a famous example:

*There is reason to believe that King Charles's spaniel has been unconsciously modified to a large extent since the time of that monarch. [p35]*

There is no doubt that unconscious selection has been a major force in the evolution of domesticated species. On unconscious selection of both domesticated plants and animals, see Jared Diamond, *Guns, Germs and Steel*. In our own time, unconscious selection goes on apace, and one ignores it at our peril. Unconscious selection in bacteria and viruses for resistance to antibiotics is only the most notorious and important example. Consider the "genes for longevity" that have recently been bred into laboratory animals such as mice and rats. It is probably true, however, that much if not all of the effect that has been obtained in these laboratory breeding experiments has simply undone the unconscious selection for short-livedness at the hands of the suppliers of those laboratory animals. The stock the experimenters started with had shorter life expectancy than their wild cousins simply because they had been bred for many generations for early reproductive maturity, and robustness, and short lives came along as an unintended (unconscious) side consequence (Daniel Promislow, personal correspondence).

Darwin pointed out that the line between unconscious and methodical selection was itself a fuzzy, gradual boundary:

*The man who first selected a pigeon with a slightly larger tail, never dreamed what the descendants of that pigeon would become through long-continued, partly unconscious and partly methodical selection. [p 39]*

And both unconscious and methodical selection, he notes finally, are but special cases of an even more inclusive process, natural selection, in which the role of human intelligence and choice stands at zero. From the perspective of natural selection, changes in lineages due to unconscious or methodical selection are merely changes in which one of the most prominent selective pressures in the environment is human activity. It is not restricted, as we have seen, to domesticated species. White-tailed deer in New England now seldom exhibit the "white flag" of a bobbing tail during headlong flight that was famously observed by early hunters; the arrival of human beings today is much more likely to provoke them to hide silently in underbrush than to flee. Those white flags were too easy a target for hunters with guns, it seems.

This nesting of different processes of natural selection now has a new member: genetic engineering. How does it differ from the methodical selection of Darwin's day? It is just less dependent on the pre-existing variation in the gene pool, and proceeds more directly to new candidate genomes, with less overt and time-consuming trial and error. Darwin had noted that in his day,

Man can hardly select, or only with much difficulty, any deviation of structure excepting such as is externally visible; and indeed he rarely cares for what is internal. [p38].

But today's genetic engineers have carried their insight into the molecular innards of the organisms they are trying to create. There is ever more accurate foresight, but even here, if we

look closely at the practices in the laboratory, we will find a large measure of exploratory trial and error in their search of the best combinations of genes.

We can use Darwin's three levels of genetic selection, plus our own fourth level, genetic engineering, as a model for four parallel levels of *memetic* selection in human culture. In a speculative spirit, I am going to sketch how it might go, using an example that has particularly challenged some Darwinians, and hence been held up as a worthy stumbling block: a cultural treasure untouchable by evolutionists: music. Music is unique to our species, but found in every human culture. It is manifestly complex, intricately designed, an expensive consumer of time, energy and materials. How did music start? What was or is the answer to its *Cui bono* question? Steven Pinker (1997) is one Darwinian who has recently declared himself baffled about the possible evolutionary origins and survival of music, but that is because he has been looking at music in the old-fashioned way, looking for music to have some contribution to make to the genetic fitness of those who make and participate in the proliferation of music. [5] There may well be some such effect that is important, but I want to make the case that there might also be a purely *memetic* explanation of the origin of music. Here, then, is my Just-so Story, working gradually up Darwin's hierarchy of kinds of selection.

## Natural selection of musical memes

One day one of our distant hominid ancestors sitting on a fallen log happened to start banging on with a stick--*boom boom boom*. For *no good reason at all*. This was just idle diddling, a byproduct, perhaps, of a slightly out-of-balance endocrine system. This was, you might say, mere nervous fidgeting, but the repetitive sounds striking his ears just happened to feel to him like a slight improvement on silence. A feedback loop was closed, and the *repetition--boom boom boom--was "rewarding"*. If we leave this individual all by himself, drumming away on his log, then we would say that he had simply developed a habit, *possibly* therapeutic in that it "relieved anxiety," but just as possibly a *bad* habit--a habit that did him and his genes no good at all, but just exploited a wrinkle that happened to exist in his nervous system, creating a feedback loop that tended to lead to individual replications of drumming by him under various circumstances. No musical appreciation, no insight, no goal or ideal or project need be imputed to our solitary drummer.

Now introduce some other ancestors who happen to see and hear this drummer. They might pay no attention, or be irritated enough to make him stop or drive him away, or they might, again *for no reason*, find their imitator-circuits tickled into action; they might feel an urge to drum along with musical Adam. What are these imitator circuits I've postulated? Just whatever it takes to make it somewhat more likely than not that some activities by conspecifics are imitated, a mere reflex if you like--of which we may see a fossil trace when spectators at a football match cannot help making shadow kicking motions more or less in unison with the players on the field. One can postulate reasons why having some such imitative talents built-in would be a valuable adaptation--one that enhances one's genetic fitness--but while this is both plausible and widely accepted, it is strictly speaking unnecessary for my Just-so Story. The imitative urge might just as well be a functionless byproduct of some other adaptive feature of the human nervous system. Suppose, then, that for no good reason at all, the drumming habit is *infectious*. When one hominid starts drumming, soon others start drumming along in imitation. This could happen. A

perfectly pointless practice, of no utility or fitness-enhancing benefit at all, could become established in a community. It might be positively detrimental: the drumming scares away the food, or uses up lots of precious energy. It would then be just like a disease, spreading simply because it *could* spread, and lasting as long as it could find hosts to infect. If it was detrimental in this way, variant habits that were less detrimental--less virulent--would tend to evolve to replace it, other things being equal, for they would tend to find more available healthy hosts to migrate to. And of course such a habit *might* even provide a positive benefit to its hosts (enhancing their reproductive chances--a familiar dream of musicians everywhere, and it might be true, or have been true in the past). But providing a genetic benefit of this sort is only one of the paths such a habit might pursue in its mindless quest for immortality. Habits--good, bad and indifferent--could persist and replicate, unappreciated and unrecognized, for an indefinite period of time, provided only that the replicative and dispersal machinery is provided for them. The drumming virus is born.

Let me pause to ask the question: what is such a habit made of? What gets passed from individual to individual when a habit is copied? Not stuff, not packets of material, but pure information, the information that generates the pattern of behavior that replicates. A cultural virus, unlike a biological virus, is not tethered to any particular physical medium of transmission. [6]

## Unconscious selection of memes

On with our Just-so Story. Some of the drummers begin to hum, and of all the different hums, a few are more infectious than the rest, and those hominids who happen to start the humming in these ways become the focus of attention, as sources of humming. A competition between different humming patterns emerges. Here we can begin to see the gradual transition to unconscious selection. Suppose that being such a focus of humming happens to feel good--whether or not it enhances one's genetic fitness slightly (it might, of course; perhaps the females tend to be more receptive to those who start the winning hums). The same transition to unconscious selection can be seen among viruses and other pathogens, by the way. If scratching an itch feels good, and also has the side effect of keeping a ready supply of viral emigrés on one's fingertips, the part of the body most likely to come in contact with another host, one is unconsciously selecting for just such a mode of transmission by one's myopic and uncomprehending preference for scratching when one itches--and this does not depend on scratching having any fitness-enhancing benefits *for you*: it may be, like the ant's hankering for the top of the grass stem, a desire that benefits the parasite, not the host. Similarly, if varying tempo and pitch of one's hums feels good, and also happens to create a ready supply of more attention-holding noises for spreading to conspecifics, one's primitive aesthetic preference can begin to shape, unconsciously, the lineages of humming habit that spread through one's community.

Brains in the community begin to be infected by a variety of these memes. Competition for time and space in these brains becomes more severe. The infected brains begin to take on a structure, as the memes that enter "learn" to cooperate on the task of turning a brain into a proper memenest, with lots of opportunities for entrance and exit (and hence replication). [7] Meanwhile, any

memes out there "looking for" hosts, will have to compete for available space therein. Just like germs.

## Methodical Selection of memes

As the structure grows, it begins to take on a more active role in selecting. That is to say, the brains of the hosts, like the brains of the owners of domesticated animals, become ever more potent and discerning selective agencies--still largely unwitting, but nevertheless having a powerful influence. Some people, it turns out, are better at this than others. As Darwin says of animal breeders,

Not one man in a thousand has accuracy of eye and judgment sufficient to become an eminent breeder. P32.

We honor Bach, the artistic genius, but he was no "natural" doodler, an intuitive genius just "playing by ear". He was the master musical technologist of his day, the inheritor of musical instruments that had had their designs honed over several millennia, as well as some relatively recent additions to the music-maker's toolbox--a fine system of musical notation, keyboard instruments that permitted the musician to play many notes at once, and an explicit, codified, rationalized *theory* of counterpoint. These mind-tools were revolutionary in the way they opened up musical design space for Bach and his successors.

And Bach, like the one man in a thousand who has the discernment to be an eminent animal breeder, knew how to breed new strains of music from old. Consider, for instance, his hugely successful chorale cantatas. Bach shrewdly chose, for his breeding stock, *chorales*--hymn melodies that had already proven themselves to be robust inhabitants of their human hosts, *already domesticated* tunes his audiences had been humming for generations, building up associations and memories, memes that had already sunk their hooks deeply into the emotional habits and triggers of the brains where they had been replicating for years. Then he used his technology to create variations on these memes, seeking to strengthen their strengths and damp their weaknesses, putting them in new environments, inducing new hybrids.

## Memetic Engineering

What about memetic engineering? Was Bach, in virtue of his highly sophisticated approach to the design of replicable musical memes, not just a *meme-breeder* but a *memetic engineer*? In the light of Darwin's admiring comment on the rare skill--the genius--of the good breeder, it is interesting to note how sharply our prevailing attitudes distinguish between our honoring the "art" of selective breeding and our deep suspicion and disapproval of the "technology" of gene-splicing. Let's hear it for *art*, but not for *technology*, we say, forgetting that the words share a common ancestor, *techné*, the Greek word for art, skill, or craft in any work. We retreat in horror from genetically engineered tomatoes, and turn up our noses at "artificial" fibers in our clothing, while extolling such "organic" and "natural" products as whole grain flour or cotton and wool, forgetting that grains and cotton plants and sheep are themselves products of human technology, of skillful hybridization and rearing techniques. He who would clothe himself in fibers

unimproved by technology and live on food from non-domesticated sources is going to be cold and hungry indeed.

Besides, just as genetic engineers, for all their foresight and insight into the innards of things, are still at the mercy of natural selection when it comes to the fate of their creations (that is why, after all, we are so cautious about letting them release their brainchildren on the outside world), so too the memetic engineer, no matter how sophisticated, still has to contend with the daunting task of winning the replication tournaments in the memosphere. One of the most sophisticated musical memetic engineers of the age, Leonard Bernstein, wryly noted this in a wonderful piece he published in 1955 entitled:

"Why don't you run upstairs and write a nice Gershwin tune?"  
(*New York Times*, April 1955 Reprinted in *The Joy of Music*, 1959, pp52-62)

Bernstein had credentials and academic honors aplenty in 1955, but no songs on the Hit Parade.

*A few weeks ago a serious composer-friend and I . . . got boiling mad about it. Why shouldn't we be able to come up with a hit, we said, if the standard is as low as it seems to be? We decided that all we had to do was to put ourselves into the mental state of an idiot and write a ridiculous hillbilly tune.*

*They failed--and not for lack of trying. As Bernstein wistfully remarked, "It's just that it would be nice to hear someone accidentally whistle something of mine, somewhere, just once." [p54]*

His wish came true, of course, a few years later in 1961, when *West Side Story* burst into the memosphere.

#### 4. Conclusions

There is surely much, much more to be said--to be discovered--about the evolution of music. I chose it as my topic because it so nicely illustrates the way the traditional perspective on culture and the evolutionary perspective can join forces, instead of being seen to be in irresolvable conflict. If you believe that music is *sui generis*, a wonderful, idiosyncratic feature of our species that we prize in spite of the fact that it has *not* been created to enhance our chances of having more offspring, you may well be right--*and if so, there is an evolutionary explanation of how this can be true*. You cannot evade the obligation to explain how such an expensive, time-consuming activity came to flourish in this cruel world, and a Darwinian theory of culture is an ally, not an opponent, in this investigation.

While it is true that Darwin wished to contrast the utter lack of foresight or intention in natural selection with the deliberate goal-seeking of the artificial or methodical selectors, in order to show how the natural process could in principle proceed without any mentality at all, he did not thereby establish (as many seem to have supposed) that deliberate, goal-directed, intentional selection is not a subvariety of natural selection! There is no conflict between the claim that

artifacts (including abstract artifacts--memes) are the products of natural selection, and the claim that they are (often) the foreseen, designed products of intentional human activity.

Some memes are like domesticated animals; they are prized for their benefits, and their replication is closely fostered and relatively well understood by their human owners. Some memes are more like rats; they thrive in the human environment in spite of being positively selected against--ineffectually--by their unwilling hosts. And some are more like bacteria or viruses, commandeering aspects of human behavior (provoking sneezing, for instance) in their "efforts" to propagate from host to host. There is artificial selection of "good" memes--like the memes of arithmetic and writing, the theory of counterpoint, and Bach's cantatas, which are carefully taught to each new generation. And there is unconscious selection of memes of all sorts--like the subtle mutations in pronunciation that spread through linguistic groups, presumably with some efficiency advantage, but perhaps just hitchhiking on some quirk of human preference. And there is unconscious selection of memes that are positively a menace, but which prey on flaws in the human decision-making apparatus, as provided for in the genome and enhanced and adjusted by other cultural innovations--such as the abducted-by-aliens meme, which makes perfect sense when its own fitness as a cultural replicator is considered. Only the meme's-eye perspective unites all these possibilities under one view.

Finally, one of the most persistent sources of discomfort about memes is the dread suspicion that an account of human minds in terms of brains being parasitized by memes will undermine the precious traditions of human creativity. On the contrary, I think it is clear that *only* an account of creativity in terms of memes has much of a chance of giving us any way to *identify with* the products of our own minds. We human beings extrude other products, on a daily basis, but after childhood, we don't tend to view our feces with the pride of an author or artist. These are mere biological byproducts, and although they have their own modest individuality and idiosyncracy, it is not anything we cherish. How could we justify viewing the secretions of our poor infected brains with any more pride? Because we *identify with* some subset of the memes we harbor. Why? Because among the memes we harbor are those that put a premium on identifying with just such a subset of memes! Lacking that meme-borne attitude, we would be mere *loci* of interaction, but we have such memes--that is who we are.

# COMPETING MEMES ANALYSES

By

David K. Dirlam

414 Whitaker Street

Savannah, GA 31401 USA.

dirlamd@bellsouth.net

## Abstract

Aunger (2000) and Edmonds (2002) argue that memetics is a theory without a methodology, in imminent danger of dying from lack of novel interpretations and empirical work. Edmonds challenges memeticists to conduct empirical tests. This article presents Competing Memes Analysis, an empirical methodology that can readily be applied to significant social problems. The methodology is implemented in three steps. Step 1 identifies the organization of memes within an activity. Each activity is assumed to exhibit numerous small groups of memes where each meme within a group competes with all other memes in the group and can be combined with any meme from any other group. The succession of memes that occurs with increasing experience can be a powerful clue to identifying competing memes. Step 2 collects records of activities and codes them for the presence or absence of each meme identified in Step 1. Any activity that people acquire from each other by imitation can be readily coded for the presence or absence of competing memes. Step 3 analyzes changing frequencies of each coded meme over time or space. Models of these changes can give useful clues to suggest empirical studies that will provide important social and scientific results. Ecology's Lotka-Volterra model of competing species illustrates the usefulness to memetics of population models.

**Keywords:** memetic methodology, meme, drawing, writing, scientific research, Lotka-Volterra, competition

## 1. The challenge to memetics

Aunger (2000) contains several arguments that memetics is a theory without a methodology. He concludes (p. 230) "The ultimate test—which would preempt theoretical objections—is whether memetics can produce novel empirical work or insightful interpretations of previous results. It has not yet done so, but must do so in the near future. Otherwise, it is likely that memetics will be perceived to be a misguided enterprise. The clock is ticking."

Edmonds (2002) extends Auger's argument by specifying three challenges for memetics. His first challenge is for a conclusive case study which shows a cultural process that (a) exhibits faithful replication where the replication process is transferred to many sources and (b) passes through a large enough quantity of such transfers to show adaptation resulting from verifiable advantages and consistent with current population-genetics models. He suggests nursery rhymes and legal phrases as appropriate test cases. Edmonds' second challenge is for a falsifiable theory that identifies when a memetics interpretation is appropriate and necessary. Edmonds' third

challenge is for a credible simulation of an emergent (not "designed in") memetic process, analogous to "exhibiting a simulation of the emergence of life from the interaction of chemicals." This paper describes methodological techniques for meeting Edmonds' case-study and memetics-interpretation challenges. Before describing the methodology it would be helpful to clarify a few aspects of the memetics challenge.

### **1.1 General and specific memetics**

A significant source of potential confusion about memetics resembles the confusion within scientific history that occurred with Kuhn's (1969) initial conception of "paradigm." A paradigm referred both to (a) "the entire constellation of beliefs, values, techniques and so on shared by members of a given community" and also to (b) "the concrete puzzle-solutions which...replace explicit rules as a basis for the solution of the remaining puzzles of normal science." Note that the second, specific conception enables the former general conception and thus is the deeper of the two.

Similarly, memetics is used in both a general and a specific manner. Specifically, we have the model of actions or artifacts that can be imitated with fidelity, fecundity and durability. Generally, we have the conception of memes as actions or artifacts that evolve independently of the people producing them. The specific conception enables the general one and is the deeper. But just like there are a host of biological studies that rely on but do not discuss evolutionary processes, there are also a host of possible memetics studies that rely on but do not discuss evolutionary processes. Furthermore, the methods used in general memetics studies may be directly applicable to the fundamental questions raised within the specific conception. This occurs especially when studies have a methodological rather than a theoretical focus. Thus, a vast array of ecological, geological and genetic models depends on genetic evolution without addressing it and many mathematical models in these fields preceded evolutionary models by decades. The power of evolutionary thinking was due as much to the remarkable convergence of knowledge that supported it as to its own models. It seems likely that a future principle of the memetics of science will be that converging knowledge is a necessary predecessor to the dominance of any scientific system.

### **1.2 Edmonds' First Challenge: The Case Study**

Critical to Edmonds' argument is that passing his challenges would show the "usability" of memetics. Edmonds' suggestion of nursery rhymes and legal phrases as apt test cases was based on the assumption that the examples were "of a limited nature about which good quality data is available." More ambitious studies would not be believed until such "straightforward" examples had been established.

Case studies are certainly essential, but Edmonds' suggested topics can be improved. Unambitious content by no means guarantees straightforward conduct and interpretation of any study. The failure of attempts to program translation, transcription, and even textual search belie the straightforwardness of any natural language task. Especially at this early point in memetics research, the methodology and interpretation of all studies will be questioned. The critical factor for the future of memetics is not straightforward conduct and interpretation. Rather, the critical

factor is whether the questions lead the scientific community to ignore or use and debate the results. The potential for use and debate, rather than obscurity, varies directly with the author's skill, the author's community support and the significance of the topic. Furthermore, the dynamics of the transmission of memes will be affected by the importance of the topic. Observing natural memetic transmission and adaptation is likely to be far more difficult for small niche topics, like those Edmonds suggested, than for topics that have more urgent and widespread social content. A primary goal of this paper is to show that although the quick-fix case study does not exist, there are methods that make the consequential studies no more difficult than those with minimal importance.

### **1.3 Edmonds' Second Challenge: The falsifiable theory of when to use memetics**

In his challenge for a falsifiable theory of when to use memetics, Edmonds rightly argues for a biological advantage of large brains independent to the hosting of memes. The issue of deciding when to apply memetics arguments, however, might be better conceived as methodological than theoretical. Though a theory might be easier to apply to ongoing social change, researchers would benefit more from operational definitions. Once several operational definitions have become established, the common elements would point to the needed theory.

The ideal, operational definition would involve direct observation of imitation. Short of that ideal, indirect studies could use situations involving developmental or historical change that (a) minimize the influence of intra-individual processes and (b) allow for a high level of opportunity for imitation. One way of precluding intra-individual processes is to study only a single record per person. A high level of opportunity for imitation occurs by examining situations that contain a high probability of frequent contact between persons being studied.

## **2. Competing Memes Analysis**

Competing Memes Analysis is a method that makes studies of consequential topics no more difficult than those with minimal importance. It is a three step process that

- (1) organizes activities into groups of competing memes that often emerge in developmental successions,
- (2) codes records of activities for the presence or absence of each meme, and
- (3) constructs models of the changing frequencies of the memes.

The methodology has been constructed through studies of the ontogenetic development of drawing (Dirlam, 1996, contains a preliminary analysis), writing (Dirlam, 1982), and the historical development of research methods in developmental psychology (Dirlam, Gamble, and Lloyd, 1999). The purpose of the drawing and writing studies was to design objective pedagogical alternatives to educational testing that could be used with natural products of classroom activities. The purpose of the study of research methods was to determine whether the analytical models that fit the ontogeny of drawing and writing also fit historical development.

## **2.1. Creating multidimensional classifiers from successions of memes**

The first step in the Competing Memes Analysis of an activity is to create multidimensional classifiers of competing memes (Dirlam, 1980). Each activity is assumed to exhibit numerous small groups of memes each of which comprises a dimension. A group of memes is a dimension when each meme within the group competes with all other memes in it and can be combined with any meme from any other group (see Appendix A for an example). The succession of memes that occurs with increasing experience can be a powerful clue to identifying competing groups. A memetic classifier for an activity consists of several dimensions of memes.

Memetic classifiers can be readily constructed by anyone with modest expertise in a topic. A straight forward approach is to observe differences between people performing the activity who have various levels of experience. When one approach to the activity is commonly replaced by another, the two approaches are competing memes. Sometimes the later appearing memes contain earlier appearing ones. In any case, since the memes can be defined to be mutually exclusive, each succession becomes a dimension.

Dimensions can and should be defined as exhaustive by generalizing to include all other cases in the definition of one meme within each dimension. Thus, every record of an activity being investigated is assigned to one meme within each group of competing memes. When an emerging meme becomes frequent enough to separate from a generalized meme, it will be necessary to redefine and recode the dimension. Analysis of large collections of records might also reveal that some memes are so rare that they should be logically combined with others. Experts in an activity can help to construct definitions of each meme so that coders can reliably distinguish which is being used. For most important activities, there is theoretical literature that suggests a rich variety of memes. For example, more than half of the dimensions of competing memes used in the study of developmental research methods mentioned above were based on concepts from Danziger (1990). The memetic classifier for the study is presented in Appendix A. The study involved coding 912 articles written from 1930 to 1992 for the presence or absence of each meme in the classifier.

The number of dimensions in a competing memes analysis is a matter of researcher choice and depends on the state of knowledge in the subject at the time of the study. It should be noted, however, that gathering and examining records is usually much harder than coding them. Therefore, studies with several dimensions require considerably less resources per dimension than single-dimension studies. The memes themselves are like fractals—they can apply to content as fine-grained as words, lines and study locations or as general as complete discourses, complete drawings and complete research articles.

## **2.2. Collecting and coding records**

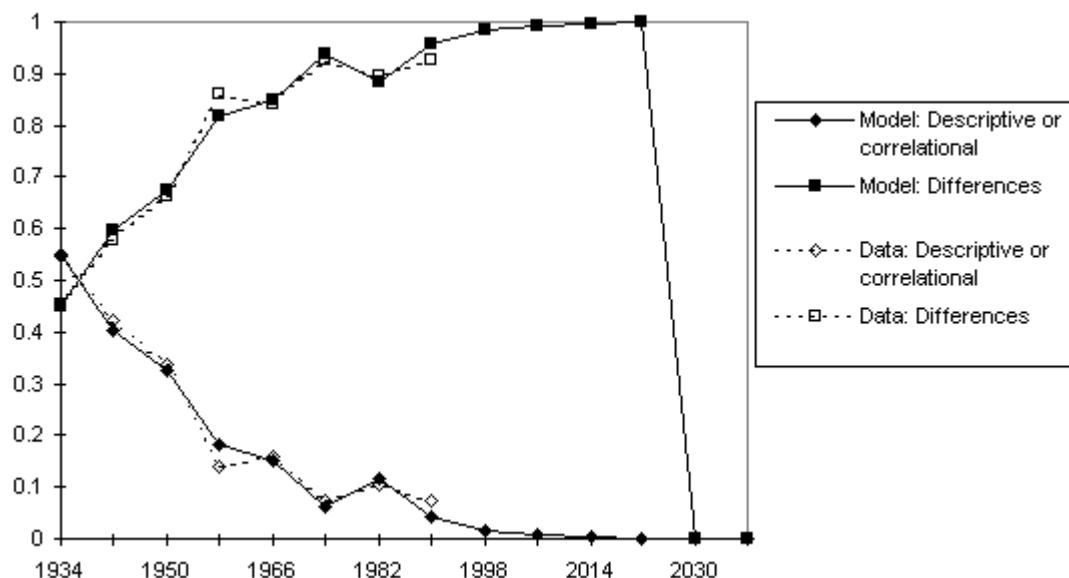
The second step in Competing Memes Analysis is to systematically collect and code records of memes. In order to permit the construction of testable models, the records should be drawn from a corpus that contains a meaningful point in time or space for all samples. For example, time points for the studies of drawing and writing were the age of each participant and time points for

the developmental research study were the publication years of the articles. Studying the spread of memes from their spatial points of origin should also generate rich, useful results. Drawings are self-contained records of the actions of those who made them. Similarly, writing samples record the actions of writers and research reports record the actions of researchers. Audio and video tapes are also records of memetic activity. Because memes are imitated, any user of a meme and many non-users can identify uses of it by others. This means that discrete samples can be reliably coded for the presence or absence of a meme. If a ten-dimensional system has been defined (as in section 2.1 above), each record will contain ten memes.

One clear implication of memetics is that humans are uniquely capable of identifying memes. The only rival species, apes and parrots, are readily surpassed in speed, variety and complexity of imitation by preschool children. Examining the diversity of memes in the three studies that have used Competing Memes Analysis reveals that the complexity or importance of memes did not affect our ability to identify them. Hence, consequential studies are no more difficult than those with minimal importance.

### 2.3. Modeling frequency changes over time or space

Once records have been coded, it is possible to count the uses of memes in situations with important consequences, whether they be economic (e.g., the use of a new product), social (e.g., the spread of juvenile crime), political (e.g., the repetition of a candidate's message), educational (e.g., the way students and programs are evaluated) or scientific (e.g., the spread of a methodology). For example in the study of developmental research methods, the model projected that the growth of difference statistics was so fast that if a competitor did not emerge in the next few decades, data analysis would consume so much of available resources (i.e. social acceptance) that it would implode (see Figure 1). During the 1930-1992 study period, data modeling was too rare to analyze separately from difference statistics. But since it is a valued and slow-growing alternative, it is a possible solution to the collapse of the analysis dimension.



**Figure 1.** Best fitting Lotka-Volterra model for competing data analysis memes

Memetics skeptics would most readily accept studies that operationally define their memetic nature by setting up situations for observing imitation. However, imitation can be inferred rather than directly observed, and can be interpreted broadly as reproduction. This opens up the potential to use archival data (e.g. research reports, crime incident reports, corporate audits and so forth). A dramatic memetics success would occur if memes were discovered that effectively competed with such practices as crime among low income juveniles, self and community destructiveness among religious fundamentalists and fraud among corporate CEOs faced with poor results. Key evidence for successful competition would come from changing frequencies of memes over time or space.

Beyond identifying memes, Competing Memes Analysis may involve a search for

- (1) the path of succession from one meme to the next,
- (2) the resources required,
- (3) the growth rate and
- (4) the competitive strength of such practices.

Results of these types of studies may help us to understand such issues as

- (a) how memes that are seen as "weeds" contribute to general survival and
- (b) how the harmful effects of overgrowth of such memes can be controlled.

An example of (a) would be marketing processes that make it possible to establish products in extremely competitive markets, but permit monopoly-like overgrowth in noncompetitive markets. Examples of (b) would be finding competitors to the emergence of monopolies ranging from antitrust laws to innovations that disrupt them. The problems solved while undertaking a rich variety of such studies could provide the sort of clarity needed to meet Edmonds' challenge for a falsifiable theory of when to use memetics.

Definitive answers to questions about resources, growth rate and competitive strength may be difficult to obtain without detailed experimental studies. However, once hundreds of records of memes have been systematically collected from a population, placed within a meaningful distribution of time or space, coded and counted, it is possible to develop models of the frequency changes. Such models can reveal characteristics of memes that are difficult to observe directly. They also make detailed, testable predictions.

A model that influenced the competing memes analysis, proposed here, is the Lotka-Volterra model of species competition (see Appendix B). This model describes the population of species that compete within an ecosystem as depending on four parameters: the initial population, the maximum sustainable population, the growth rate, and the competitive strength. Given these parameter values, at least four life cycles of memes can be identified (see Appendix A). Rapid growth rate is high enough to create chaotic fluctuations in a noncompetitive environment. Data

analysis modeled in Figure 1 is an example. High competitive strength reduces the rapid growth of competitors to moderate levels or less.

Life cycles	Parameter values		
	Initial prevalence	Growth rate	Competitive strength
Default	High	Near zero	Near zero
Niche	Low	Slow	High
Pioneering	Low	Rapid	Low
Dominant	Low	Moderate	High

**Table 1.** Lotka-Volterra values for four memetic life cycles

Once such vital statistics of a group of memes have been worked out, scientifically significant and socially valuable experiments can be conducted to determine the factors that influence the long-term vitality of particular memes. For example, revisiting the species analogy, we note that pioneering species are often those that adapt to the presence of many predators by growing very rapidly. If such species find an environment where the predators are absent, their populations can grow so rapidly that they consume too much for the environment to sustain them. The formerly adaptive high growth results in a local extinction. It takes little imagination to realize what powerful social consequences would occur when the analogous process operates mimetically. Pioneering growth may be as essential to responding to social threats or establishing new products as it is to establish species in predatory environments. But planners need to be aware of the social or economic danger that these new responses or products pose if their memetic competitors were to suddenly vanish.

In conclusion, Competing Memes Analysis provides memetics with a method for conducting precise studies of memetic processes found in any human activity. It can be summarized as a 3 step process. Step 1 identifies the organization of memes within an activity. Each activity is assumed to exhibit numerous small groups of memes where each meme within a group competes with all other memes in the group and can be combined with any meme from any other group. The succession of memes that occurs with increasing experience can be a powerful clue to identifying competing groups. Step 2 collects records of activities and codes them for the presence or absence of each meme identified in Step 1. Step 3 analyzes changing frequencies of each coded meme over time or space. Models of these changes can give useful clues to suggest empirical studies that will provide important social and scientific results.

## Appendix A: Memetic system for the historical development of developmental research

DIMENSIONS	Life cycles (see Table 1 for explanation of life cycles)			
	Default	Niche	Pioneering	Dominant
<b>DEPENDENT VARIABLES</b> What was measured in the study?	Limited Behaviors Easily counted, or Categorized Coded free behavior	Summed or Enumerated Standardized tests or sums of ratings of loosely connected items)		Categorized and Other coded free behavior combined with a second dependent variable
<b>DATA ANALYSIS</b> What kind of statistics were used?		Descriptive Statistics like counts, means or correlations	Difference Statistics t-tests, ANOVAs non-parametric difference tests or modeling	
<b>DESIGNS</b> How often were comparable measurements taken?	One Session Per Task	Micro-longitudinal Repeated measurements taken weeks apart	Longitudinal Repeated measurements taken years apart	
<b>AGE</b> Was it used to measure time or to assign people to groups?	Single Age Group	Dependent Variable Using age as a measure of how long it takes to develop		Multiple Age Groups
<b>SOCIAL CONTEXT</b> Who was present with the people being studied?		Significant Other Alone or With Test	Test Alone or With Experimenter	
<b>LOCATION</b> Where was the study done?	Unspecified	School, Home, Other, Lab, or Multiple		
<b>BACKGROUND</b> How many fields used as sources?			Interdisciplinary	Disciplinary Only one
<b>APPLICATIONS</b> Who benefited from the study?	Researchers Only			Others Professionals or individuals

**Table 2.** The life-cycles and dimensions of a Competing Memes Analysis for developmental research methods

## Appendix B: Lotka-Volterra competition

Lotka (1925) and Volterra (1926) independently formulated competition between species in an ecosystem. The basic idea is an expansion of Verhulst's logistic law of a century earlier, which in turn built on Malthus' concept of exponential growth, where the population at a particular moment in time equals the population at the preceding moment multiplied times one plus the growth rate.

$$(1) x' = x * (1+r)$$

Populations do not grow beyond the available resources, but rather either stabilize at an equilibrium that depends on the resources or collapse after exhausting the resources. Letting  $k$  denote the equilibrium population, Verhulst corrected Malthus' notion by reducing the growth rate as the population approached the equilibrium.

$$(2) x' = x * [1 + r * (1 - x/k)]$$

When the population  $x$  equals the equilibrium population  $k$ , the growth rate will be 0. The result for slow or moderate growth rates is the well known S-curve that reaches an asymptote at the equilibrium. Also, this is the equation that produces wild fluctuations with very high growth rates. Its age, simplicity and practical usefulness have made it a favorite of chaos theorists.

Lotka and Volterra further refined Verhulst's equation by reasoning that a competing species would further reduce the growth rate. They argued that the effect of each competing species would depend on its population,  $y$ , multiplied by a characteristic competitive strength factor,  $c$ . Of course, each competing species would reduce the population of the target species.

$$(3) x' = x * [1 + r * (1 - x/k)] - \sum c_i y_i$$

Since, the model is commonly applied to discrete breeding cycles, it is easy to develop a spreadsheet program where the populations at a particular point of time are found in a single row and are used to calculate the populations in the successive moment in the following row. Further details of applying the model to developmental and historical data can be found in Dirlam, Gamble, and Lloyd (1999).

## References

- Aunger, R (2000). Conclusions. In R. Aunger (ed.), *Darwinizing Culture: The Status of Memetics as a Science*, Oxford, Oxford University Press, 205-232.
- Danziger, K. (1990). *Constructing the Subject: Historical Origins of Psychological Research*. Cambridge: Cambridge University Press.
- Dirlam, D. K. (1980). Classifiers and cognitive development. In S. & C. Modgil (Eds.), *Toward a Theory of Psychological Development*. Windsor, England: NFER Publishing, 465-498.

Dirlam, D. K. (1982). Theoretical Framework. In NE NY Board of Cooperative Educational Services, The Second R. Bureau of English Education, New York State Educational Department.

Dirlam, D. K. (1996). Macrodevelopmental analysis: From open fields to culture via genres of art and developmental research. *Mind, Culture, and Activity*, **3**, 270-289.

Dirlam, D. K., Gamble, K. L., & Lloyd, H. S. (1999). Modeling historical development: Fitting a competing practices system to coded archival data. *Nonlinear Dynamics, Psychology, and Life Sciences*, **3**, 93-111.

Edmonds, B. (2002). Three Challenges for the Survival of Memetics. *Journal of Memetics - Evolutionary Models of Information Transmission*, **6**. [http://jom-emit.cfpm.org/2002/vol6/edmonds\\_b\\_letter.html](http://jom-emit.cfpm.org/2002/vol6/edmonds_b_letter.html)

Kuhn, T. (1969) *The Structure of Scientific Revolutions*. Chicago: University of Chicago Press

Lowenfeld, V. (1957). Creative and Mental Growth(3<sup>rd</sup> ed.). New York: Macmillan.

Moffett, J. (1968). Teaching the Universe of Discourse. Boston: Houghton Mifflin.

Piaget, J. & Inhelder, B. (1967). The Child's Conception of Space. New York: Norton.

## Acknowledgments

Most of the integrating work for this paper was accomplished in 1997-1998 while I was a James McKeen Cattell Fellow at the Laboratory of Comparative Human Cognition of the University of California San Diego. I would like to thank Michael Cole for comments on earlier versions of this manuscript and for his extensive advice and counsel over the last seven years of this project. Kurt Danziger, James Moffett, and Jerry Balzano also provided important theoretical insights on the road to refining Competing Memes Analysis. The readability and usefulness of this article was greatly improved by the careful editorial commentary of Martin De Jong. Paul Marsden also offered numerous very useful suggestions. Especially, the treatment of pioneering growth was more balanced as a result of his involvement.

# USING MEMETICS TO GROW MEMETICS

By  
**David K. Dirlam**  
Memetics Consulting  
414 Whitaker Street  
Savannah, GA 31401 USA.  
[ddirlam@memeticsconsulting.com](mailto:ddirlam@memeticsconsulting.com)

## 1. Crises in Memetics and Psychology

As sciences, psychology predates memetics by a century. But both are facing crises. For memetics, this issue marks the end of the first incarnation of the Journal of Memetics. In psychology the United States government is reorganizing its National Institutes of Mental Health. It will now focus on helping people with mental and behavioral disorders while substantially narrowing its support for basic behavioral research in social psychology, personality, learning and memory and animal studies. The February issue of American Psychological Society Observer devoted both its feature article and its presidential column to the crisis. APS Fellow Daniel Gilbert of Harvard was quoted saying “This is an absolutely devastating blow to the infrastructure of basic science.”

The two crises are of fundamentally different sorts. The crisis in memetics stems from a dearth of empirical studies. In an age when governments, universities and foundations support empirical studies, a science without a track record has a difficult time raising funds. Established sciences, like psychology, are stronger competitors. Memetics may be a future oak tree, but currently the forest of established social sciences is using up its resources.

The crisis in psychology actually stems from a problem identified more than a half century ago by one of the greatest of early psychological scientists, Karl Lashley (1951). Psychology has no true unit of analysis for spontaneous and complex human behavior. Reflexes (including neural responses), conditioned responses, test and questionnaire responses all restrict behavioral possibilities so much that researchers are often unable to distinguish their own biases from the behaviors being studied. When compared to other sciences, these restricted-possibilities units have caused psychology to be more fraught with unresolved conflicts and less clear about scientific community directions for addressing societal problems related to its topic.

De Jong, Konstantions and Mamadou (2002) found convincing evidence that system upheaval or severe performance crises are more fertile breeding grounds for pervasive change than other occasions. They conclude that “reformers do best when they are prepared for oncoming windows of opportunity and save their energies to act at such moments.” Though a harrowing time for the psychological research community in the United States, psychology’s crisis is a window of opportunity for memeticists to begin demonstrating how its evolutionary approach can bring practical results that people hoped psychology would address more effectively than it has.

Key to realizing the opportunity for memetics is establishing the meme as the unit of analysis for

spontaneous and complex human behavior. The view that memes are methodologically complex hampers their growth as such a unit. There are two aspects of this supposed complexity: the research triad and evolutionary data. The research triad refers to the fact that identifying memes requires three people: a model to exhibit the meme, a learner to replicate it and an experienced judge to reliably determine whether the model and learner are producing the same meme.

Replacing models with artifacts, such as videos, texts or diagrams will reduce some of the methodological complexity in the research triad. The fundamental aspect of this complexity, however, stems from the beginnings of psychology. The community trying to establish psychology as a science made an early mistake in attempting to simplify the research setting. Since then it has practiced that mistake so long, that it may as difficult to undo as to start over as a science. Danziger (1994) documented how psychology eliminated the judge from the experimental paradigm. Memetics theory, however, implies that reliable judges are not the problem. The very foundation of memetics is the fact that humanity evolved a unique capacity to reproduce the actions of other humans. The ability to judge the similarity of memes, therefore, is an evolutionarily maximized quality of humanity.

In support of the memetics implication that reliable judges are not the problem in psychological research is the fact that even the mighty computer is no competition for the human in the realm of judging the similarity of memes. The poverty of computerized speech recognition and translation compared to recognition and translation even by experienced children proves the poverty of the “eliminate-the-judge” meme for psychological research. If the computer cannot assess behavioral similarity, surely the lowly test and questionnaire are not up to the task. Memes are not methodologically over-complex; rather psychology as a science is suffering from an artificial assumption of methodological simplicity.

The second aspect of methodological complexity of the meme as unit for spontaneous, complex human behavior is its emphasis on evolutionary data. Dirlam, Gamble and Lloyd (1999) showed that developmental psychology systematically collected data on gradual changes over time only with respect to physical growth. Worse, the field gradually eliminated such studies a half century ago in favor of simplifying to one, two or a few developmental data points. Other social sciences have addressed the distribution of behaviors in space, but psychology does not even consider such distributions an issue. With such a methodological bias, it is no wonder that psychology has missed the memetic evolutionary aspects of human behavior. Again memetics is not methodologically too complex; rather, psychology is methodologically too simple.

For spontaneous and complex human behavior, the meme is a better unit of analysis than any that psychology has constructed. The problem, however, is how to grow usage of the meme in the context of overwhelming hegemony of those previously constructed units.

## 2. A Memetics Solution

Memetics can contribute to the growth of the meme as a unit of analysis in psychology. First, reflexes, responses, tests and questionnaires are all memes. Dirlam (2003) showed that the growth of memes obeys ecology’s Lotka-Volterra model of growth. Thus, four general factors control the growth of memes: initial strength, characteristic growth rate, resource availability and

competitive strength. In current strength, psychology's restricted-possibilities units are several orders of magnitude stronger than the meme. The problems with psychology's oversimplification of the research setting strongly suggest that the meme's characteristic growth rate and competitive strength are greater than psychology's restricted-possibilities units. This means that those interested in growing the meme must focus on the resources.

Certainly the scientific apparatus of journals, university departments, funding agencies and societies are resources. But memetics allows us to look at a more general level of the resource problem. We have a thorough conception of the general resources required by species, such as temperature, food, water etc. But at this general level, what are the resources that lead to meme growth? The psychologist, Albert Bandura (1977) devoted his career to the study of "observational learning" or imitation. In his analysis, there were four essentials: attention, retention, reproduction and motivation. This is a good starting point for identifying resources for meme growth. If the "meme-as-a-unit-of-analysis-for-psychology" meme is to spread, memeticists need to provide it with attention, retention, reproduction and motivation. Memetics history of self critique suggests that motivation is the central problem. Therefore, Bandura's criteria will be discussed in the reverse order.

## 2.1. Motivation

Judging from the philosophical articles on memes in both the Journal of Memetics - Evolutionary Models of Information Transmission and in collections of readings such as Aunger (2000), the central resource for growing memetics is motivation. Each of the authors had clearly attended to, retained and reproduced memetic thinking. But the question of long-term support seemed ever present. In Aunger's words "The ultimate test—which would pre-empt theoretical objections—is whether memetics can produce novel empirical work or insightful interpretations of previous results. It has not yet done so, but must do so in the near future. Otherwise, it is likely that memetics will be perceived to be a misguided enterprise. The clock is ticking." The impression is that there is a plethora of theoretical critiques and a scarcity of data. But what would motivate a change to memetic data collection both within and outside of the memetics community?

The most insightful article on motivation for change that I have encountered was the brilliant analysis of dietary change by economist/public policy analyst, Carl V. Phillips (1999). When we eat, we start with a strategy that will satisfy our needs for good nutrition, good taste, ease of obtaining and familiarity (four key resources for eating). The high fat, high sugar diet that we get at fast food chains is an example. As we practice, we get better and better at satisfying our needs with this food. In time, we get so good that when we try to change, we are usually less satisfied. From a nonlinear dynamics point of view, we have found a relative maximum. Trying to cut back within our favorite diet leaves us feeling hungry, resulting in a lower overall utility value. We easily backslide into our old habit patterns.

Eventually for some people, the strategy fails. They take a step back and consider making a radical change. They might try a vegetarian diet. This whole new strategy results in a whole new set of experiences, at first not as satisfying as the old diet. As they keep trying, they gradually get better at satisfying their four needs with the new approach. While they are working at satisfying their needs with the new approach, they frequently backslide into the old one. When they get as

good at satisfying their needs with vegetarian food as they did with fast food, however, they rarely backslide. A radical change in strategy that needs practice to get right might be called a “memeplex replacement.” Evidence presented in the next section shows that it applies not just to eating, but to all the life strategies used in complex human activities.

For memetics to succeed in psychology there needs to be a radical change in the activities of psychological researchers. To summarize the dietary analogy, people are dying from the “fast food” diet. That is motivation to try the leap to the vegan diet. Trying to maximize the four utilities without radical change will not work. It helps to have models who have shown that they can successfully maximize their utilities with the new approach, but each individual must commit to the change and then learn to make it work. To apply these concepts to psychology, maximizing the utilities of psychological research will not motivate people to change. A crisis, such as the US NIMH (the National Institutes of Mental Health) precipitated, must motivate researchers first to commit to change memeplexes and then follow up by learning how to make the new approach work.

People have died from fast food for decades. That fact has not been sufficient to create change. A rich set of data has confirmed that fast food impairs people’s health and lifestyle. A public critique of fast foods is growing. The food industry is creating new foods in response and there is a growing list of healthy food cookbooks coming from publishers. There is less data about the failings of psychology, but a public critique is beginning to emerge. Several core concepts in psychology lack hard empirical support. These include developmental stages, intelligence, the genetic transmission of psychological traits, and psychological disease. Also, if there really was a science of mind that worked, social scientists would be as much in demand in corporations and governments as biologists and engineers. Instead, at least in the U.S. policy makers seem committed to seeking biological solutions to social problems, no matter how weak the solutions may be.

Much of the early writing about memetics has involved philosophical critiques. If someone wants to help some friends change their diets, it is certainly useful to avoid suggesting a diet that would make matters worse. But the critiques of memetics have focused too much on the problems of memetics and not enough on the problems of the alternatives.

Clearly many memetics solutions are not going to be imminent, but some are. For example, the developmental stage is an almost mystical concept that at best prevents people from trying to teach concepts to children without adequate preparation. If psychologists replaced the concept of “stage” with one of memeplex change, however, it would cause educators to look for the utilities involved in the activity and to create “recipe” books that might help learners to maximize these utilities.

Likewise, psychological “disease” is a mystical concept that causes people to look for medications that alleviate the symptoms. That is roughly equivalent to solving addictions by giving drugs that block the effects of the addicted drug. A perspective that leads to some interesting possibilities for memetics is to view psychological disorders as self maintained addictions – people learning to manipulate their own neurotransmitters until their brains begin to malfunction. Such a view would motivate clinical researchers to look at the utilities of the

neurotransmitter manipulation and at how to create the commitment to change. Ultimately it might lead to developing “recipe books” for how to achieve those utilities without manipulating the neurotransmitter. As an example consider that television watching produces attentional problems (Christakis, Zimmerman, DiGiuseppe, and McCarty, 2004). In the U.S. more and more children are being given drugs for attentional problems. From a memetics point of view, a more lasting and less damaging solution would be to discover ways to serve children’s entertainment needs other than television.

Many psychologists are already struggling against the current diet of such “fluff” concepts as developmental stages and psychological diseases. They are working in areas such as naturalistic observation, qualitative data analysis, ecological psychology, observational learning, etc. The memetics community should invite such psychologists to contribute to memetics methodology.

## 2.2. Reproduction

From an empiricist’s point of view, empirical research in memetics is currently difficult to reproduce. No one has yet written the methodological recipe book for memetics research. A good starting point is to make analogies with ecology. Before attempting ecology’s Lotka-Volterra model, I spent a frustrating decade looking for a mathematical model that fit data from over 1,200 drawings made by 5 to 18 year olds and coded for nearly 30 drawing qualities by three independent judges. Probability theory offered a reasonably good fit using the generalized gamma law developed in the context of radioactive decay from one element to another. But why development would proceed by decay was a mystery. The ecology equation, however, fit the data like a statistical skin-tight glove. When the same equation fit data on the historical development of developmental research methodologies (as coded by two independent judges in nearly 1000 journal articles written from 1930 to 1992), it opened up several powerful theoretical insights:

- A competing-agent action of the qualities in both settings fits well with the meme concept.
- An outside-of-the individual status of the qualities also fits the meme concept.
- Decay happens because later memes compete with earlier ones.
- “Development” refers to changes in the frequency of use of competing memes described above as memeplex replacement.
- An easy to study developmental succession of memeplexes replaces the fluffy stage concept.
- The easily researched process of commitment to radical change followed by relearning readily replaces the inferred process of developmental stage acquisition.
- Controlling factors of meme populations are initial frequency, growth rate, resources and competitive strength.

Given constant resources, four growth patterns identify four categories of memes:

- Default – high initial strength, low growth rate and competitive strength.
- Pioneering – high growth rate (enough to exhaust all resources in the absence of competition), low initial and competitive strengths.
- Niche – high competitive strength, low initial strength and growth rate.
- Dominant – high competitive strength, low initial strength, moderate growth rate ((not enough to exhaust all resources in the absence of competition).

Other analogies from ecology should produce similarly powerful insights not only for human development but for memetics research related to all the social sciences.

### 2.3. Retention

Key factors that influence retention are repetition, organization, context similarity and distinctiveness. Repetition, organization and context similarity factors depend on growth. Based on Google searches “Memetics” currently appears on the internet about 0.1% as often as the term “internet” itself (controlling for growth in the internet as a whole), as often as “psychotherapeutic”, 3% as often as anthropology or sociology and 1% of psychology. It’s ranking among such terms grew rapidly last fall, but has declined this spring. The crisis represented by the journal change may be a key reason for the recent decline. Since there are none of the other traditional academic resources mentioned above for memetics, a large impact of a change in the only journal might be expected. In this context, closing the journal would be roughly akin to people aspiring to change from a fast-food to a vegan diet by throwing out their only vegan recipe book.

The remaining factor is distinctiveness. There have been numerous distinctive contributions to memetics. Topics ranging from anti-religion and anti-self to paganism and marketing are reminiscent of the early days of psychology amplified by a century-and-a-half of self preoccupation. Few of these distinctive contributions, however, have an empirical basis. The deeper question is whether there is a distinctive memetics methodology.

Two candidate methods that set memetics apart from the other social sciences have already been mentioned. The first is the model-learner-judge triad and the other is measurement of the frequency distribution of memes in time and space. Again a comparison of ecology and psychology is illustrative. I counted the frequency of tables and charts in a standard textbook of ecology and one of psychology. From beginning to end of the ecology textbook, about 75% of the charts and tables involved distributions in time or space. The same was true for the first portion of the psychology text that dealt with sensation, perception and conditioning. The last portion of the psychology text, however, dealt with personality and social psychology. For that portion, the ratio was reversed—only 25% of the tables and graphs involved distributions in time or space. Because psychology views the mind as being inside the brains of individuals, it does

not consider mapping the distribution of complex human behaviors in time and space. Because memetics views the human mind as dependent on relationships with other humans, past and present, distributional data is central to memetics methodology. Distributional data made the theory of evolution both possible and established. It will also yield the distinctive findings needed to establish memetics.

#### **2.4. Attention**

Memetics needs empirical studies. Empiricists pay attention to empirical results. Those that attract the most attention solve intractable societal problems. If memetics is to help solve such problems it will be through its emphasis on adaptive growth. Growing a reduction in crime within a city, a more equitable way to distribute food in an underdeveloped country, a reliable reduction in dependency on addictive drugs or self-manipulated neurotransmitter changes in individuals, or an improvement of skill in a school population would all be spectacular results that would certainly attract attention. If memetics researchers tackled the worst problems in their environment by collecting evolutionary data using the model-learner-judge triad, the results would establish memetics as a science.

### **3. Conclusions**

There is a window of opportunity for memetics resulting from the current crisis in U. S. funding for psychology. To realize the potential of this window memetics must provide other social sciences with the four essential resources for imitation: motivation, reproduction, retention and attention. The discussion showed that the following four strategies will grow the use of memes in the social sciences:

- (1) Motivate social sciences by critiquing current paradigms, by calling for a commitment to a radical shift of units of analysis to those involving experienced human judges and the collection of evolutionary data and by encouraging methodological articles in memetics by those already using such units.
- (2) Help researchers reproduce memetics by providing methodological models which borrow heavily from analogs of approaches found to be successful in ecological and evolutionary research.
- (3) Help researchers retain the memetics approach by repeatedly emphasizing the model-learner-judge triad and the charting of changes in frequency across time and space.
- (4) Attract researcher's attention to memetics by tackling spectacular problems.

### **References**

Aunger, R (2000). Conclusions. In R. Aunger (ed.), *Darwinizing Culture: The Status of Memetics as a Science*, Oxford, Oxford University Press, 205-232.

Bandura, A. (1977). *Social Learning Theory*. Englewood Cliffs, NJ: Prentice-Hall.

Danziger, K. (1990). Constructing the Subject: Historical Origins of Psychological Research. Cambridge: Cambridge University Press.

De Jong, M., Konstantions, L. and Mamadouh, V. (2002). The Theory and Practice of Institutional Transplantation: Experiences with the Transfer of Policy Institutions. Dordrecht: Kluwer.

Christakis, D. A. Zimmerman, F. J., DiGiuseppe, D. L., and McCarty, C. A. (2004). Early Television Exposure and Subsequent Attentional Problems in Children. *Pediatrics*, **113**, 708-713

Dirlam, D. K. (2003). Competing Memes Analysis. *Journal of Memetics - Evolutionary Models of Information Transmission*, 7. <[http://jom-emit.cfpm.org/2003/vol7/dirlam\\_dk.html](http://jom-emit.cfpm.org/2003/vol7/dirlam_dk.html)>

Dirlam, D. K., Gamble, K. L., & Lloyd, H. S. (1999). Modeling historical development: Fitting a competing practices system to coded archival data. *Nonlinear Dynamics, Psychology, and Life Sciences*, **3**, 93-111.

Lashley, K.S., (1951). The problem of serial order in behavior. In L.A. Jeffress (Ed.), *Cerebral mechanisms in behavior*. New York: Wiley, 112-136.

Phillips, C. V. (1999) Complex systems model of dietary choice with implications for improving diets and promoting vegetarianism. *American Journal of Clinical Nutrition*, **70**(3), 608S-614S.

# SIMULATING COLLECTIVE MISBELIEF

By  
Jim Doran (1998)

Journal of Artificial Societies and Social Simulation vol. 1, no. 1,  
<http://www.soc.surrey.ac.uk/JASSS/1/1/3.html>

To cite articles published in the *Journal of Artificial Societies and Social Simulation*, please reference the above information and include paragraph numbers if necessary

Received: 29-Oct-1997   Accepted: 13-Dec-1997   Published: 3-Jan-1998

## Abstract

It appears that what the agents in a multiple agent system believe is typically partial, often wrong and often inconsistent, but that this may not be damaging to the system as a whole. Beliefs which are demonstrably wrong I call *misbeliefs*. Experiments are reported which have been designed to investigate the phenomenon of collective misbelief in artificial societies, and it is suggested that their results help us to understand important human social phenomena, notably ideologies.

**Keywords:** Multiagent system, collective belief, ideology

## Introduction

### 1.1

Using computers to simulate naturally occurring societies is a rapidly growing area of research (Gilbert and Doran, 1994; Gilbert and Conte, 1995). Much of the ongoing work is set at the level of societies in which each individual is relatively low on the cognitive scale, but some aims to gain insight into human society, addressing such issues as social action, cognized models, group formation, planned cooperation and "macro-level emergence". The techniques of distributed AI (O'Hare and Jennings, 1996) may be used to support this latter work.

### 1.2

Some explanation and justification of methodology is appropriate. By working in the computational domain we are able to study multiple agent systems *independently* of any particular modeling interpretation. We can establish rigorously and objectively what consequences, including non-obvious consequences, flow from what assumptions. We can therefore hope to develop an abstract theory of multiple agent systems and then to transfer its insights to human social systems, *without an a priori* commitment to existing particular social theory. Of course, at a certain level assumptions must be built into whatever systems we create and experiment with, but the assumptions may be relatively low level, and their consequences (including emergent properties) may be discovered, not guessed.

## Agents and Artificial Societies

### 2.1

Agents may be characterized as computational mechanisms situated in an environment that repeatedly select and perform actions in the light of their current input from the environment (perception) and their current internal state. In general, the actions that agents may perform include the sending of messages to other agents.

### 2.2

A variety of different agent "architectures" have been designed in recent years and their properties partially explored (see, for example, Wooldridge and Jennings, 1995; Wooldridge, Muller, Tambe, 1996). One simple type of agent architecture, often called a *reflex agent* (Russell and Norvig, 1995), comprises the following main components:

- A *working memory* (comprising a changing set of tokens, determined by processes of "perception")
- A set of *rules* of the form IF <CONDITION> THEN <ACTION>. Such rules are often called "situation-action" rules or "condition-action" rules. The condition part of the rule typically specifies a required conjunction of tokens in the working memory (the specification may involve variables which must be bound to tokens in a consistent way) and the action part an executable procedure.
- A process that repeatedly identifies a rule whose condition part matches the contents of the working memory, and then executes that rule's action procedure -- thereby causing the agent to perform one or more actions in its environment.

Thus the essential functioning of the agent is that the results of perception are deposited as tokens in the working memory, and matching rules are "fired" -- with the result that actions are performed by the agent in its environment. The heart of such an agent is thus the set of condition-action rules which it contains.

### 2.3

A standard extension to this design is to add a semi-permanent body of (symbolically coded) information to the working memory, an *agent memory*, and to include rules which update the contents of both the working memory and, from time to time, the agent memory. The agent is then, in effect, able to maintain and update some sort of representation of and beliefs about its environment. This type of agent, which may be called an *extended-agent*, will be that at issue for the remainder of this paper.

## 2.4

A computational multiple agent system is then a set of agents (here extended-agents) which share an environment and which interact and possibly pass messages in some form one to another. The agents in a computational multiple agent system are sometimes referred to as an *artificial society* (Gilbert and Conte, 1995; Epstein and Axtell, 1996).

## Beliefs and Collective Beliefs

### 3.1

I take a simple but sufficient view of what it means for an extended-agent to *believe p*. This is that *p* may be "decoded" from the agent's memory -- which implies that there are systematic processes by which suitably coded propositions are added to and deleted from and manipulated in the agent's memory and which, as viewed by an outside observer, are consistent with the working of its perceptual and effector processes. I also limit consideration to beliefs which are essentially descriptive, and ignore the matter of degrees of belief. For a much more developed notion of belief see Mack (1994).

### 3.2

By a *collective belief p* of a group of agents, I mean merely that the great majority of agents in the group in question have the belief *p*. This definition is about the simplest possible. There is no requirement, for example, for agents to believe that others hold the belief *p*, still less any notion of mutual belief or belief associated with a role. The collective beliefs (or *collective belief system*) of a group of agents are then those beliefs that occur widely amongst the agents. Again, this simple concept of collective belief, imprecise though it is, is sufficient for our purposes. For more developed theories of group belief see Tuomela (1992) and Wooldridge and Jennings (1998).

## The Impact of Belief upon Behaviour

### 3.3

An extended-agent's behaviour is determined by a combination of its beliefs and its rules. At first sight, it may seem somewhat counter-intuitive that to change an agent's behaviour one may change either its rules *or* its beliefs, but this is clearly so. If the beliefs in an extended-agent's memory are changed then, assuming a fixed rule set, the actions it performs will change. In human terms, I will behave differently when I awake in the morning if I believe it is Sunday rather than Monday. More dramatically you can perhaps make me "commit suicide" if you can make me believe that a glass of nitric acid is a glass of water -- or that the dead go to a "better world".

## Misbelief

### 3.4

The examples of the foregoing section raise the issue of the accuracy of beliefs. When an agent believes something which is (from the point of view of an external observer or experimenter) demonstrably NOT the case in its environment, I speak of the agent's *misbelief*. There is an analogous notion of *collective misbelief*. Of course, an agent may have beliefs whose soundness is difficult to determine. An agent may, for example, believe that there are two types of agent in its world, those that should be attacked if possible and those that should not. How does one assess the soundness of such a belief?

### 3.5

In general, sources of limited or mistaken belief in a multiple agent system may easily be identified. They include:

- Limited and faulty perception. An agent may be unable to perceive what is the case (e.g. because it is too far away) or may misperceive.
- Limited and faulty communication. One agent may fail to pass information to another, or may do so in such a way that errors are introduced.
- Beliefs becoming inaccurate with the passage of time. An agent may persist in a belief long after it has ceased to be accurate.
- Faulty generalization.

### 3.6

To the extent that these factors are at work within the agent community partial, inconsistent and errorful belief may occur and may well be the norm. Further, to the extent that there is *belief harmonization* between spatially neighboring agents, for example by agents communicating possibly erroneous beliefs one to another ("Have you heard, the Martians have landed?"), collective misbelief will also be the norm. Belief harmonization may be achieved in many ways. All that is required is some process which tends to bring the beliefs that neighbouring agents hold into agreement, irrespective of their soundness.

### 3.7

It is natural to suppose that collective misbelief must be detrimental to the society of agents in question. However, it is easy to see that this is not always the case. To take a human example, consider a group who (incorrectly) believe that a certain stream is holy so that water may not be taken from it. This may be to their detriment. But it may also be to their benefit if the stream is dangerously contaminated in some non-obvious way. Indeed it has been argued that collective misbelief may often be *necessary* for the survival of those who hold it. Thus the anthropologist

Roy Rappaport, discussing the importance of the Tsembaga people's *cognized model* (which loosely corresponds to what is here called "collective belief") remarks: "It can thus be argued that the cognized model is not only not likely to conform in all respects to the real world... but that it must not" (Rappaport, 1984, page 239). He is referring particularly, but not exclusively, to the way in which socially beneficial ritual truces between different sections of the population are reinforced by mutual fear of the anger of their (dead) ancestors should a truce be violated.

## Experimental Scenarios

### 4.1

Now I discuss experimental work intended to explore the genesis and impact of collective (mis)belief. A typical experiment involves the "evolution" of (the beliefs of) a population of agents in a specified environment, so that the emergent properties of the population may be related to its initial properties and to the properties of the environment. The evolutionary process itself is akin to the techniques of genetic algorithms and genetic programming, and the emphasis on the evolution of beliefs is similar to Dawkins' (1989) notion of evolving populations of self-propagating ideas or *memes* (to which I shall return later).

### 4.2

In Doran (1994) I reported the properties of an artificial society set "on a line" which was simple enough to be given a full mathematical description, and whose simulation properties demonstrated the potential impact of a certain type of collective misbelief which may be called *attribute error*. Attribute error is present where agents do not hold false beliefs about the existence of entities in their world, or the types of those entities, but are in error about their more detailed properties -- such attributes as colour or dimension. In the experiments reported it was shown, in brief, that in certain circumstances the population of agents would typically evolve to a steady state in which their collective beliefs about their own precise locations on the line were inaccurate, but had the effect of enhancing individual agent survival and hence increasing average population size. Here I report experiments with more complex computational systems which are no longer capable of simple mathematical description, but which nevertheless enable the investigation into collective misbelief to be taken further.

## The Scenario-3 Testbed

### 4.3

The experiments to be described used the software testbed SCENARIO-3 (written in the programming language C) which supports in simulation:

- a two dimensional spatial environment
- mobile agents and immobile resources
- (limited) agent perception of their surroundings
- agent internal representations of other agents and of resources (including memory limitation and forgetting)

- agents moving towards and harvesting resources for energy (in mutual competition)
- agent death by starvation or by aging
- (asexual) reproduction
- one agent killing another

The passing of time is simulated within the testbed as a sequence of "time units", within each of which events at different locations (e.g. the movements of agents) take place simultaneously.

#### **4.4**

Words such as "harvesting" and "killing", used to aid intuitive understanding, denote relatively simple events within the testbed. For example, "harvesting" is said to occur when an agent located at a resource reduces the energy level of the resource to zero, and increments its own internal energy store by a corresponding amount. Energy is used up by an agent (reducing its energy store) as an agent moves around in the testbed. "Killing" involves two agents meeting and one possibly becoming "dead" (and deleted from the world) and the killer acquiring the killee's energy store. One agent maintains a representation of another when it holds information (not necessarily accurate) about the other and about some of the other's characteristics.

#### **4.5**

In the SCENARIO-3 testbed, many important parameters are under the control of the experimenter. These include the perceptual range of the agents, their rate of movement, the range over which agents can communicate one with another, and their memory capacity. And many potentially complex sequences of sub-events, for example those which in principle would determine each particular agent reproduction, are bypassed by an appeal to suitably chosen probability distributions.

### **Experiments I: The Impact of Pseudo-Agents**

#### **4.6**

The set of experiments I shall now describe focuses on awareness by agents of other agents and of self-renewing energy resources, and on action to harvest them. It particularly concerns the impact of agents (mis)believing in non-existent agents -- what may be called *existence errors* (in contrast to the *attribute errors* of the work briefly outlined earlier). The experimental scenario includes simple forms of agent death and reproduction.

### **The First Experimental Scenario in Detail**

#### **4.7**

$N$  agents and  $M$  resources are located on a plane. They are initially randomly distributed in the unit square with corner points  $(0,0)$ ,  $(0,1)$ ,  $(1,1)$ ,  $(1,0)$  in rectangular coordinates, but agents may leave this square as the experiment proceeds. Resources have fixed locations. Each agent has an energy level, which declines with time and which is restored by resource consumption.

Resources renew periodically, which implies that there is a maximum carrying capacity for the environment

#### 4.8

If an agent's energy level declines to zero, the agent dies and is removed. Further, if an agent enters a particular "fatal" zone (the circle centered at [0.5,0.5] and with radius 0.25) it immediately dies. In each time unit, there is a small probability that any given agent may reproduce, that is, create an exact copy of itself located at an adjacent point).

#### 4.9

The actions taken by an agent in each time unit, in order, are the following:

- Harvesting and consumption of immediately adjacent resources
- A small random movement
- Movement towards the nearest resource *to which it believes itself the nearest agent*. This movement will only occur *if there exists such a resource*. The underlying idea is that there is no point in an agent trying to harvest a resource when some other agent is better placed to harvest it.

#### 4.10

In these experiments, agents are aware of the current locations of all resources and agents in the environment. Additionally, they may (mis)believe in the existence (essentially, the location) of a small number of agents that do *not*, in fact, exist in the environment. I refer to these as *pseudo-agents*. Pseudo-agents do not move and cannot, of course, consume resources. It follows that an agent may pass up a good opportunity to harvest a resource from belief in a pseudo-agent.

#### 4.11

Initially, each agent independently (mis)believes in a small set of randomly generated pseudo-agents. Different agents believe in different pseudo-agents. Such beliefs are typically handed on from 'parent' to 'child', with a possibility of random variation when agents' energy levels are low. In addition, and importantly, the testbed may be set so that agents tend to harmonize their beliefs with their nearest spatial neighbours. Specifically, agents are scanned in turn in each time unit, and each takes over the beliefs (about the locations of pseudo-agents) of its nearest neighbour where they differ from its own. If the circumstances are right, therefore, a particular pattern of (mis)beliefs may spread through the population.

#### 4.12

Each agent decides its movements by reference to all its current beliefs, including beliefs about pseudo-agents, not by reference merely to the actual state of the environment. The agents in no way distinguish between real and pseudo-agents when deciding their actions.

## **Results and Discussion**

### **4.13**

In systematic experiments it has been found that the agent society enters semi-stable states in which agent survivability is demonstrably enhanced by collective misbelief in small sets of pseudo-agents. Typically, most of the agents in the society come to share (mis)beliefs in a small number of pseudo-agents in the fatal zone. This causes these agents to seek resources elsewhere, away from the fatal zone and therefore more safely. The effect is to maintain a substantially (about 50%) greater agent population than would otherwise be the case (as determined in control trials). In evolutionary terms, what happens is that groups of agents with particular collective beliefs "compete", and those groups with beliefs which help them to survive tend to increase their numbers at the expense of the remainder.

### **4.14**

Simple though these experiments are, they do illustrate the socially beneficial impact of what is perhaps the most straightforward type of misbelief that can occur in a multi-agent system -- existence errors.

## **Experiments II: "Cults"**

### **4.15**

The second set of experiments involves a significantly different type of misbelief, *category error*, where an agent assigns an entity to the wrong type category and reacts to it accordingly. Thus in this particular experimental scenario, agents can come to believe that a non-agent, that does indeed exist, is an agent. There is also a more complex interaction between misbelief and behaviour.

### **4.16**

These experiments again use the SCENARIO-3 testbed, but now using that feature of the testbed which enables agents, in certain circumstances, to "kill" one another. Also new are a particular type of agreement, a *friendship* and, crucially, the notion of a *resource agent*.

## **Friendship**

### **4.17**

An agent may decide that another agent, which it happens to meet, is its *friend*. The chance of such an outcome to a meeting is determined by the experimenter. If an agent X does come to "think of" an agent Y as a friend, then *X passes information about resources to Y (e.g. their locations) whenever Y is within message passing range AND X never attempts to kill Y*. Note that friendship is not necessarily symmetric. X may treat Y as a friend whilst Y does *not* so treat X.

#### **4.18**

There is another important aspect of friendship. *An agent X will not attempt to kill an agent Y (even though X does not view Y as a friend) if it is the case that X and Y have a believed friend, say agent Z, in common.* This latter aspect of friendship is crucial to what follows.

### **Resource Agents and Cults**

#### **4.19**

The testbed may be set so that from time to time an agent "agentifies" a resource, that is, wrongly comes to think of a resource as if it were an agent. We may call a pseudo-agent like this a *resource agent*. In its thinking an agent does not distinguish between resource agents and real agents, so an agent may even regard a resource agent as a friend and act towards it accordingly. Of course, messages sent to a resource agent go nowhere.

#### **4.20**

Once an agent forms a representation of a resource agent, that representation may be passed to other agents by inter-agent communication. It may also be passed from one agent to its offspring. If the circumstances are right, therefore, the representation may spread. When a set of agents all come to believe in the same resource agent, and that they have that resource agent as a friend, we may call that a *cult*. The resource agent in question may be called the cult *head*. It follows immediately from the properties of the friendship relation given earlier that the members of a cult will not kill one another.

## **Results and Discussion**

#### **4.21**

What is found experimentally is that even a very low-frequency possibility of agents coming to believe in resource agents can regularly lead to the formation of large and enduring cults and that, all other things being equal, killing in the society is then greatly reduced and the average population of the society over time increased.

#### **4.22**

To give the reader a feel for these experimental trials and what happens within them, there follows a summary account of key events in one typical trial (and see Figure 1). The agent and resource identifiers are exactly as they appear in the testbed and its output:  
Resource number 11 was initially 'conceived' as a friendly agent, 110000000rrr, by agent number 248 in time unit 251.

At time 268 this resource agent had just one 'host' agent in a population of only 3 agents - not agent 248, which died before time 268, but agent 8000254. A cult around 110000000rrr built up

thereafter (comprising descendants of 8000254), typically containing about 30 member agents, and lasted for hundreds of time units.

Initially the cult was also around a 'dead' real agent 20000267 which itself had lived for only one time unit, but was also regarded as a friend by agent 8000254. However, memory of 20000267 was lost in time unit 287.

#### 4.23

Note the appearance of a "dead" agent in the account. The potential of dead agents as cult heads was not anticipated, though obvious enough in hindsight. In fact, a dead agent is not as effective as a resource agent as a cult head because the latter, unlike the former, can be "seen" and awareness of it thereby refreshed. What matters most for the formation and indefinite survival of the cult is the effective "immortality" of its head.

#### 4.24

It should be made clear that these experimentally observed phenomena are not entirely straightforward to obtain. As was indicated earlier, the agent society and world embody many parameters (e.g. dimensions, the rate of renewal of resources, agent perception range, agent rate of movement, agent life span, agent memory span, the probability of a new friendship in any time unit, the probability of an agent seeing a resource as a resource agent in any time unit), and different combinations of settings for these parameters often lead to very different outcomes. A heuristic is that the chance of cults is greater the greater the contrast between the "immortality" of the potential cult head and the "mortality" of the potential cult members. Thus relatively short agent life and memory spans make the formation of cults more likely.

#### 4.25

Current experiments are focussing on the dynamics of multiple competing cults each exploiting several resources, and with each agent possibly a member of several cults. These experiments raise such issues as competition between cults, and also the formation of hierarchies topped by a cult head.

### Future Work

#### 4.26

The two sets of experiments reported have demonstrated that it is possible to evolve collective misbelief systems in an artificial society which interact in such a way with prior rules governing agent behaviour that the society as a whole benefits. Some of the relevant conditions have been established. But much more experimental work needs to be done to approach a full understanding. Thus the connection between different patterns of misbelief and their impact upon different types of environment needs further examination. Is it the case that in more realistic scenarios the possible benefits of collective misbelief tend to be outweighed by disadvantages?

#### **4.27**

For any particular collective belief system, two different questions may be asked: how effective is it in context? And how may it come into existence? These two questions are not as well separated as they should be in the experiments reported here. And the answer to the latter question is surely more than the *ad hoc* random processes incorporated in Scenario-3 (see next section).

#### **4.28**

Finally, what may be called *biased* belief systems merit investigation. These are belief systems whose effect is to favour one subset of the agents in the system against another. It might be, for example, that agents with certain characteristics are rewarded by the belief system in a way that would not be so were another belief system in place.

### **A Computational Theory of Ideologies?**

#### **5.1**

The foregoing experiments and associated discussion related to agents in computer based artificial societies. But any discussion of collective belief and disbelief in a system of non-trivial agents, and its functional significance, prompts consideration of a connection with the "social construction of reality" and with "ideologies" in the sense that they are studied in social science. By "ideologies" social scientists typically mean systems of belief held in a society which enable and validate action (especially group motivated action) and/or erroneous belief systems which enable or support domination of one human group or class by another (e.g. Thompson, 1984). It seems reasonable that insofar as ideologies may be regarded as observable features of human society and may be objectively studied (which is admittedly a matter of some controversy), then studies of collective belief and disbelief in DAI agent systems must be one way to further our understanding of their dynamics.

#### **5.2**

Approaching ideologies from an AI and artificial societies standpoint, with a consequent emphasis on the connection between micro and macro social phenomena, makes it natural to suggest that the origins of ideological beliefs lie in the inherent cognitive ability of individuals to manipulate beliefs, by adaptation, generalisation or processes of "generate and test", in a manner which is heuristic rather than fully reliable. Manipulation of kin beliefs and concepts may be especially significant (compare Read, 1995), a possibility explored in the context of the EOS project by Mayers (1995). It also seems likely that a significant impact of ideologies is to coordinate the actions of people and to limit, in effect, their individual autonomy by creating "social norms" so enhancing survival. This is the view argued by Rappaport, 1971. Conte and Castelfranchi, 1995, and Ephrati et al., 1995, have recently reported computer based studies from a similar standpoint.

### **5.3**

In all this work there is a close relationship with *memes* (Dawkins, 1989), that is, ideas or beliefs which propagate and evolve in the *noosphere*, the collective mental space of a population. Computer based experimental work by Bura (1994) and Hales (1997) has studied in detail the dynamics and functional impact of memes and meta-memes in simple artificial worlds. Further, the work reported here should also be compared with that of Reynolds (1995). The collective beliefs studied by Reynolds, however, are differently represented, and are primarily directed to effective problem solving.

## **Conclusions**

### **6.1**

It is clear that belief and collective misbelief are inevitably at the heart of the behaviour of natural and artificial societies. Further, collective misbelief is not necessarily something to be avoided -- it may be functional if matched to the agents' environment. This observation is not in itself new. What is new and demonstrated here is our emerging ability to address particular questions and conjectures about such issues by way of precise and targeted computer-based experiments in which functional collective misbelief (involving attribute, existence and category errors) is evolved in an agent society. This has important implications for the study of ideologies in human societies.

## **Acknowledgements**

Much of the experimental work described here has used supercomputer facilities made available at the Centro di Calcolo Interuniversitario dell'Italia Nord-Orientale, under the EC CINECA ICARUS programme.

## **References**

- BURA S (1994) MINIMEME: of Life and Death in the Noosphere. In *From Animals to Animats 3*, Proceedings of the Third International Conference on Simulation of Adaptive Behavior (eds D Cliff, P Husbands, J-A Meyer & S W Wilson) MIT Press: Cambridge Mass. & London, England. pp 479-486.
- CONTE R. and Castelfranchi C. (1995). Understanding the Functioning of Norms in Social Groups through Simulation. In *Artificial Societies* (eds N Gilbert and R Conte). pp. 252-267. UCL Press: London.
- DAWKINS R (1989) *The Selfish Gene* Oxford University Press: Oxford & New York (new edition).
- DORAN J (1994) Modelling Collective Belief and Misbelief. In *AI and Cognitive Science '94* (eds M Keane et al). Dublin University Press. pp 89-102.

EPHRATI E, Pollack M E and Ur S (1995) Deriving Multi-Agent Coordination through Filtering Strategies, *Proceedings of the Fourteenth International Joint Conference on Artificial Intelligence* (IJCAI-95), Montreal, August 20-25th 1995. pp. 679-685.

EPSTEIN J M and Axtell R L (1996) *Growing Artificial Societies: Social Science from the Bottom Up*. The Brookings Institution Press: Washington D.C. & The MIT Press: Cambridge, Mass.

GILBERT N and Doran J eds. (1994) *Simulating Societies: the Computer Simulation of Social Phenomena* UCL Press: London.

GILBERT N and Conte R eds (1995) *Artificial Societies: the Computer Simulation of Social Life*. UCL Press: London.

HALES D A P (1997) Modelling Meta-Memes. In *Simulating Social Phenomena* (eds R Conte, R Hegselman, P Terna). Lecture Notes in Economics and Mathematical Systems, 456. pp. 365-384. Springer: Berlin.

MACK D (1994) A New Formal Model of Belief. *Proceedings of the Eleventh European Conference on AI* (ECAI-94). John Wiley.

MAYERS S D (1995) Modelling the Emergence of Social Complexity using Distributed Artificial Intelligence, Masters dissertation, Department of Computer Science, University of Essex, Colchester, UK.

O'HARE G and Jennings N eds. (1996) *Foundations of Distributed AI* Wiley Inter-Science.

RAPPAPORT R A (1971) The Sacred in Human Evolution. *Annual Review of Ecology and Systematics* 2:23-44.

RAPPAPORT R A (1984) *Pigs for the Ancestors* (2nd edition) Yale University Press. New Haven and London.

READ D W (1995) Kin Based Demographic Simulation of Societal Processes. *Journal of Artificial Societies and Social Simulation* vol. 1, no. 1,  
<http://www.soc.surrey.ac.uk/JASSS/1/1/1.html>

REYNOLDS R G (1995). Simulating the Development of Social Complexity in the Valley of Oaxaca using Cultural Algorithms. Pre-Proceedings of *Simulating Societies '95 Symposium (SIMSOC'95)*. Boca-Raton, Florida, September 1995.

RUSSELL S & Norvig P (1995) *Artificial Intelligence: a Modern Approach*. Prentice- Hall.

THOMPSON J B (1984) *Studies in the Theory of Ideology*. Polity Press: Cambridge.

TUOMELA R (1992) Group Beliefs. *Synthese* 91:285-318.

WOOLDRIDGE M J and Jennings N R eds (1995) *Intelligent Agents* LNAI 890, Springer: Berlin.

WOOLDRIDGE M J and Jennings N R (1998) Formalizing the Cooperative Problem Solving Process. *Readings in Agents*. (eds M N Huhns and M P Singh) pp. 430-440. Morgan Kaufmann: San Francisco, California.

WOOLDRIDGE M J, Muller J P, Tambe M eds (1996) *Intelligent Agents II*, LNAI 1037, Springer: Berlin.

# **THE REVEALED POVERTY OF THE GENE-MEME ANALOGY: WHY MEMETICS PER SE HAS FAILED TO PRODUCE SUBSTANTIVE RESULTS**

By  
**Bruce Edmonds**  
Centre for Policy Modeling  
Manchester Metropolitan University

Over two years ago in this journal I challenged the memetics community to meet three challenges (Edmonds 2002). My prediction was that if these were not substantially met that memetics as a distinct approach would not survive. Despite some attempts to do this these challenges were not met, at least not within the terms I had proposed. Now *JoM-EMIT* in its current form has ceased due to a lack of quality submissions. Whilst I do not claim that the former caused the latter, I do claim that the failure to answer those challenges was indicative of the poverty of the memetics project resulting in a lack of demonstrable progress which, in turn, has meant that it has failed to interest other academics.

Here I distinguish between what might be called the "broad" and the "narrow" approaches to memetics. The former, broad, approach involves modeling communication or other social phenomena using approaches which are evolutionary in structure. Work within this approach is often done without appealing to "memes" or "memetics" since it can be easily accommodated within other frameworks. In other words, it does not require an analogy with genetics. The later, narrow, sense involves a closer analogy between genes and memes – not necessarily 100% direct, but sufficiently direct so as to justify the epithet of "memetics". What has failed is the narrow approach – that of memetics. Work continues within the broad approach, albeit under other names, and in other journals.

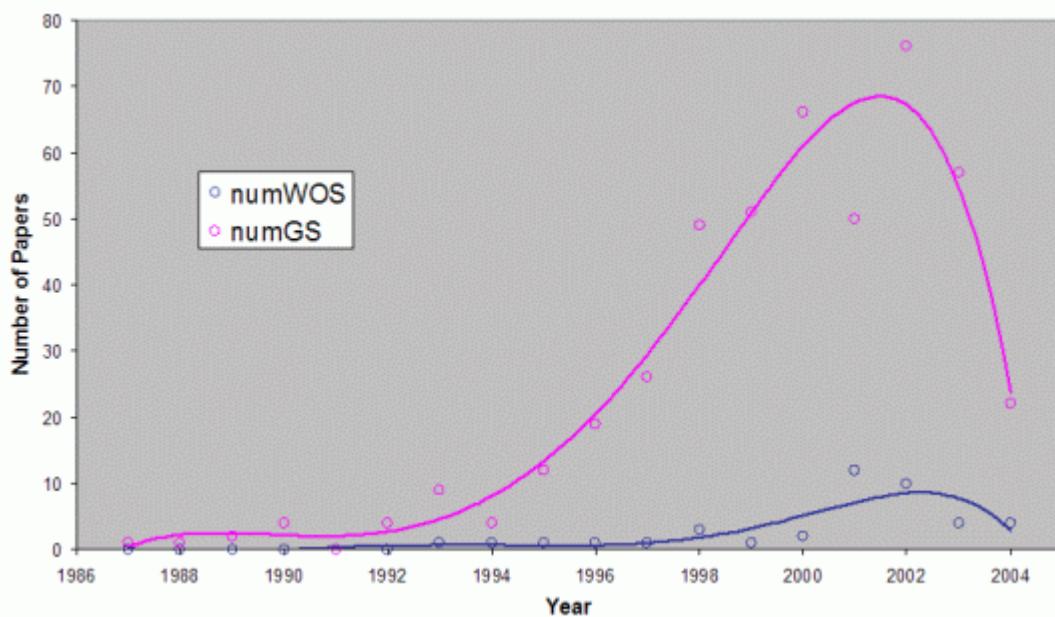
I claim that the underlying reason memetics has failed is that it has not provided any extra explanatory or predictive power beyond that available without the gene-meme analogy. Thus whilst the idea of memes has retained its attractiveness for some in terms of a framework for thinking about phenomena, it has not provided any "added value" in terms of providing new understanding of phenomena. The fact that some who wear the theoretical spectacles (Kuhn 1969) of memetics insist of redescribing a host of phenomena in these terms despite the lack of substantive results merely confirms other academics' opinion of the approach. The ability to think of some phenomena in a particular way (or describe it using a certain framework), does not mean that the phenomena has those properties in any significant sense.

Academics who seek to study memetics in serious ways have suffered in the respect that they are often confused with those on the penumbra for whom memetics is a fad. However, this mistake is grounded in an element of truth. The study of memetics has been characterized by theoretical discussion of extreme abstraction and over ambition. Thus for example, before any evidence is available or detailed causal models constructed, attempts have been made to "explain" some immensely complex phenomena such as religion in general [note 1] or consciousness. This sort of discussion shifts any study of memetics from the realm of science to that of a philosophy and,

on the whole, this philosophy has adopted the subsumption tactic (Hull 2001), seeking to generalize explanation rather than been productive of essentially new insights.

In the broader sense it would be extremely surprising if there were no social processes with evolutionary aspects. However, the intricacy of social phenomena means that understanding the effects of different mechanisms in systems of multiply interacting actors requires a finer tool than that of discursive analogy. Thus unraveling some of the conditions for observed social processes (which may have evolutionary aspects) requires such tools as agent-based computational simulation which can track long-chains of intricate interactions. This approach also allows for a wide variety of different mechanisms and processes to be explored and not merely those amenable to description via biologically-rooted analogies. Of course, there is a successful community of social simulators who study, among other things, evolutionary models of information transmission [note 2]. Similarly there is work in computer science, applying evolutionary ideas to computational processes and work in theoretical biology studying non-genetic evolutionary processes. Thus this wider work will continue as subsets of other projects, but not under the discredited label of memetics.

To illustrate how the memetics bandwagon may have peaked I used ISI's large citation index (<http://www.isi.org>) and Google Scholar (<http://scholar.google.com>) to estimate the number of papers that mention the word memetics. I picked "memetics" rather than "meme" (which has a longer history) because of the existence of the common French word "meme" (as in le meme chose). I subtracted those papers about "memetic algorithms" because these are not about memes in any meaningful way [note 3].



**Figure 1.** Number of papers mentioning "memetic\*" (but not "memetic algorithm\*") each year according to Google Scholar (numGS, pink circles) and on the ISI's citation index (numWOS, blue circles). Lines are 6<sup>th</sup> degree fitted polynomial trend lines of the respected series

The fact is that the closer work has been to the core of memetics, the less successful it has been. The central core, the meme-gene analogy, has not been a wellspring of models and studies which have provided "explanatory leverage" upon observed phenomena. Rather, it has been a short-lived fad whose effect has been to obscure more than it has been to enlighten. I am afraid that memetics, as an identifiable discipline, will not be widely missed.

## Notes

Here I want to distinguish woolly explanations of religion in general from the more careful study of the phylogeny of particular institutions such as in (Lord and Price 2001).

The success of the Journal of Artificial Societies and Social Simulation (<http://jasss.soc.surrey.ac.uk>) contrasts markedly with that of JoM-EMIT.

"Memetic Algorithms" are Genetic algorithms but where hill-climbing is used to locally optimize solutions. No spread of solutions is involved (unlike a few other evolutionary computation approaches which I have included).

## References

Edmonds, B. (2002). Three Challenges for the Survival of Memetics. *Journal of Memetics - Evolutionary Models of Information Transmission*, 6.  
[<http://jom-emit.cfpm.org/2002/vol6/edmonds\\_b\\_letter.html>](http://jom-emit.cfpm.org/2002/vol6/edmonds_b_letter.html)

Hull, D. (2001) In Search of Epistemological Warrent. In Cecilia Heyes and David Hull (eds) Selection Theory and Social Construction: the evolutionary naturalistic epistemology of Donald T. Campbell. State University of New York Press, Albany, NY, 2001, 155-167.

Kuhn, T. (1969) *The Structure of Scientific Revolutions*. Chicago: University of Chicago Press.

Lord, A. and Price, I (2001). Reconstruction of organizational phylogeny from memetic similarity analysis: Proof of feasibility. *Journal of Memetics - Evolutionary Models of Information Transmission*, 5. [<http://jom-emit.cfpm.org/2001/vol5/lord\\_a&price\\_i.html>](http://jom-emit.cfpm.org/2001/vol5/lord_a&price_i.html)

# **THE EVOLUTION OF LANGUAGE AND SCIENCE STUDIED BY MEANS OF BIOLOGICAL CONCEPTS**

By

**Börje Ekstig**

Department of Teacher Training  
Uppsala University, Uppsala Sweden  
[borje.ekstig@ilu.uu.se](mailto:borje.ekstig@ilu.uu.se)

## **Abstract**

This study examines certain mechanisms underlying the evolution of language and science — including mathematics — using concepts developed in the field of biological evolution. Developmental processes are particularly emphasized. Analysis of developmental processes, such as human embryonic development, children's verbal development, and adolescents' scientific conceptual development reveals the unifying principle referred to as "condensation" — the successive shortening of developmental stages. The mechanism of condensation is coupled to the rate of evolutionary change.

The analysis examines the applicability of the concept of the meme. Regarding the evolution of language, we suggest a cooperative combination of genetic and memetic replication; while early on in the evolution of science only memetic replication is envisaged.

**Key words:** Evolution, development, memes, cultural evolution, language, science, mathematics.

## **1. Introduction**

This paper examines the mechanisms underlying the evolution of language and science, applying concepts developed in the field of biological evolution. Common to these various manifestations of evolutionary process is developmental growth, as observed in the embryo, child, and adolescent. The significance of the developmental process in its various guises forms the main subject of my study. Such a broad approach is highly multi-disciplinary, and in this introduction I will comment on some relevant fields of research.

In an early work, Stephen J. Gould (1977) reviews the widespread notion of a coupling between the developmental and evolutionary processes. More recently, our rapidly growing insight into genetics has lent this controversial field renewed respect. One such new field of research focuses on the developmental genetic machinery that underlies embryological phenotypes (Arthur 2002). It is recognized that evolutionary change occurs not by direct transformation of adult ancestors into adult descendants, but rather when developmental processes produce the features of each generation in an evolving lineage (Raff 2000). Wagner et al. (2000) have examined another theme, the role of development in the origin of evolutionary innovations. In this study I will build on the same principal insight of the central role of development, but restrict my discussion to the lineage of humans.

Concerning cultural evolution, there is an overwhelmingly abundant literature on the relationship between cultural and biological evolution within which one can discern several different approaches. One such approach is socio-biology. As Edward O. Wilson has suggested, certain cultural norms may survive and reproduce better than other competing norms, causing culture to evolve along a track parallel to and usually much faster than genetic evolution. Wilson, however, contends that the connection between genes and culture is never completely broken and discusses this connection in terms of gene-culture co-evolution (Lumsden and Wilson 1981), a concept that supports the general notion of a close relationship between cultural and biological evolution.

Another approach that examines the connection between cultural and biological evolution is memetics—an approach likely familiar to readers of this journal. Memeticists attempt to explain cultural evolution without reference to any benefit for genes. The Darwinian principle of natural selection is expanded, and is regarded as working not only on genes but also on memes. Works by Dawkins (1976), Dennett (1995), and Blackmore (1999, 2000) introduce and develop the concept of memes. These authors attempt to grant memes an analogous role in cultural evolution to that played by genes in organic evolution. I will analyze this idea at length in this investigation.

In studying children's psychological development, psychologists have attempted to apply concepts from evolutionary theory. Advocates of this influential field of research, called evolutionary psychology, emphasize that understanding the principles governing evolution is indispensable in attempting to understand human nature (Tooby and Cosmides 1992, p. 50). Such a general statement may stand as a guiding principle for the present investigation. However, if I have understood evolutionary psychologists correctly, they also seem to imply that all behavior ultimately comes back to genetic advantage. Such a view is questioned, for instance by Dennett and Blackmore, and I will demonstrate that genes are insufficient to explain many characteristics of cultural evolution.

Language is an important constituent of culture, and there is a tremendously rich literature dealing with the enigmatic question of the origin and evolution of language. There has recently been exciting debate as to whether or not language has evolved by means of a principle analogous to natural selection; Dennett provides a survey of this debate (1995, pp. 384-400). I suggest that any explication of the evolution of language, while incorporating elements from natural selection, should have recourse to additional mechanisms as well.

There is a widespread notion that the evolution of science has much in common with biological evolution. Thus Blute (2003), for instance, discusses a new field of research called evolutionary ecology. I will refer to such analogies, and in particular analyze their implications for children's and adolescents' learning of science.

The referenced literature testifies to the many notions of connections between cultural and biological evolution. These connections are elaborated in the analysis I will propose in this paper. It builds on my previously introduced model (Ekstig 1985, 1994), and here I will focus on the evolution of language and science. First I will recapitulate the main features of the model.

## 2. The pattern

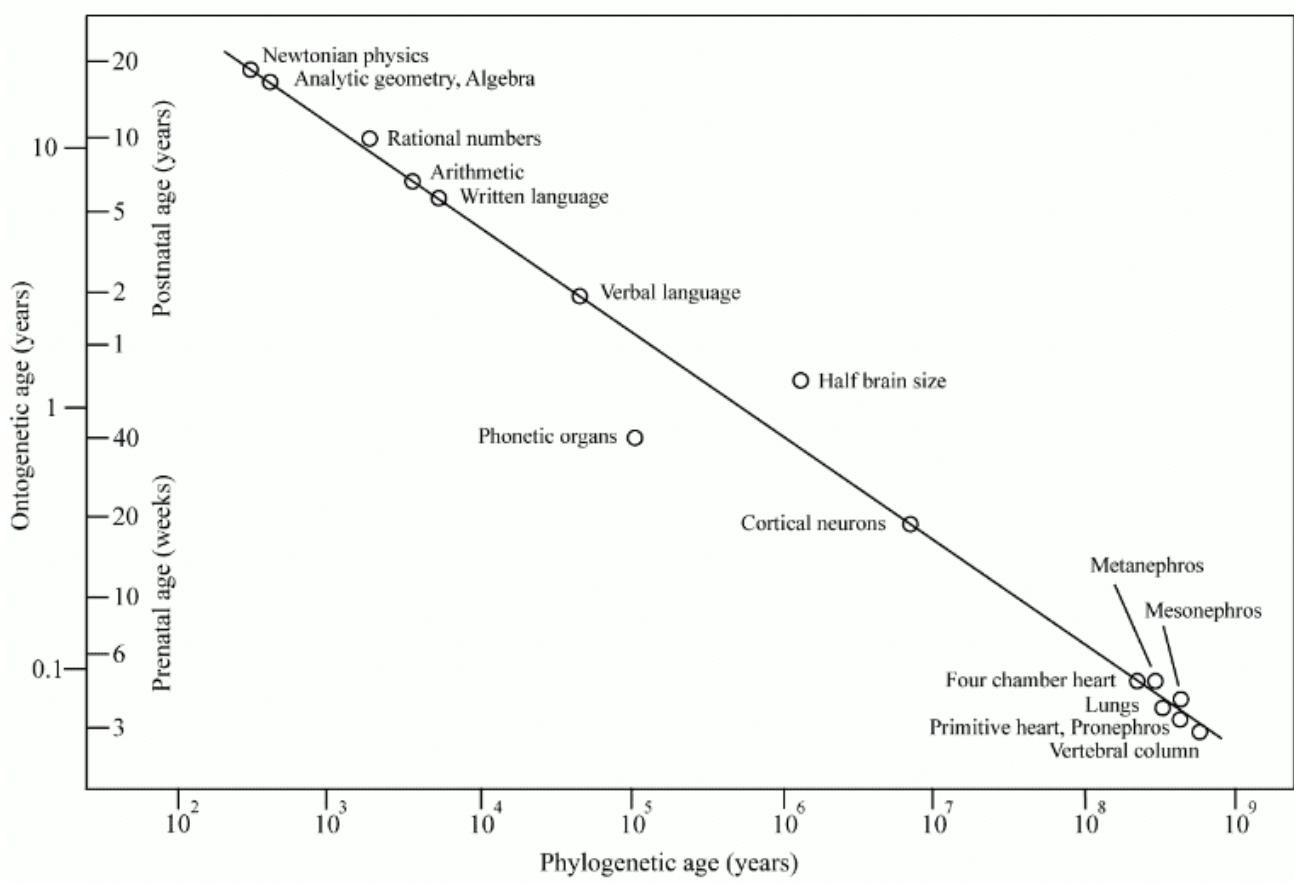
My study assigns the developmental processes of individual human beings a central role in the analysis of biological and cultural evolution. The role of development in biological evolution stems from the fact that genes carry and transmit instructions for self-replication and for development, not for evolution, and because organic evolution proceeds as a result of continuous modifications in the developmental program. The continuity of this process means that many vestiges of early evolutionary stages are stored in the genes and manifested as phenotypic developmental stages in the present-day embryo. Or, as Darwin long since observed, “the embryo comes to be left as a sort of picture, preserved by nature, of the former and less modified condition of the species” (1859, p. 338). Such common stages or traits, called parallels, are thoroughly examined by Gould (1977).

Likewise, the continuity of the transmission of cultural traits over generations means that stages from our cultural history are now discernible as developmental stages in the present-day child and adolescent. Such parallels in the cultural realm have long been observed and discussed, and are also surveyed in Gould’s review. In the field of mathematics and physics, Piaget and Garcia (1983) have reported a thorough study in which they find striking coincidences.

I have earlier (1994) analyzed the relationship between individual development and the evolutionary process, taking as a point of departure the pattern that appears when the developmental age at which each trait appears is compared to the period of time that has elapsed since it first arrived on the scene in evolutionary history. The essential features of the model appear in Figure 1. The pattern encompasses both the organic and mental aspects of human development, and indicates an intimate relationship between biological and cultural evolution.

As is immediately obvious, the diagram includes very few of the traits found in a growing organism. The principle behind the selection of the traits forming the pattern, organic as well as cultural, is that these traits appear in an inflexible order in the individual’s developmental program as well as in the evolutionary process. I suggest that such traits be denoted cumulative traits.

Organic evolution proceeds through the successive addition of novel developmental traits, thus forming the evolutionary pathway of the species. These novel traits predominantly manifest themselves towards the end of individual development, i.e., just before sexual maturity, because, by and large, the progeny’s genetic constitution is not influenced by selection exerted on the parent after conception. These novel traits are called terminal additions, and since they manifest themselves before sexual maturation they tend to lengthen the development process and therefore also generation length. However, there is a counteracting process as well.



**Figure 1.** Diagram of ontogeny versus phylogeny in the lineage of humans

The ontogenetic age is measured from the moment of fertilization. The ordinate also gives the individual's postnatal and fetal ages. Both scales are logarithmic. The diagram is a revised version of a diagram appearing in Ekstig (1985), and the reader is referred to that reference for comments. It should be mentioned that the uncertainty regarding the Phonetic organs point is very great. If the phylogenetic age is assumed to be one million years, which is not unrealistic, the point will be situated on the line.

### 3. Condensation

Lineages with shorter generation lengths increase their relative number in the gene pool. Hence there is a ubiquitous selection pressure favoring shorter generations, as analyzed in life-history theory. Stearns (1992) shows that for organisms that mature earlier, generation length is shorter and populations grow more rapidly. There is thus a trade-off between terminal additions and acceleration of maturation. Delayed maturation caused by terminal addition of a novel trait is selectively favorable only if the benefit of the novel trait is greater than the loss in fecundity due to the resulting prolonged generation length.

Shortening of generation length may be realized by eliminating unnecessary traits or by gradually shortening early developmental stages. Such gradual shortening of developmental stages is called condensation in the literature on evolution (the term “condensation” is also used

in psychology, but with a different meaning).

Condensation and terminal addition act simultaneously and independently. Thus even if generation length increases—the most common trend—condensation may simultaneously have been at work. This implies that it is difficult to separate out and individually observe or measure the two processes. However, in the lineage of humans, condensation is indirectly revealed even by superficial analysis of the diagram in Figure 1. With time, all points in the diagram are displaced to the right, but due to the logarithmic form of the time scale, at unequal distances in the diagram. Therefore, linearity would not prevail unless the points were simultaneously displaced downwards at an appropriate rate. This reasoning rests on the assumption that the pattern will persist, because it is very improbable that it should appear just in our own time. I presume that the displacement of the traits towards earlier ages is accomplished by condensation, a conclusion that, in the case of biological evolution, is made plausible by the general selection pressure discussed above. The model also predicts the action of condensation in the cultural realm, and this subject will be explicitly discussed later in this paper.

One may conclude that the condensation of an evolutionarily old trait, appearing early in development, is less than that of a more recent trait appearing later in development. This seems reasonable, since one may assume that condensation of a specific trait is harder to accomplish the more condensed the trait has already become. Actually, mathematical analysis of the pattern revealed in Figure 1 (Ekstig 1985) indicates that the rate of condensation for a particular trait is inversely proportional to the age of the trait. For instance, the condensation of the development of the embryonic heart is 3% over 10 million years, while Newtonian science is condensed at a rate of 5% per decade. Thus, according to the present model, condensation of the manifestations of memes is about one million times faster than that of genes. This does not imply a claim to measure the rate of the evolutionary process; it only indicates the relative rates of change of biological and cultural evolution.

#### 4. Parallels in cultural evolution

This study will attempt to apply concepts that have emerged from a biological interpretation of the pattern in Figure 1 to the evolution of human culture. Therefore, my analysis of cultural evolution will bring the child's mental development to the fore and examine how successive changes in the child's developmental program are initiated by different modes of selection pressures and how they add up to evolutionary changes.

A conspicuous feature of the pattern in Figure 1 is that it includes the realm of cultural evolution, forming a continuous extension of the trend established by biological evolution. Indeed, just because organic change happens so slowly, compared to cultural change, as to be imperceptible over human time scales, it does not mean that biological evolution does not proceed simultaneously with cultural evolution. Cultural evolution comprises the cumulative development of many cultural manifestations involved in complicated interactions. Nevertheless, a few of the multifarious traits, namely cumulative traits associated with language and science, form the pattern. Other well-known cultural manifestations, such as mythologies, religions, and political ideologies, are not relevant to this discussion since they are not cumulative.

The selected traits exhibit parallels between their manifestation in the individual developmental process and in cultural history. This point is crucial to present model. Consider the ability to perform arithmetic calculations using whole numbers. This faculty emerged several thousand years ago and was later developed into the ability to handle the more sophisticated rational numbers. For an individual human it was and still is necessary first to be acquainted with whole numbers. As a matter of fact, it is generally impossible to omit a former stage of mathematical knowledge before acquiring a latter one, nor is it possible to study math issues in reverse order. This is characteristic of cumulative cultural traits. Therefore, the individual acquisition of these stages follows the same order in which the corresponding stages are observed to appear in the historic record. This cumulative characteristic is not related to psychology, education, or any other external factors, but is a logical necessity inherent in many mathematical concepts. It is only in cultural features possessing cumulative developmental stages, I repeat, that one can expect to find such parallels.

Not all stages of mathematical development, however, are strictly cumulative in the present sense. Thus negative numbers can be acquired somewhat earlier by modern youths than could be expected from their first appearance in history. This illustrates the fact that the points in the diagram are not arbitrary, but are selected according to the criterion of being cumulative.

During the course of cultural evolution new cumulative cultural manifestations are added to the individual's developmental program analogously to the way terminal additions are added in biological evolution. Thus the time needed for a child's mental growth has increased, and this can be broadly verified by referring to the evolutionary record. The infants of chimpanzees, to whom our ancestors are supposed to be related, reach adulthood faster than human children do. Relying on the continuity of the evolutionary process, this indicates that human childhood — viewed from a long-term perspective — has grown longer.

Simultaneously, according to our present approach, condensation has been acting on many cultural traits acquired before maturation, resulting in a continuous shortening of cumulative cultural stages. This hypothesis will be discussed with specific reference to the evolution of language and of science.

## 5. The evolution of language

Noam Chomsky is reluctant to accept a Darwinian interpretation of the evolution of language, and instead suggests that an inherited universal grammar is the crucial, operative concept. Gould concurs, and further contends that "the universals of language are so different from anything else in nature..., that origin rather than as a simple advance in continuity from ancestral grunts and gestures, seems indicated" (Gould 1989, p. 14). Dennett (1995, pp. 384-400), however, vehemently rejects such non-Darwinian explanations. In considering language as an instinct, Steven Pinker tentatively approaches natural selection in suggesting that if language is like other instincts, it presumably evolved by natural selection (Pinker 1994, pp. 354–355).

Terence Deacon (1997) has articulated a more comprehensive Darwinian framework in his analysis of the evolution of language; I will take this as a point of departure for my discussion. Deacon accentuates the role of development by stating that children are the vehicle by which a

language gets reproduced (*ibid.*, p. 109), and that “the structure of a language is under intense selection because in its reproduction from generation to generation, it must pass through a narrow bottleneck: children’s minds. … Language operations that can be learned quickly and easily by children will tend to get passed on to the next generation more effectively and more intact than those that are difficult to learn. So, languages should change through history in ways that tend to conform to children’s expectations; those that employ a more kid-friendly logic should come to outnumber and replace those that don’t” (*ibid.*, p. 110).

Deacon indicates the existence of a mutual adaptation between language and humans, children in particular, in declaring that, “we should not be surprised to find complex human adaptations to language on the one hand, whose purpose is to ensure that language is successfully replicated and passed from host to host, and language adaptations to children on the other, whose purpose is to make languages particularly ‘infective’ as early as possible in human development” (*ibid.*, p. 113). He thus arrives at a notion of language as an autonomous process: “The evolution of symbolic communication … created a mode of extrabiological inheritance with a particularly powerful and complex character, and with a sort of autonomous life of its own” (*ibid.*, p. 409).

If we assume that such an autonomous language has evolved as a result of some kind of selection process, we must make intelligible the very basis of such a selection principle. To this end we may take hold of Susan Blackmore’s suggestion: “I have assumed that people will both preferentially copy and preferentially mate with people with the best memes — in this case the best language” (Blackmore 1999, p. 104). Blackmore’s assumption thus implies two kinds of selection, one acting directly through imitation, a central principle in her works on memetics, and one through sexual selection.

I will now discuss these ideas in light of the present model, and attempt to distinguish between two particular evolutionary mechanisms. The first mechanism comprises the reproductive benefits induced by language, in other words, language as evolved by natural selection. The second comprises the mechanisms that might have caused language to evolve without reference to any influence on the genes of the host, the speaking human. As a consequence of these, I will introduce a third mechanism that has to do with the time children have available for learning.

The existence of the first mechanism is commonly accepted. It is associated with conventional natural selection and thus coupled to genetic changes, in particular those implying the growth of the human brain. I would like to be more specific about natural selection. Following Blackmore, I think that in the case of language, sexual selection is the most powerful mechanism; I will outline this view as follows.

By and large, an adult of high verbal ability will gain a high social ranking. Such people are supposed to enjoy more frequent access to mates and certainly also better access to food and other necessities of survival. In other words, they have greater reproductive success than the average citizen. But such high verbal ability is made possible by a genetically determined high brain capacity, and hence the selection on behalf of language also benefits genes for great brain capacity. In Deacon’s words: “The remarkable expansion of the brain that took place in human evolution … was not the cause of symbolic language but a consequence of it” (Deacon 1997, p. 340).

In attempting to tackle the question of language evolution in terms of memes, Susan Blackmore has apparently come up with a similar line of reasoning. She suggests what she calls a meme-gene co-evolution that functions as follows. People are assumed to mate preferentially with those possessing the best language ability. “These people then pass on genetically whatever it was about their brains that made them good at copying these particularly successful sounds. In this way, brains gradually become better and better able to make just these sounds” (Blackmore 1999, p. 104, emphasis in the original).

Of course, other factors quite independent of language contribute to the selection pressure for a large brain, factors such as skill at hunting, foraging, tool making, and warfare. These are mainly coupled to natural selection and contribute to the co-evolution of language and the brain. Actually, one must assume that such abilities were the target of selection for increased brain capacity before humans crossed the symbolic threshold, to use Deacon’s terminology.

My second item, language evolution without reference to influence on genes, brings us back to Blackmore’s suggestion that language is acquired by imitation. The important point in this view is that one must assume that the best speakers have the greatest impact. In other words, imitation implies a process of selection: the best speakers are selected as models.

This process is most important in children’s learning. In their imitative activities, children are strongly influenced by peers with more developed language ability, in particular by older children. But of course, children predominantly learn from their parents and this circular coupling, parents-children-parents, opens up the possibility of an additional mechanism that I will call feedback.

Let us imagine a human being at the dawn of humankind who is endowed with a somewhat better than average talent for the primitive spoken language of the time. This person probably influenced his or her children such that they acquired language more efficiently than did other children. When these children became parents the process was repeated, but, as one may assume, at a somewhat higher level. Such positive feedback improves the language ability of those participating in it. Of course, one can also imagine a negative feedback process having a corresponding opposite effect. Thus, although not having any direct total effect on the evolutionary rate of change, such feedback affects the degree of variation in language ability that, in a second step, influences the rate of change as follows.

Every feature exhibits a certain variation, formed as a Gaussian clock curve, as measured with regard to its fitness in the actual environment. Natural selection works on this variation both by favoring the best-fit features and by disfavoring the less fit. In this way, the clock curve is successively displaced, which illustrates the continuity of evolutionary change. Obviously, the selection pressure is stronger the more a feature deviates from the mean value, and this has an important bearing on the functioning of the discussed feedback process in the evolution of language. Both the positive and the negative feedback processes contribute to broadening the distribution curve of language ability. However, since the selection pressure is strongest on the most deviating features, such broadening will strengthen the selection pressure. So when considering the evolution of language as a selection process at which people preferentially copy those with the best language, we may conclude that the feedback process will increase the pace

of evolution towards higher language ability.

The suggested feedback process differs somewhat from what is characteristic of feedback circuits. In an electronic amplifier, say, a limited part of the output signal is fed back to the input of the same unit and causes a reinforced amplification. In the feedback that reinforces the evolution of language, the information is fed back to children in the next generation, and thus not to the same “unit.” However, to label the suggested mechanism as feedback is legitimate, I think, because of the otherwise great similarity in function.

Of course, someone could object that other animals’ ordinary reproductive processes also involve such feedback in the transition from parent to offspring. However, I think that the difference lies in the learning process human children undergo in childhood. This learning process enhances parental influence and it is that very process that may be seen as analogous to amplification in electronic amplifiers, whereas in animals the offspring’s behavior is determined mainly by its genetically unchangeable constitution.

Before leaving the discussion of the evolution of language I will suggest a third process that is a consequence of the mechanisms so far discussed. As we have seen, these mechanisms imply enhanced childhood learning and thus earlier language acquisition. This contributes to the evolution of language as follows. By acquiring language earlier in life the child will have more time to practice talking before adulthood, hence improving its verbal competence as an adult.

To sum up, there are two categories of mechanisms that I believe underlie the evolution of language. The first is the co-evolution of language and the brain, and this has an impact on genes. This mechanism works mainly through sexual selection. The second mechanism has no direct impact on genes, but works through imitation and an associated selection process. In addition to this selection process I have suggested a feedback process that either enhances or weakens parental influence, thus reinforcing the selection pressure and increasing of the rate of evolutionary change in favor of higher language ability. Finally, I have proposed a third mechanism according to which the enhanced and hence earlier acquisition of language lengthens the time children have available for learning. This mechanism would also contribute to increasing of the rate of evolutionary change.

The suggested processes are mutually reinforcing, resulting in progressively enhanced language sophistication. They also demonstrate the importance of children’s learning for the evolution of language, thus supporting the present model of condensation as an intrinsic principle of all evolution.

## 6. The evolution of science

Let us now examine the most recent phase of cultural evolution, the evolution of science. I use the word science to refer to the natural sciences, including mathematics.

There is a widespread notion that science evolves in a way analogous to biological evolution. Thus Hull (1988) suggests that scientific ideas develop in lineages in an evolutionary process. A similar notion is expressed by Plotkin (1993), in his discussion of science as a product of a

“Darwin machine.”

Holton and Brush (1985, p. 196), in their cogent account of the evolution of scientific concepts, analyze the analogies between the biological and scientific evolutionary processes in a most clarifying way. These authors compare the evolutionary mechanisms of species and those of science, finding four common points.

First, both processes presuppose continuity. This means that a species or a science can persist only if there is some stable means for handing on its structure from generation to generation. In science this continuity is identified as the operational and quantitative nature of concepts.

Second, there is the mechanism of mutation, the opportunities for individual variations. In science, these variations are assured by the boundless creativity of the human mind.

Third, there is the mechanism of multiplicity of effort. Science and species alike must rely on a large number of individual attempts from which ultimately come those few that turn out to be useful. The innumerable pages of scientific research documentation of past years testify to the wastefulness of this process.

Fourth, scientific theories are subject to a selection process not unlike Darwinian selection acting on mutant forms. Scientists create a multitude of competing theories and concepts; these are subject to tests in which the internal rules of science are applied, as are external conditions such as applicability, potential for further development, or contribution to social welfare.

It must be emphasized that the type of selection acting on concepts is not coupled to selection acting on human genes. In other words, there is no coupling between scientific ability and reproductive success, as we posited in the case of language. A scientist does not give birth to more children than the average citizen. I therefore conclude that the leash between genes and culture, to use Edward Wilson’s oft-cited metaphor, is broken as far as the scientific part of culture is concerned. Another consideration is that science has only been in existence for a few thousand years, too short a time to have had any influence on genes.

Though the similarities between scientific and biological evolution as discussed in the cited literature are descriptive analogies, that does not mean that they lack explanatory power. Selection in particular is commonly regarded as the decisive cause of scientific progress. In our own time such selection is intentional, performed mostly by scientists themselves. However, I doubt that such intentional selection was really operative in earlier epochs during which scientific thinking was tentatively forming its own character.

If one wants to posit a non-intentional selection process acting on the early evolution of science, one must pose the critical question: who would benefit from this selection? Since the benefit cannot be for the genes, the remaining possibility is that science has evolved to its own advantage, an idea in line with the notion of how memes work.

When humans passed the threshold of the symbolic, they obviously increased their ability to imagine abstract concepts. But in mathematics and science the employment of abstract thinking

is still more accentuated, mathematics being totally abstract. The fact that we nevertheless put considerable effort into these activities shows, I believe, that such conceptual thinking must have increasingly attracted people. It must have been, as it still is, irresistibly attractive to the human mind to create the abstract conceptions that have led to the metaphysical notions of mythologies and religions, as well as to the logical and rational explanations of natural phenomena that have led to science. Deacon (1997, p. 421) reminds us in this context of the story of Archimedes running naked through the streets yelling “Eureka!” This myth illustrates the fascinating experience of recoding a familiar observation into abstract scientific concepts, and I think such experiences have had vital influence on the evolution of scientific thinking. Even more importantly, children’s minds must also have developed such a fascination in understanding abstract concepts.

We find another approach to selection in the reasoning Terence Deacon applies to the case of language. Regarding science, we may tentatively infer that scientific concepts are under intense selection pressure: in their reproduction from generation to generation, they must pass through a narrow bottleneck — children’s minds. Scientific concepts that can be learned quickly and easily by children and adolescents will tend to get passed on to the next generation more effectively and more intact than those that are more difficult to learn. For example, we may consider the invention of the position system that made arithmetic calculations much easier. This efficient although more abstract system quickly replaced the Roman way of denoting numbers and thus contributed to the evolution of mathematics.

Scientific concepts are, I believe, typical memes and as such must have proliferated because they are attractive to the human mind. This is also selectively advantageous for their own reproduction in their specific environment, the human brain. In this enterprise they have continuously been exposed to competition with other memes, most pronouncedly, with those of religion. It is thought provoking to interpret White’s (1896) classic account of the “warfare” — to use his own word — between science and religion as a struggle between memes.

The discussed process works by adapting concepts to children’s learning abilities and also, possibly, by adapting teaching content and practices in view of the requirements imposed by science itself. As a result, the learning of specific scientific concepts is being accomplished at successively earlier ages. In other words, our discussion supports the notion, as predicted in the present model, of the condensation of scientific concepts.

Let us discuss a second example. Euclidian geometry in its original form has been taught to children for about two thousand years. Nowadays in many countries such axiomatic geometry is replaced by abridged versions. This is because children can generally learn such a simplified geometry in less time, allowing room in the curricula for other parts of mathematics that have developed in the meantime and are considered as more important.

Earlier learning enables young adults to begin contributing to scientific knowledge from an earlier age, and hence contribute for a greater total time. These people will thus exert a greater than average impact on the growth of science and, presumably, on education.

The accumulation of concepts is equivalent to the terminal addition that we discussed in the case of biological evolution. By and large, this process causes a prolonged learning period in spite of

the superimposed action of condensation.

It should be noted that the present model does not claim to cover the modern phase of scientific evolution, i.e., the last three hundred years or so. In our own time, however, a school system has been developed with the explicit purpose of, among other things, enhancing science learning. Moreover, recognition of the importance of science education has in recent years resulted in the development of the specific research disciplines of mathematics and science education, the aim of which is to improve instruction methods and thereby, of course, learning. Extensive educational projects with such a purpose have been developed, such as Project 2061 (AAAS 1989). No doubt, the school system has strongly enhanced concept learning in science, presumably also causing a rapid increase of condensation.

To sum up this discussion of the evolution of science, in addition to the general notion of intentional selection, for earlier epochs of science history I have suggested two co-operative mechanisms for the selection process. The first is a self-selection of scientific concepts for their own benefit in their own environment, the human brain. This can be considered a memetic explanation. The second is related to the way children and adolescents acquire scientific concepts, implying an adaptation of concepts to those most easily learned. This second mechanism leads to earlier acquisition, thus extending the individual's active period as scientist, and hence contributing to the progress of evolution.

## 7. Summary

Analysis of various manifestations of evolutionary processes in the human lineage indicates some common principles. One such principle is the crucial impact of development as manifested in embryonic development, in the verbal development of the child, and in the conceptual development of the adolescent mind. These developmental processes are characterized by a specific kind of trait, cumulative traits, which turn out to reveal a pattern that unites the evolutionary process in its biological, verbal, and scientific manifestations. This pattern is a consequence of condensation, a regular shortening of developmental stages. An explication of condensation remains to be made.

The present model shares central ideas with universal Darwinism in as much as it points out mechanisms common to both the biological and cultural realms of evolution. In so doing it also demonstrates the explanatory power of the notion of the meme, since symbols and concepts, as typical memes, are the cornerstones of culture. There is another, somewhat more speculative conclusion. If the observed trend toward the condensation of developmental stages, maintained since the emergence of early life, is not to be broken in our own time, the continued improvement of science and mathematics education is imperative.

I have also discussed several principles for the selection of information to be transmitted between generations in the various manifestations of the evolutionary process. Biological evolution employs natural selection of genetic information. As to the evolution of language, I suggest a process of selection acting cooperatively on both genes and memes. Finally, in the evolution of science, I conjecture that early in the evolution of science memetic selection was in charge, while in our own time we humans have taken over, and carry out an intentional selection. This

sequence of principles of selection should be regarded in the light of the pattern visible in Figure 1, whereby one can perceive how the various principles of selection have contributed to the steadily increasing rate of evolutionary change.

## References

- American Association for the Advancement of Science (AAAS). 1989. Project 2061: Science for All Americans. Oxford: Oxford University Press.
- Arthur, W. 2002. The Emerging Conceptual Framework of Evolutionary Developmental Biology. *Nature* **415**:757–764.
- Barkow, J. H. et al. 1992. The Adapted Mind. New York: Oxford University Press.
- Blackmore, S. J. 1999. The Meme Machine. Oxford: Oxford University Press.
- Blackmore, S. J. 2000. The Power of Memes. *Scientific American* **283**(4):52–61.
- Blute, M. 2003. The Evolutionary Ecology of Science. *Journal of Memetics—Evolutionary Models of Information Transmission* **7**(1). [http://jom-emit.cfpm.org/2003/vol7/blute\\_m.html](http://jom-emit.cfpm.org/2003/vol7/blute_m.html)
- Darwin, C. 1859. On the Origin of Species. London: Murray.
- Dawkins, R. 1976. The Selfish Gene. New York: Oxford University Press.
- Deacon, T. W. 1997. The Symbolic Species. New York: W. W. Norton & Company.
- Dennett, D. C. 1995. Darwin's Dangerous Idea. New York: Simon & Schuster.
- Ekstig, B. 1985. The Regular Condensation of Development as a Mechanism of the Evolution of Man and Culture. *Evolutionary Theory* **7**:195–204.
- Ekstig, B. 1994. Condensation of Developmental Stages and Evolution. *BioScience* **44**(3):158–164.
- Gould, S. J. 1977. Ontogeny and Phylogeny. Cambridge, MA: Harvard University Press.
- Gould, S. J. 1989. Tires to Sandals. *Natural History April*, 8–15
- Holton, G. and S. Brush. 1985. Introduction to Concepts and Theories in Physical Science. Princeton, NJ: Princeton University Press.
- Hull, D. 1988. A mechanism and its metaphysic: an evolutionary account of the social and conceptual development of science. *Biology and Philosophy* **3**:125–155.
- Lumsden, C. J. and E. O. Wilson. 1981. Genes, Mind and Culture. Cambridge, MA: Harvard University Press.

- Piaget, J. and R. Garcia. 1983. Psychogenèse et Histoire des Sciences. Paris: Flammarion.
- Pinker. S. 1994. The Language Instinct. London: Penguin Books.
- Plotkin, H. C. 1993. Darwin Machines and the Nature of Knowledge. London: Penguin Books.
- Raff, R. A. 2000. Evo-devo: The Evolution of a New Discipline. *Nature Reviews Genetics* **1**:74–79.
- Stearns, S. C. 1992. The Evolution of Life Histories. Oxford: Oxford University Press.
- Tooby, J. and Cosmides, L. 1992. The Psychological Foundations of Culture. In *The Adapted Mind*, edited by Barkow, J. H. et al. New York: Oxford University Press.
- Wagner, G. P., C-H Chiu and M. Laubichler. 2000. Developmental Evolution as a Mechanistic Science: The Inference from Developmental Mechanisms to Evolutionary Processes. *Amer. Zool.*, **40**:819–831.
- White, A. D. [1896] 1960. A History of the Warfare of Science with Theology in Christendom. Reprint, New York: Dover Publications.

# A JUSTIFICATION OF SOCIETAL ALTRUISM ACCORDING TO THE MEMETIC APPLICATION OF HAMILTON'S RULE

By  
John R. Evers

## I. Introduction: Genes and Altruism

Charles Darwin described the sterility of certain castes of social insects, and more generally, the reproductive self-sacrifice such organisms represented, as "one special difficulty, which at first appeared to me insuperable, and actually fatal to my whole theory."<sup>[1]</sup> In the 1960's, W.D. Hamilton "inaugurated" the theory of "kin selection," which offered a brilliantly simple explanation for such altruistic behavior.<sup>[2]</sup> As Holldobler and Wilson explain, "Hamilton recognized the importance of a measure he called *inclusive fitness*, which incorporates both the individual's personal reproduction (classical fitness) and its influence on the reproduction of collateral relatives."<sup>[3]</sup>

The essentials of kin selection and inclusive fitness are summarized according to a simple equation, called "Hamilton's Rule," which is expressed:  $C/B < b$ . "This says that the cost C (which is the loss in expected personal reproductive success through the self-sacrificing behavior) divided by the benefit B (the increase in the relatives' expected reproductive success) must be less than  $b$ , the probability that the relatives have the same allele,"<sup>[4]</sup> if the altruist gene is to survive natural selection.

Inclusive fitness begins by focusing solely at the level of the gene, but then widens that focus to encompass the entire group or population comprising any given gene's extant copy-set. For example, assume organism X is a carrier of gene-A and that gene-A causes X to behave in a certain way which kills X prior to the production of any offspring. If X's suicidal behavior directly enhances the survival and reproductive potential of X's siblings, gene-A still has a chance at survival and even proliferation - not in X or X's non-existent offspring, but in X's siblings and their offspring since at least some of X's siblings will carry copies of gene-A. One copy of gene-A is just as good as the next; any single gene-copy is just as transitory as the organism which carries it.<sup>[5]</sup> Indeed, inclusive fitness demonstrates that the true fundamental unit of natural selection is a gene's entire copy-set.

For this reason, in the expression  $C/B < b$ , everything hinges on  $b$  - the probability that the altruist's siblings carry copies of the same altruist gene. As Dawkins observed, "[k]in selection accounts for within-family altruism; the closer the relationship, the stronger the selection."<sup>[6]</sup> Thus, the "family" is merely a short-hand approximation of a representative portion of any given gene's copy-set (according to the degree of relation within the family), and changing degrees of relatedness will directly affect levels of intra-family altruism.<sup>[7]</sup>

The obvious limitation of Hamilton's Rule is that it only justifies altruism expressed within the family. Alternative theories, such as reciprocal altruism, can account for the natural selection of generally cooperative behavior regardless of familial relationships, but theories such as reciprocal altruism do not necessarily account for altruism.<sup>[8]</sup> As Wright points out, "[reciprocal

altruism] doesn't involve sacrifice for anyone who doesn't ultimately reciprocate."<sup>[9]</sup> Hence, reciprocal altruism does not account for "sacrifice" at all. For the purposes of this analysis, an organism that is to be considered an "altruist" must suffer a *net* cost (in terms of reproductive potential) as a result of its behavior, as it was primarily this feature that established the monumental problem which challenged Darwin in 1859 and which was solved in part by Hamilton's Rule.

The intra-family limitation of Hamilton's Rule is a result of a certain feature of genetic reproduction; namely, that genes can only produce copies by making new carrier organisms in which to house these copies. In other words, gene copies are only distributed *vertically* (down family trees) within a population. Because genes distribute their copies in this fashion, the only easily identifiable groups of organisms that will have some *a priori* reason for possessing high *b* values in Hamilton's Rule is the genetic family. Thus, while Hamilton's Rule may have solved the paradox of completely self-sacrificial behavior within the confines of the genetic family, it simultaneously precludes the possibility of altruism directed beyond the genetic family since it is unlikely that a random, non-family member will share a copy of the relevant altruistic gene. When C/B is less than *b*, self-sacrificing behavior can be seen as successful altruism; but, when C/B is greater than *b*, the same behavior begins to look more like genetic suicide. In fact, altruism of any degree should face the threat of competitive elimination where the *b*-value is zero.

Moreover, any benefit enjoyed by an organism not carrying the altruist gene (a "non-carrier") will necessarily dilute the net benefit that would otherwise accrue to the altruist gene's copy-set. Therefore, if an altruist gene is to succeed, some mechanism must evolve which can eliminate (or, at least, offset) this potential dilution. By functioning to decrease any potential benefit to competitors and increase potential benefit to the altruist gene's copy-set, this mechanism will manifest itself generally as an increase in competition between carriers of the altruist gene and their non-carrier competitors. Thus, the inverse operation of Hamilton's Rule dictates levels of inter-family competition generally proportional to any degree of intra-family altruism.

However, if some non-genetic unit of natural selection can allow for a sufficiently high *b*-value among non-siblings (i.e., between any randomly selected members of a population), then the threat posed by the inverse operation of Hamilton's Rule is removed. Richard Dawkins introduced the generally accepted name for this alternative unit of natural selection: it is the "meme."<sup>[10]</sup>

## II. Memes and Altruism

Any meme can be defined generally as a *rule of behavior*, encoded by functional neuronal groups or pathways. Behavior is action, whether mental or physical. Ideas such as *tying shoe-laces* or *opening a door* represent rules of physical action, i.e., rules of patterned neural-muscular interaction. Concepts such as *apple*, *seven*, or *causality*, represent rules of mental action, or rules of cognition, i.e., rules of patterned neural-neural interaction. Hence, physical movement is governed by memes which represent rules of physical action and thought is governed by memes which represent rules of mental action.<sup>[11]</sup>

If memes are units of natural selection (allowing for a memetic application of Hamilton's Rule); and, if memes copy horizontally (allowing the "memetic family" to extend beyond the scope of the genetic family); and, if memes can be directly responsible for altruistic behavior, then memetically driven and inclusively fit altruism can extend to the whole of any given (freely communicating) human population according to the mathematical purity of Hamilton's Rule.

### A. Memes as units of selection

Most, if not all, models of selection share three basic elements: (1) variability of base units; (2) competition between base units for reproductive resources, or selective pressure; and, (3) selective reproduction (or, selective longevity).[12] Memetic variation is demonstrated by the approximately one trillion (1,000,000,000,000) neurons which make up the average human brain, with the cerebral cortex accounting for at least 50 billion of these nerve cells.[13] Moreover, because anyone who wants to share any idea with another person *must* create a physical substrate for the idea (e.g., language, whether spoken or written), the full scope of memetic variability is most accurately described by combining the subjective knowledge of *each* living human carrier with the full body of extant objective knowledge (e.g., knowledge encoded in books, or computers).

Considering a meme as a discrete neural group or pathway, it follows that memetic competition is fueled by the need for (finite) neurochemical resources and (limited) neural space.[14] That neural patterns or neural groups compete for space and neurochemical resources is not a novel proposition.[15] Thus, memetic competition for the limited space and biochemical resources available in the brain can be restated simply as competition between neurons or functional groups of neurons for *stimulation*.

Rules of behavior which are followed and reinforced to increase the probability of future stimulation are *remembered*. Memes which fail to acquire the necessary resources for reproduction, whether manifested in physiological or non-organic material substrates, will face competitive elimination, i.e., they will be forgotten. Every human activity, from basic cognition to daily "decision-making," can be seen as a process of natural selection during which functional neuronal groups compete to occupy a behavioral (whether mental or physical) niche within a carrier organism's nervous system. Hence, certain memes are selected over others based on the neurochemical characteristics of the underlying neuronal group(s) or pathway(s), as well as their resulting classical and inclusive fitness.

### B. Horizontal Memetic Reproduction and Altruistic Memes

The second and third demonstrations, that memes are capable of horizontal reproduction and that memes are capable of encoding altruistic behavior (assuming such behavior exists at all), are easily made as they are both self-evident propositions.

Genetic reproduction occurs through an elaborate process of cellular fusion and the consequential growth and maturation of newly formed carrier organisms. Memetic reproduction occurs through an equally elaborate process of imitation and/or linguistic communication, whereby memes are copied from one nervous system into another. That memes are capable of

*horizontal* reproduction (i.e., that memes are not limited to copy distribution within genetic families) is a self-evident proposition, provided the recipient of the proposition is not a genetic sibling of the one advancing the proposition. Indeed, any attempt to dispute the proposition that human beings *generally* are capable of sharing ideas is to advance or attempt to share the idea that ideas cannot be shared. Therefore, it is literally beyond argument that memes are capable of horizontal reproduction.

Finally, while genes encode the assembly of polypeptides and thereby transcribe physiological structure, memes represent patterns of neural stimulation and thereby transcribe *behavior*. Therefore, if anything is to be *directly* responsible for altruistic behavior it will be a meme, whether the meme is genetically encoded (i.e., innately acquired), or learned (i.e., environmentally acquired).[16]

### **III. Memetic Application of Hamilton's Rule**

As currently applied, Hamilton's Rule is based solely on the static nature of genetic identity: because an organism's genetic makeup is fixed for the duration of the organism's lifespan, any two organisms will either share a copy of a gene or they will not, and this objective fact is measured according to probability and familial relation. However, since memes are capable of horizontal reproduction (whether through imitation or linguistic communication), it is possible that any meme, including an altruist meme, could express itself in favor of (and, therefore, presumably in the presence of) a non-carrier competitor *and thereby copy itself into the memetic (i.e., neuronal) structure of the competitor*. In short, a competitor directly experiencing the benefits of another's self-sacrifice *might* subsequently imitate such altruistic behavior for no other reason than having experienced such behavior first-hand. Consequently, one should expect various factors (whether genetically encoded, memetically encoded, or both) to evolve throughout a population which would enhance or detract from the probability of such imitation (e.g., empathy, admiration, so-called "open-" or "close-mindedness," etc.).[17] Any given meme or set of memes could likewise evolve various mechanisms to increase the probability of such imitation (e.g., cognitive or psychological appeal, etc.).[18] Hence, it is conceivable that some factor or set of factors could be delineated that would determine the relative *rate of conversion* (from non-carrier-competitor to carrier-clone) for any given meme within a certain population.

This rate of conversion shall be expressed as a new variable, to be added to the right side of Hamilton's Rule since it will operate to increase the likelihood that the beneficiary organism(s) will carry a copy of the same altruistic meme (following the altruistic encounter). Because the rate of conversion will have no application where a beneficiary already carries a copy of the altruist meme, it will only modify the probability that a beneficiary does *not* carry a copy of the altruist meme, which is described by  $(1-b)$ . Therefore, the memetically adapted Rule is:  $C/B = b + c(1-b)$ , where the rate of conversion [ $c$ ] measures the *likelihood that expression of a meme will "infect" a non-carrier and thereby convert the non-carrier into a carrier*.

Operation of this adapted version of Hamilton's Rule can be stated as follows. Assume that there exists a 1/4 (or 25%) degree of memetic relation within an intra-communicating population. Hence, each member of this population shares roughly 1/4 (or 25%) of the same memes. Assume also that a member of this population carries an altruist meme that has a conversion rate of 3/4

(or 75%). Hence, for every four non-carriers, three will be converted into carriers upon contact with the meme. According to these figures, there is a 13/16 (or, 81.25%) chance that expression of this meme in favor of *any* randomly selected member of the population will operate to confer a direct benefit on a representative member of the altruist meme's copy-set. Whether a *net* benefit is likely to be conferred by expression of that meme will depend on comparing this figure with the Cost-Benefit ratio defined by the left side of Hamilton's Rule.

#### IV. Conclusion

According to the memetic application of Hamilton's Rule, the key factor to success for altruistic behavior is not a high probability of clone status (with regard to the altruistic gene) based on genetic familial relation, but a high probability of clone status (with regard to the altruistic meme), based on a population's horizontal reproductive fertility. Indeed, a highly infectious, altruistic meme (e.g., 80% conversion rate) existing in a small fraction (e.g., 10%) of the population would still enjoy a very high probability (72%) that expression of any copy, in favor of any random member of the population, would serve to benefit that meme's (growing) copy-set (all other things being equal). Assuming acceptable criteria and empirical data can be established to give real meaning to the "rate of conversion" for any given meme, the memetic application of Hamilton's Rule offers a comprehensive justification for general (intra-societal, or intra-cultural) altruism. Meanwhile, it can be said with newfound certainty that purely altruistic behavior is *possible* within any memetically fertile population.

#### Footnotes

[1] Holldobler and Wilson, *The Ants*, (Harvard University Press, 1990), at p.181.

[2] Id.

[3] Id.

[4] Id.

[5] A sequence of nucleotides which dictates the construction of a specific polypeptide can be classified as a "gene," and can be analyzed as being fit or unfit according to natural selection; but, it cannot be called a "survivor." Likewise, a discrete quantum of functionally synchronized neurons can be classified as a "meme," and can be analyzed as being fit or unfit according to natural selection; but, it cannot be called a "survivor." Any particular sequence of nucleotides or quantum of neurons will degrade prior to or contemporaneous with the death of the carrier organism. Indeed, nucleotides and neurons are as much a part of the organism as is any protein, enzyme, cell, organ, or physiological system. The only thing that is *carried* by an organism (in the sense that it can outlast the organism) is the underlying *pattern* of nucleotides or neurons, which exists purely as a speculative construct of subjective consciousness. The concept of the potentially immortal gene/meme can be made to represent a concrete reality only by redefining the functional "gene" or "meme" as the total extant copy-set of any given particular genetic/memetic manifestation, i.e., as a *population* of gene/meme copies. Any individual member of such a population may be transitory, but the population itself can survive indefinitely.

[6] Richard Dawkins, *The Selfish Gene* (Oxford University Press, 1989), at p. 94.

[7] See e.g., Robert Wright, *The Moral Animal* (Vintage Books, 1994), at 164.

[8] See Dawkins, *The Selfish Gene*, supra note 6, pp. 202-233 for explanations of various strategies for cooperation and mutual benefit. See also, Wright, *The Moral Animal*, supra note 7, at pp. 189-209. According to Wright, reciprocal altruism is the only available explanation for inter-family altruism and, therefore, "wins by default." Id. at 202.

[9] *The Moral Animal*, supra note 7, at p.207.

[10] *The Selfish Gene*, supra note 6, at p. 192.

[11] A possible third category of action is linguistic action, which represents a means of converting physical action into mental action, i.e., language is movement that encodes thought.

[12] See e.g., Edelman, *Neural Darwinism* (Basic Books, 1987), at p. 9.

[13] Kandel & Schwartz, eds., *Principles of Neural Science* (N.Y.: Elsevier Science Publishing Co., 2nd Ed. 1985), at 223.

[14] It also follows that memes compete for access to the limited and finite materials required for manifestation in non-organic memetic carriers, e.g., book space, computer storage space, etc.

[15] For example, Edelman describes the "Mnemon hypothesis," proposed by J.Z. Young in 1965, which describes a means for selective competition amongst neuronal groups competing for stimulation. (*Neural Darwinism*, supra note \_\_\_, at 14-15). Edelman's own theory of neuronal group selection is based on the concept of neuronal competition for stimulation. (Id. at pp. 45-46).

[16] For the purposes of this analysis, innate behavior describes behavior which is the immediate result of memetic expression, but which is ultimately the result of genetic expression, i.e., the responsible memes are formed as a result of genetic expression. Learned behavior is the product of environmentally acquired memes which are, in turn, the product of epigenetic, somatic selection. An innately acquired meme that is incapable of horizontal reproduction is the functional equivalent of a gene.

[17] See generally Aaron Lynch, *Thought Contagion* (Basic Books, 1996), at pp.2-16.

# **STROLLING THROUGH THE MEMETIC MINE FIELD**

**By**

**Leon Felkins**

WRITTEN: 28 AUGUST, 1995

REVISED: 9 SEPTEMBER, 1995

## **INTRODUCTION**

Isn't it strange that in the past people were so peculiar in their tastes, their prejudices, their sexual attitudes, their treatment of animals, and their family relationships? Even now, people in some parts of the world have some really strange ideas about how to feel and behave. Were our ancestors stupid? What about the great scientists and philosophers of the past, who seemed very intelligent, yet showed little understanding for the problems of women and minorities? Many of the great leaders and statesmen of the past even had slaves. Are we just smarter now -- or what?

No, those strange folks in other times and other lands were making decisions in the same way you do now -- decisions based on the memes and genes that inhabit their minds.

We all know what genes are, but maybe we are not too clear on "memes". Basically, memes are beliefs that spread from individual to individual; that is, they replicate somewhat as genes do. Note that the concept of "memes" itself is a meme!

In this essay I will give a brief introduction to how the mind is influenced by memes and genes and give examples of the many memes that have great control over our lives.

## **ME AND MY MIND**

While the subject of the mind and exactly how we humans control ourselves is very complex and somewhat controversial, a limited review is possible that will be adequate for the purposes of this essay. (An excellent reference on this subject is [DENN]).

Our mind controls our body somewhat like an operating system controls a computer. While the task of genes is primarily to determine the physical characteristics of our bodies, including our brain and nervous system, they also influence decisions of the mind. For example, the fact that genes make us feel great pleasure in sexual activity certainly has a great influence on the mind.

But the mind is also strongly influenced by beliefs -- and that is where memes come in. In his book, *The Selfish Gene* [DAWK1], Dawkins defines a meme as a replicating information pattern that uses minds to make copies into other minds. That is, memes are thoughts, ideas, beliefs, prejudices, etc. that replicate. Dawkins writes:

*"Examples of memes are tunes, ideas, catch-phrases, clothes fashions, ways of making pots or of building arches. Just as genes propagate themselves in the gene pool by leading from body to body via sperm or eggs, so memes propagate themselves in the meme pool by leaping from brain to brain via a process which, in the broad sense, can be called imitation."*

(Several good papers on this subject are available on the Internet, for example see [HENS] and [HALE]).

The memes associated with morals are often referred to by the term "mores" which my dictionary defines as, "folkways of central importance accepted without question and embodying the fundamental moral views of a group".

For our purposes here we will assume that the individual's mind makes its decisions by means of analysis of the inputs as augmented by the genes and memes. Details on just how genes and memes influence the mind will be discussed in the next sections.

A comparison to computers may be helpful: Genes correspond to the ROM that is built into computers. Memes correspond to the programs that run on the computers, along with the associated data, and are loaded subsequent to the initial boot. The mind corresponds to the Operating System. Roughly.

Note that the origin of a meme could be based on fact or logical analysis. Memes should not be equated with myths. But the meme once established in the mind is just a belief -- regardless of its origin.

Memes are comparable to genes in the fact that they replicate and that they influence our behavior. But how memes influence the mind is quite different from the way genes do. Let's look at each approach.

## **Genes**

Our genetic psychological factors are powerful forces that tell us what to do but, unfortunately, are slow to adapt to a changing world. Tragically, they are just not adequate for humans interacting with a rapidly changing world. We are struggling today with emotions -- provided by our genes -- appropriate for the caveman.

Genes cause problems because they provide psychological directions for a particular situation that is no longer appropriate. The evolved genetic solutions for certain problems apparently worked at one time but may no longer do the job. Unfortunately, it takes thousands of years for the genes to be modified.

Genes affect our behavior in two ways: physical characteristics and sensual enhancement.

## **Physical Characteristics**

Many of the characteristics of humans, both physical and mental, are determined at birth. We all are unique when it comes to capabilities, emotions, and physical makeup. One person's sexual drive can be as different from another's as their ability to play music or do mathematics is.

## **Sensual Enhancement**

Our physical characteristics, our hormonal makeup, and our other inherited factors color the sensations sent to the brain from which it must make decisions. That is, from a given set of external conditions, the message received by the brain is not the same for every individual. Moreover, the brain's ability to process these messages is vastly different from individual to individual. Therefore the decision made by two different brains, based on the same external conditions, could, and likely will, be quite different for two different individuals.

Now, to complicate the matter further, the memes get involved.

## **Memes**

To the brain, memes are additional facts that are combined with the genetic input (discussed above) from which a decision is made. Even though they are just beliefs, for all practical purposes to the brain they are facts.

For efficiency, beliefs are accepted as factual and the brain moves on. If this approach were not taken, our brain would get bogged down in analysis to the point that it would never be able to make a timely decision.

Unlike genes, these memes, or "facts on file", may be added, removed or changed many times in a person's life time. The ability for humans to rapidly change our society and our surrounding environment, made it necessary for some adaptive mechanism to be built in our mind that would allow for these rapid societal and environment changes.

As we noted above, memes are more effective than genes in dealing with modern problems due to their ability to rapidly develop. That is, while genetic forces have evolved over many centuries to adapt to particular problems, memes can evolve in a lifetime or less. The price to pay for this quick reaction is that there is little chance for corrections when the memes may be in error. In fact, the behavior they evoke is not necessarily best for humans (but may be best for the memes - - see [DAWK2]). They just evolve, mainly based on short term phenomena, limited information and often, ignorance.

Unfortunately, memes generate psychological forces and emotions that are just as powerful as genetic forces. Sexual mores, to be discussed shortly, best illustrate the power of memes.

It is difficult for each of us to make decisions without using memes. Since memes are involved in our thinking process, we may have to use memes to evaluate other memes! Even when applying our best effort to think precisely and logically, we may still be basing our decision on "facts" that are really just memes. After all, the "scientific method" itself is a meme.

## **How Memes get established**

Memes can evolve spontaneously in a group in response to some problem. That is, the meme evolves, somewhat like genes, by adapting to conditions and continuously changing. Most likely

our meme for compassion probably evolved this way since it made life better for everyone in the group.

It is obvious, however, that many memes are "planted" in the group by certain individuals for their own reasons. The politicians, the schools, the press, the clergy and others are guilty of sowing a constant stream of memes within the populace and encouraging their growth. Examples will be given in the next few paragraphs.

Just like genes, the survivability of a meme depends on its environment. In that sense, memes that are compatible support each other's growth and survivability. The meme for religion, for example, is helped a great deal by the meme that says we should have "blind faith".

Yet, strangely, we do harbor contradictory memes.

### **Contradictory Memes coexist**

Since we are constantly bombarded with either new memes or the reinforcement of old memes, it should be no surprise that many of the memes might be contradictory to each other. What is a surprise is that humans seem to be able to have contradictory memes cohabiting within their mind at the same time! An example would be a person who works with computers using analysis and logic all day and yet has many beliefs in religion and sex that he accepts on blind faith.

Now let us look at how memes and genes work together to get us to do their bidding. See [DAWK1], page 192.

### **The Gene/Meme Team**

It appears that memes enhance or amplify the influence of genes in many situations. For example, genes insure that a young man is sexually excited when he views a young voluptuous girl in the nude. Memes, on the other hand, might make this young girl even more attractive *with clothes on* if she were only *partially* clothed; or if she is his best friend's wife; or if she admits having just had sex with someone else; or she is seen buying porno flicks; etc.

Practically, the gene/meme team determines our outlook and our attitudes toward the world.

### **Our Individual World View**

The genes we have inherited and the memes that we have adopted determine what our world view is. They provide the perspective in which we view the world and the basis for our decisions. It is ridiculous for us to make judgments on how other people act in other times or other societies. For our judgments are based on our set of memes, and their actions were caused by their set of memes. Our memes are not superior just because we happen to possess them.

We simply should not be critical of other societies based on nothing more than our own meme/gene makeup. For example, to judge that societies in which marriages are arranged or

where free love is common are inferior to ours is stupid. They live by their memes and we live by ours. That's all.

In fact, it is important to keep in mind that any judgment you make about anything is directly based on -- or at least contaminated by -- memes and genes. Even your opinion of this essay comes from a foundation of memes that this essay may be critical of! So, could you just turn off your memes until you finish reading this article please! Just kidding.

## **Logical Thinking**

Can we think without being influenced by memes and genes? While we can't completely eliminate the influence of memes and genes we certainly can diminish their impact. Further, we can adopt the "Scientific Method" which requires that all conclusions be subject to verification. That is, we can keep an open mind for errors in our thinking and try to minimize accepting anything on blind faith.

## **THE MEMETIC MINE FIELD**

In this section I will list some of the common memes we live with. It is difficult not to pass judgment on some of these memes as many appear not to be in the best interests of our society. But the primary purpose here is to list the memes, not to judge them. Let's start with the richest source of funky memes -- our love life!

### **Sex, Love, Marriage**

#### **"Sex is Immoral outside of Marriage"**

This well known, apparently religious based, meme is blindly followed by many people who aren't even into religion! Many people feel guilt if they have sex outside of marriage and have given little thought to the origin of that feeling.

### **Relations between the Sexes**

There is always a great deal of memetic activity in this area!

### **Sexual Harassment**

There is no doubt that humans, particularly men, have done some pretty heartless and cruel things to members of the opposite sex just to get sexual favors. To combat this problem, laws have been passed and memes have been established. Both, unfortunately, have gone too far.

The current "sexual harassment" meme now deems sexual harassment what used to be harmless and entertaining sexual flirtation between the sexes. Apparently this meme can be turned on or off since some women admit that the actions that would be deemed "sexual harassment" in one situation might be deemed "desirable flirting" in another.

It is interesting that in the past, for a man to have refrained from flirting would have hurt some women's feelings due to the sexual meme that they possessed then.

## **Male/Female Relationships**

This essay is far too short to cover all the memes alive and well in this area! Instead I will only discuss the prejudicial memes and that will be done below in the "Prejudices" section.

## **Death**

Funerals and other ceremonies surrounding a death are powerful and ancient memes that remain unchallenged in the most modern of communities. The United States has spent millions retrieving the bones of service men and others from foreign countries without a thought as to what could be in those bones that make them so valuable.

## **Cleanliness**

The meme that makes us want to be forever spotlessly clean probably, on balance, is a good one but does get a bit silly at times. Lovers, husband and wives, etc. refuse to share a tooth brush but otherwise swap bodily fluids with abandonment!

Even people who actively support environmental causes still insist on a shower and clean clothes everyday without the tiniest consideration of a rational justification for wasting energy on a somewhat fastidious purpose.

## **Politics and Government**

There are too many memes associated with government to enumerate. I will mention a couple, not necessarily the most important.

"The government can fix all our problems"

This powerful political meme causes all of us a lot of grief and expense. In reality, it is quite rare that government ever solves any problems without creating far greater ones, and, at the same time, relieving the taxpayers from a lot of their money. The fact that government has been so successful in instilling this meme in most of the population and in keeping it alive (in spite of all evidence of their success to the contrary), is quite astounding!

"Let the government pay for it"

Another meme that the government has been very successful in establishing is that they can provide services for free. How often do you hear someone say, "It doesn't cost anything - the government is paying for it!" Right!

"The Sanctity of Government"

This meme has taken some hits here in the USA in recent times but will probably survive. Recently, President Clinton had to admonish the people that it was wrong to be critical of the government.

Well, maybe so, but a rational person must realize that the government, including law enforcement agencies, is composed of people looking after their own best interests and who, in fact, have a long history of corruption and deceit. We would probably be better off if we simply gave the government the respect it deserves -- based on its record.

"People of Country X are Evil"

During World War II, the Allied countries very successfully implanted the meme in to the minds of the citizens that the Germans and the Japanese were evil monsters. Later, to a lesser degree, the Russians were made to look evil. Of course, in each case, when peace is made the memes have to be squashed quickly! People fight wars better if they hate each other!

"The primary factors for selecting (male) politicians are his sex life (or, more likely, the lack thereof) and how much hair he has"

Strange but true. To be fair, we did accept Clinton and his sexual escapades, but he is a goner if we find out that he is wearing a toupee!

## **Environment**

The "Concern for the Environment" meme is a very interesting meme. Since our genes promote selfishness, without restraint, the genes would likely have us trash the environment. But let us not be too hard on the genes. Their selfishness was established in ancient times before we had the power to lay waste thousands of acres in a day. Again, the genes are not up to handling a modern situation.

So, the solution is to establish memes that tell us that trashing the environment is wrong. This has been amazingly well done by environmentalists and the government with the help of the news media.

Memes, like genes, by necessity are simplistic so as not to take up too much space in our limited brain capacity. We see people making sure their own trash is picked up at a camp site even though there is a huge pile already there! We see people using vast quantities of water for bathing and electricity for air conditioning while religiously making sure that their old newspapers are carried to the recycle depot.

## **Compassion**

Without compassion, civilization, or at least a decent society, would probably collapse. On the other hand, we know that humans have the capability and the tools to be incredibly cruel to each other.

Periodically, we are confronted with the facts that humans will do horrible things to each other -- especially if they are not held accountable. How can this be if we are compassionate?

The answer is that compassion is most likely not genetic. No, compassion is a meme that is grilled into our heads from childhood. Thank goodness for this one!

## Cause and Effect

The "Cause and Effect" meme is an interesting one. This meme, established by the scientific revolution, says that most situations, such as health, mental states, relationships, etc., have simple causes. A popular example is, "If you eat the right foods and avoid the wrong foods, you will not get cancer or heart disease".

To be fair to the scientists, the statement "established by the scientific revolution" needs some qualification. Actually the meme is a misunderstanding of scientific principles. Scientists may publish a report that says -- I'm making this up -- that there is a correlation between eating a lot of stale hot-dogs and the loss of sexual desire. The facts may be that people obsessed by sports spend a lot of time in stadiums (where they eat a lot of hot-dogs) and have little time or interest in sex. But the news media simply reports that "hot-dogs kills sexual appetite" and that becomes the meme.

That there is a simple cause and effect relationship for most events in the world is a concept that is not necessarily obvious. Yet, strangely, even people who are hostile towards technology and science still believe in this meme ("if you will just take a handful of Vitamin C pills every day, you will not get a cold").

## Prejudices

The dictionary gives several definitions for "prejudice", the first two in mine are:

*"1. an unfavorable opinion or feeling formed beforehand or without knowledge, thought, or reason. 2. any preconceived opinion or feeling, either favorable or unfavorable."*

A common and well established meme in our society would have us believe that the first definition is the correct one (yes, our prejudice against prejudices is itself a meme). Yet the second one is probably more practical since it would allow "prejudices" to refer to both good and bad beliefs.

However, in recent times, due to the tragedies associated with some unfortunate historical prejudices, a powerful meme has been established that prejudices are bad. Since the crux of the definition of "prejudice", "opinion or feeling formed beforehand or without knowledge, thought, or reason" sounds almost like the definition of a meme, we might simply define a "prejudice" as an unfavorable meme.

So, it appears that prejudices are just a subclass of memes that are believed to be bad. Therein lies a "can of worms" for we are all in possession of thousands of memes and whether they are good or bad is a judgment call (made with the help of another meme!).

Let us look at an example or two. Almost everyone agrees that racial prejudice is bad. How about prejudice against people that seem stupid? Speak with a Southern accent? Wear unstylish clothes, are economically disadvantaged (poor), are fat, and so on?

The fact is prejudices have the same advantages and failings that other memes do. They allow the brain to make a quick judgment which is generally necessary for survival. If the prejudice is based on logic and/or facts (definition number 2), then it is useful. If it based on faith or "opinions without knowledge", like memes in general, then it can be bad.

Let's look at a couple of well known examples:

### Racial prejudices

Racial prejudices have caused much grief, cruelty and suffering in the world. Genes may be the culprit here -- in an indirect way. Genes would have us to **only** look after the best interests of those folks that are closely related to us. [DAWK1] People with a different color or from a foreign land certainly don't meet that requirement.

There appears to be some success in diminishing racial prejudice in much of the world. This success is the result of diminishing the control of the "racial prejudice" meme and by establishing another meme that says you are a bad person to have racial prejudices.

### Sexual prejudices

Obviously the sexes have some differences. However, it is not conducive to peace and harmony to believe that one sex is inferior to the other. Yet, creating a meme that says there is no difference in the sexes, as some are promoting, seems to be a bad approach. A better meme would be simply "neither sex is inferior to the other".

### Cool Sayings and actions

Expressions that are considered "cool" come and go and woe to those who use an out of date one!

An interesting aside here is the survivability of the word "cool" itself. It has been around for a couple of generations.

How about wearing your shoes untied and your baseball cap on backwards?

## **Advertising**

Companies know that getting the right memes established can result in profits of millions of dollars! Trademark and trade name recognition along with the associated belief in superior value is a common and profitable type of meme.

A few months ago, while at a party, a friend of mine was induced to take a blind folded taste test of various bourbon whiskies -- including her favorite that is very pricy (another meme!: "quality equates to price"). To her embarrassment, she flunked the test, picking "Jim Beam" as the best. Did that change her buying habits? Of course not!

Brand name acceptance is a powerful meme that the majority of the people unhesitatingly follow. However, there are recent trends that indicate that people are beginning to drift away from that bias and to try products that have generic labels. Maybe there's hope.

## **Clothing**

An amusing but serious application of memes is our attitudes towards clothing styles. Without any particular reason, clothing styles change with time: skirts get short and skirts get long, ties get fat and then they get slim, and so on.

The "clothing styles" meme allows us to clearly see the power of memes. We genuinely feel negative toward a person dressed in clothes that are "out of style" but were in style some years ago. Now, rationally, we know this is ridiculous. How can there be anything wrong with a style that once was a popular style?

Here we have evidence that memes can make us feel badly towards something that is not supported by rational reasoning -- in other words, prejudice. With this recognition in mind, we might want to reflect on the many other memes we so blindly accept.

Many of our young people truly believe that if they will only buy and wear the right brand of athletic shoes, they will be able to do athletic feats that defy the laws of physics. It is a sad reflection on our society that advertisers have been so successful in planting this meme that really exploits the young, especially the very poorest young.

## **Food**

In some parts of the world people would no more consider eating a cow or a pig than you would a dog.

## **Music**

A real tragedy here. Essentially all varieties of music have at least some examples that are enjoyable and some that are quite outstanding. Yet memes dictate, particularly to the young, what "brand" of music they should "enjoy". And they seem to only enjoy that designated music! Amazing, the power of memes!

## **Family**

Our attitudes towards members of our family are greatly affected by memes. These attitudes vary greatly over time and in different societies. In the 60's, the meme got established, with some justification, I suppose, that parents could not be trusted and were generally out of touch with the activities and interests of the current society. This meme has caused a lot of grief and suffering for everyone and is one that needs to be replaced by something more realistic. How about, "Parents are (reasonably) out of date, a little sluggish with the thinking and may have some interests of their own, but they are all I have and it would be in my best interests to stay on their good side".

## **Political Correctness**

This good intended but rather hopeless meme has had amazing success, particular on the college campuses. The driving force seems to be to eliminate any possibility of unfairness or hurting anyone's feelings. To accomplish this, we are told that one word is better to use than another word, even if they have the same meaning, because the words themselves can hurt. So we are told that "intellectually impaired" is better than "stupid", for example.

In view of the rather outlandish requirements that this series of memes puts on our belief systems, it is doubtful that they will survive. There is a limit to what you can make reasonably intelligent people believe!

## **Religion**

Religion has had a great deal of success with memes. In fact, for anyone wanting to do research on the impact of memes, this is probably your best laboratory. There are many memes associated with religion, of which only a few are listed here. The memes listed here are associated with the Christian religion but in most cases are applicable to all religions.

- Infallibility of the Bible
- If you don't rely totally on blind faith, you are a creep!
- The power of prayer
- Heaven will be your home after this life
- Converting others will help you get to heaven
- Sending your money to God's representatives is what God wants you to do

All of these memes have proven to be very successful in promoting Christianity.

## **Academia**

Much of the discussion above on "Political Correctness" applies to academia. However, there are a few more peculiar memes that thrive in that environment.

## **Academic Freedom**

This well entrenched meme holds that academic freedom is more important than the welfare of the students -- or maybe even the university. During my personal experience as a professor in a public university, no administrator ever sat in my class room to evaluate my capabilities! Instead of Computer Science, I could have been teaching flatulence control (and doing a bad job of it, possibly) for all they cared.

"We need more research to understand the world better"

Poppycock! We need research to promote the selfish interests of the universities and the professors. This meme costs the taxpayers a few bucks!

"Doctoral Dissertations make major contributions to our understanding of the world"  
Ha! Go read some of them!

## **Social Relationships**

"We should cooperate"

Social scientists say that the genes would have us look only after our own selfish interests even when cooperating would be better. But, thanks to some powerful memes, most people seem to want to cooperate even when it may not be in their best interests to do so. This behavior is apparently caused by a good meme, the idea that we should cooperate.

"We are all created equal"

Of course, we are not. We do have -- somewhat -- equal protection under the law, but we have vastly different capabilities, most of which result from birth and the circumstances we were born into.

Is this a good meme? Consider this your homework assignment. :-)

## **TAKING CONTROL**

The quite incomplete list of memes presented above constantly and continuously try to control our lives. It is obvious that a few are good but, in my judgment, many are bad. One thing is for certain: it is not in your best interests to totally resign to the control of memes. But can you do anything about it? Certainly.

If we accept that many of our actions are meme based and therefore can be changed, we have the potential to improve our lives by evaluating the memes and overriding those that we believe have the potential to do us harm.

But how do we override memes? They are just as powerful as genes and they color our thinking, right? True, but they can be overridden by applying the rules of logic.

For example, after a person has accepted that certain memes are enhancing or interfering with their love life, that person should examine these memes under the microscope of rigid logic. Up until recently, there was a meme that said it is bad to masturbate. OK, let's take a look at it. What exactly is bad about it? Is there really any correlation between masturbation and our visual difficulties? Not that anyone can prove. Is it sinful? You will have to decide that. What else? Not much. Another meme for the trash!

One useful tool in examining sexual mores is to ask, "What do animals do?" After all, we are part of the animal kingdom and most of our sexual genes are common. But, as far as we know, the rest of the animals are not burdened (or blessed) with sexual memes. Based on this approach, you might ask, for example, is it OK for a female to have several lovers? Looks fairly common amongst the higher mammals. You take it from there. :-)

But be careful. Some rules are not universal in the animal kingdom. For example, monogamy or polygamy; which is normal? No universal agreement here.

By selecting the good memes and chunking the bad ones, you should be able to greatly enhance your life and to avoid dropping a pile of your hard earned cash at the local shrink. Your friends will appreciate it too.

## REFERENCES

[HENS]

"Memes Meta-Memes and Politics"  
By H. Keith Henson, (on the Internet)

[HALE]

Memetics: A Systems Metabiology, Version 950220, by Ron Hale-Evans, rwhe@netcom.com  
(on the Internet)

[DENN]

Dennett, Daniel C. (1991). Consciousness Explained. Boston: Little, Brown and Co.

[DAWK1]

Dawkins, Richard. (1989). The Selfish Gene. New York: Oxford University Press. Second edition

[DAWK2]

Dawkins, R. (1993b) "Viruses of the Mind" in: Free Inquiry, summer 1993, vol 13 nr 3 (on the Internet at: <<http://phenom.physics.wisc.edu/~shalizi/Dawkins/viruses-of-the-mind.html>>)

**Leon Felkins says:** *I am a retired engineer (Computer Systems) and educator. I now happily spend much of my time on the study of philosophy -- a lifelong interest but one that could not be pursued in depth while I was employed. When my fingers tire of communicating with my cyber-friends, I relax by playing with my garden and my animals here on a small farm on the edge of the Cumberland Plateau. My non-technical publications include one essay on "A Theory of Cooperation" in Free Inquiry; Spring, 1989.*

## **VIRAL MARKETING: SPREAD A COLD, CATCH A CUSTOMER**

**By  
Frank Fiore**

The Business Journal of Phoenix - June 25, 1999

They call it "viral marketing." Having someone else spread the word to drive more visitors to your site.

Let's face it. The best kind of marketing is the kind you don't have to do yourself. Especially if you're a small business on the web with a limited advertising budget. Viral marketing is like it sounds. Call it word-of-mouth, spawning, self-propagation -- organic.

Great new idea, right? Nope. Viral marketing has been around forever. Spreading the word through word-of-mouth was the world's first form of marketing. But the Internet has taken this organic from of marketing to new heights by making communications better and communities of people tighter -- thus making word-of-mouth even more effective.

When you use viral marketing as a tool, you're using the Net the way it was meant to used. Viral marketing on the Net has a long history. The first use was by Netscape. The small "Designed for NETSCAPE" icon was first used as a status symbol by webmasters to show that their site was the latest in web-page design. It didn't take Mr. Gates very long to see the power of viral marketing and soon "Designed for MS Internet Explorer" icons quickly spawned next to Netscape's.

A couple of years later, the creative marketers at Amazon took viral marketing beyond this simple link. When Amazon created the first affiliate program on the Net, they spawned a whole new generation of viral partners eager to promote Amazon's book site to their web-site visitors.

Today, they have more than 100,000 affiliates promoting Amazon and gaining a small commission on each sale directed to them.

Affiliate or associate programs are now one of the hottest viral marketing programs on the web. There are hundreds of small, medium and large companies offering their programs to any web site that wants to earn a few extra bucks.

But you don't have to have a costly or complex affiliate program to use viral marketing in your business. There are other ways to motivate your customers and visitors to do the marketing for you.

Here are just a few.

- Pass this on, will ya: The granddaddy of the new viral marketing has to be HotMail which soon was copied by Yahoo! and anyone else who had a free e-mail service. It was simple. Attach your URL to every e-mail that your users send out. Before they knew it,

HotMail had more e-mail users than the largest ISP. You can do the same. Make sure that every e-mail sent out from your company has your URL on it.

- Tell 'em Sam sent you: Reward your steady visitors for bringing new visitors to your site. Create a special "referrer program" that your steady visitors can sign up for. Have them invite their friends to visit and if they do, have them mention the referee's e-mail address and the referee earns something free from your site.
- I'll scratch your back if you scratch mine: Your site sells cigars. Their site sells humidors. A relationship made in web heaven! Look for sites that sell complementary products or services to your own. Then contact them and offer a reciprocal linking arrangement. You can go even one step further and offer their visitors a discount if they join your referee program.
- Got any good jokes lately?: I don't know how many times I've forwarded jokes or scam and virus alerts I received in my e-mail to friends and associates. So, create a funny newsletter or an e-mail alert that someone would pass on to friends.
- Take my site, please!: This is similar to "Got any good jokes lately" but in this case you give away your best assets. Don't just send out a notice about new content on your site -- send out summaries by e-mail and ask people to forward it on (with copyright and URL attached, of course). A variation on this is to let other sites reprint your content on their site, with appropriate credits and links to yours.
- And the winner is: Create an award for the best sites in your category or subject area. Then send them an e-mail informing them they won the award and where to acquire the award icon to place on their site. I do this with my Mining Co. Online Shopping site.

I award a "Best of the Net" to certain worthy sites and ask them to place the "best of" icon on their site -- which is then linked back to my site.

The one downside to viral marketing is that you're letting others do your marketing for you. Though this will save you money, your message and your brand are in the hands of someone else. There's a fine line here of spreading the word and diluting your brand.

The challenge is to exercise some control of how your message is delivered and how others perceive it. But when you master this technique, your message and your site can spread as quickly as the common cold.

Finally, don't expect a viral marketing program to pay off immediately. Like a real virus, viruses don't become epidemics until they reach critical mass. Your virus must propagate through the host population until it reaches a certain threshold of visibility and scale.

Think of it this way. Suppose a real-world virus doubles every year. In the first few years it's scarcely detectable. But within a few years after that it suddenly becomes an epidemic. You

should understand that you're playing the same game. Viral marketing takes time. So be patient, be fruitful -- and go out and multiply.

*Frank Fiore is an e-commerce consultant, author of "Dr. Livingstone's Online Shopping Safari Guidebook," the Online Shopping Guide for the Mining Co. ([onlineshopping.miningco.com](http://onlineshopping.miningco.com)) and the web producer for mySimon Inc., <http://www.mySimon.com>. He can be reached at 602-951-8757*

# CULTURAL R/K SELECTION

By  
**Agner Fog**

Ørnevej 29, 2400 Copenhagen NV, Denmark.

[agner@login.dknet.dk](mailto:agner@login.dknet.dk)

## Abstract

A new model of cultural r/k-selection is introduced. This model, which provides a classification of cultural processes based on the factors that influence memetic fitness rather than on selection mechanisms, predicts that cultural evolution will go in different ways depending on the balance between internal and external conflicts of a society. A society dominated by external conflicts or war will evolve in a direction called regal, whereas a society in a peaceful or sparsely populated area will evolve in the opposite direction, called kalyptic. The regal or kalyptic characteristics of a society influence the evolution in many areas of culture, including religion, political structure, art, music, etc.

**Key words:** r/k-selection, vicarious selection, regal, kalyptic, sociology, religion, art, music, architecture.

## 1 Introduction

The new model introduced in this article provides a way of explaining important social phenomena based on cultural selection theory. The model is defined in terms of memetic fitness determinants, and some possible mechanisms are outlined. The article finally discusses how the theory can throw new light on human cultural history and contribute to a better understanding of various cultural phenomena, including religion, ideology, art, and music.

It is assumed that the reader is familiar with memetics or cultural selection theory. A popular introduction to this paradigm is given by Lynch [19]. A more thorough description of mechanisms in cultural selection can be found in Boyd & Richerson [5] and Campbell [6]. A more sociologically oriented version of the theory is given by Schmid [30].

The theory presented here is described in more detail in Fog [14].

## 2 Fitness Determinants

The fitness of a replicator (gene or meme) often depends on several different factors. Some of these factors are important because they are responsible for a large part of the variation in fitness, while other factors are less interesting because they have only little influence on the fitness, or because they do not vary significantly within the boundaries of the system being studied. In systems which are too complex to analyze in detail it may be useful to concentrate on those factors which have the highest effect on fitness. The most important factors define what I call the main *fitness determinants*.

The concept of fitness only makes sense relative to a specified process of reproduction and selection, be it genetic or cultural, and a specific environment. It is important to recognize that fitness is a *relative* concept, depending on the selection mechanism and external conditions. Different selection conditions can lead the process in different directions, and an examination of the fitness determinants is necessary to predict the direction of evolutionary change. The failure to acknowledge this dependency has led to the often criticized unilinear theories of cultural evolution.

In order to illustrate the relativity of the fitness concept, I am going to give an almost classical example: The habit of tobacco smoking has spread to most of the world because it gives a subjective feeling of pleasure and because it is difficult for the smoker to quit when the unfortunate consequences turn up. But smoking undermines the reproductive health in many ways which reduce the probability of producing healthy children (Abel [1] and ARHP [4]). We must therefore conclude that smoking is promoted by cultural selection but counteracted by genetic selection. Tobacco smoking has a positive fitness in cultural selection but a negative fitness in genetic selection. If we pin down the process of cultural selection into partial processes, such as hedonic selection (Martindale [21]), rational selection, economic selection, etc., then we will see that the broad label of cultural selection comprises many different mechanisms each pushing in its own direction. The study of conflicts between different genetic selection mechanisms pushing in different directions has lead to important results in sociobiological theory, and a similar study is highly needed in the area of cultural processes.

A fitness determinant is not the same as a selection mechanism, but it is determined by the selection mechanism and in particular by the external conditions and selective forces working on the system. I will explain what I mean by fitness determinants by referring to the example of economic competition. Industrial enterprises may compete to produce the cheapest products of a particular quality. One possible mechanism in this process is that those factories which use the cheapest sources of energy, manpower, and other resources outcompete less efficient producers which then go bankrupt and disappear. A more efficient mechanism is that intelligent managers consciously seek the cheapest resources and production methods, thus avoiding bankruptcy. The latter mechanism is faster than the former, but they both lead the evolution in the same direction because they have the same fitness determinant: cheap production. Knowing the fitness determinants without knowing the mechanism, we may predict the direction of evolution, but not its speed.

### 3 Classification Based on Fitness Determinants

An analysis of the fitness determinants of a system can be useful for predicting the direction of evolution. This approach is particularly attractive when the system comprises so many complex, and possibly conflicting, selection mechanisms that a detailed mathematical modeling is impossible.

One example of a classification based on fitness determinants is the distinction between specific and general evolution as defined by Marshall Sahlins [28]. Specific evolution is the adaptation to a specific niche, while general evolution is the evolution of *adaptability*. Evolution of specific adaptedness may be advantageous in a constant and stable environment, while the evolution of

general adaptability or flexibility is required in an ever-changing environment. This distinction between specific and general evolution is applicable to biological as well as to cultural evolution.

Another example is the distinction between r- and K-selection in biological evolution (ALEksic [3], Takada [31], Wilson [32]). If an animal species lives under conditions where resources are ample so that there are good opportunities for expansion, but where there are also considerable dangers such as predators, then it will be advantageous for this species to use most of its resources on breeding as fast as possible and spending few resources on each offspring. This is called *r-selection*. The r is the mathematical symbol for the rate of reproduction. r-selection causes the evolution of small animals growing fast and breeding fast. Examples are mice and insects.

The opposite of r-selection is *K-selection*. This is what happens when a species lives under conditions where the population is limited by scarce resources rather than by predation. The capital K is a mathematical symbol for *carrying capacity*, i.e. the maximum number of individuals that the resources in a given habitat can continually sustain. K-selection leads to the evolution of big animals which breed slowly and utilize the given resources optimally, and which invest a considerable proportion of their resources in the care of their sparse offspring. If the animals under these conditions bred excessively, then they would have insufficient resources for nurturing each young, and they might over-exploit their habitat to the point where the resources were exhausted. K-selection is found in those animals that come last in a food-chain, such as whales, elephants, and humans.

The r/K dichotomy is useful for classifying species according to their reproductive strategy. The term *r-strategy* is used to describe animals that use most of their resources on producing as many young as possible, but do not care for their young. The opposite is a *K-strategy*, which means that the animals produce few young but spend a lot of resources on caring for and protecting their sparse offspring. We may expect a species to develop an r-strategy when reproduction rate is the dominating fitness determinant, and a K-strategy when effective utilization of limited resources is more important.

The selection mechanisms leading to these strategies may be more complicated than this simplistic model indicates. In many cases the mechanisms behind K-selection are not fully understood, although the fitness determinants are.

## 4 Cultural Selection Models

In cultural selection theory, the number of possible models may be greater than for the genetic processes, because innovation, reproduction, and selection of cultural phenomena may involve many different mechanisms and transmission media (Findlay [13], Fog [14]). All these mechanisms may interact with each other in so many complicated ways that a stringent quantitative account and classification of possible cultural processes is hardly possible, and it is even more questionable whether this would be a useful approach in applied social research.

Rather than building a taxonomy of cultural processes on selection mechanisms, I have chosen to base my classification on the social forces that give rise to selection, and the corresponding

fitness determinants. This principle is analogous to the abovementioned distinction between r- and K-selection in genetic evolution.

I am using this shortcut not only to avoid intractable mathematical problems, but also because I consider the direction of evolution more interesting than its speed - and the direction of evolution is indeed determined by fitness. This approach is attractive because it enables you to make predictions about the direction of cultural evolution based on incomplete information about the system (Fog [14]). Of course, I do not deny that other classification principles may have valuable applications.

## 5 Cultural r- and k-Selection

I want to emphasize that the analogy between genetic and cultural selection cannot be used to prove anything about either mechanism - the differences between the two mechanisms are simply too big (Boy & Richerson [5], Daly [9], Fog [14]) and the reader is warned against expecting too much from the apparent analogy between biological mechanisms and the cultural model presented below. The analogy may be useful as a source of inspiration, and should not be regarded as anything else here when I am introducing what I will call *cultural r- and k-selection*.

Cultural r-selection takes place when a group has substantial opportunities for political and cultural expansion, i.e. to defeat other groups and impose its ideology or culture on them, but at the same time has a great risk of falling victim to the expansion of other groups. In other words, the group is dominated by external conflicts and wars. By group I mean a cluster of people bound together by the feeling of a common collective identity, such as a tribe, a nation state, or a religious sect. Group membership is usually defined by religious, political, or ethnic belonging and is often symbolized by certain distinctive marks (Hogg & Abrams [15]).

Cultural r-selection results in the allocation of a high proportion of the group's resources to the fighting of external wars or conflicts or other collective dangers. The group with the strongest military force and the most effective strategy will win in the process of cultural group selection. In other words, r-selection leads to armament. This armament is not only of a technical kind, but also very much of an ideological and political nature. A strong community spirit will be fostered in connection with an ideology saying that the individual exists for the benefit of the community, that the individual should sacrifice himself for the community, where discipline and uniformity are regarded as virtues, where martyrdom is the highest honor, and where a strong central government is regarded as a sign of wealth. This kind of ideology and a corresponding political organization will make the strongest forces in political as well as ideological conflicts with neighbor groups, and will therefore have the highest cultural fitness in a situation where cultural r-selection is dominating.

The opposite of cultural r-selection is cultural *k-selection*, which takes place when a group has no opportunities for cultural expansion and is not threatened by aggression from other groups. This will typically be the case when a group is geographically isolated, for example on a solitary island, or when the cultural differences between a group and its neighbors are small compared to the internal differences within the group. The external conflicts are small or non-existent and the

only conflicts that are significant in selection processes are group-internal conflicts between leaders and subjects, between subcultures, or between individuals.

A strong military force would be a waste of resources in the absence of external conflicts. The population will not accept a despotic government that unifies and disciplines. They will rebel against powerful leaders, and the fights for freedom for everybody will be the dominating conflicts. This will lead to an ideology where society exists for the benefit of the individual, and not vice versa. There will be more freedom of choice for the individual and higher tolerance towards individual differences. The leaders will regard the life and welfare of any individual as important.

The analogy with genetic r/K-theory becomes apparent when you consider that an r-selected culture spends a high proportion of its resources on winning new territory. Or, to be more exact: The r-memes make their hosts spend many resources on winning new hosts for the same set of memes. The cultural k-strategy implies a different allocation of resources, namely on keeping the hosts you already have by making them satisfied with their society.

The fitness determinant for cultural r-selection may be characterized as military strength and political unification. It is the ability of a culture to spread to new peoples and to withstand the influence from other cultures. The fitness determinant for cultural k-selection, on the other hand, is the contentment of all individuals and thereby a minimization of conflicts between leaders and subjects. Only by satisfying the needs and wishes of all individuals as fully as possible can the culture avoid upheavals. The r-selection is determined by the reproduction of culture in space (geographic expansion), the k-selection is determined by reproduction in time (retention).

In order to avoid the impractical r- and k- terminology and to establish a distance from the flimsy analogy with genetics, I will here introduce the words *regal* and *kalyptic* to replace the symbols r and k in connection with cultural selection. The result of cultural r-selection will be termed a regal culture, and the result of cultural k-selection is called a kalyptic culture. The word regal comes from *rex*, which means king, and I have chosen this word because a dictatorship can be regarded as the prototype of a regal culture. I have formed the word kalyptic from *Kalypso*, the name of a nymph in Greek mythology, who held Odysseus captured on a desert island. This word is chosen because the most typical cultural k-selection is found on isolated islands. You may notice that the K in *genetic K-selection* is capital because the mathematical symbol it implies is so, whereas *cultural k-selection* is written with a small k because it stands for kalyptic. The concept of regal may be delineated by the following definitions:

- a regal selection is a cultural selection process dominated by inter-group conflicts or other collective dangers.
- a regal culture is the result of such a selection, *or*
- a culture which spends a high proportion of its resources on expansion or defense, *or*
- a culture which limits the freedom of the individual members and makes considerable demands on the resources of the individuals for the purpose of strengthening the group.
- a regal cultural product is a cultural phenomenon which is part of the strategy of a regal culture or otherwise a typical product of a regal culture.

The term *kalyptic* is of course defined as the opposite, i.e. a culture which is not dominated by external conflicts, which spends more resources on satisfying the individual than on strengthening the group and which attaches importance to individual freedom. The words should preferably be applied as relative graduations, rather than as absolute ideal types. It makes more sense to say that culture *X* is more regal than culture *Y*, than to just say that culture *X* is regal. I will use the term *regalization* for an evolution in the regal direction, and *kalyptization* for the opposite.

Occasionally, you may observe a correspondence between the strategic implications of genetic and cultural r/k-selection. In a regal society, the population usually spends a lot of resources on producing many children, but invest little in the individual child. In the kalyptic society, people have few children, but spend many resources on giving each child the best possible education. This is in perfect agreement with the *genetic* r- and K-strategies. This similarity between genetic and memetic strategies is due to the fact that I have classified the evolutionary systems according to fitness determinants rather than selection mechanisms. If two different evolutionary systems have similar fitness determinants, then they are likely to evolve in similar directions, no matter how different the selection mechanisms. The concordance between the two systems is far from perfect, though. If a regal culture expands by conquering new territory for its people, then this is in agreement with a genetic r-strategy. But if the regal culture wins new hosts for its memes by proselytizing, then this is an r-strategy only in the context of the cultural scheme, not the genetic.

Due to the weakness of the analogy, I prefer to regard the cultural r/k-model as an independent theory where the r and k do not have the same meaning as in the genetic model, although the similarity in nomenclature is intended.

## 6 Mechanisms in Cultural r/k-Selection

As explained above, cultural r- and k-selection may be defined by the fitness determinants or driving forces pushing the evolution in one or the other direction. The most important driving forces behind regalization are intergroup conflicts and other collective dangers, while the driving forces behind kalyptization are conflicts within a group, or to be more specific: between leaders and subordinates. However, a driving force is not the same as a mechanism. I will therefore explain some possible mechanisms behind cultural r- and k-selection.

The fundamental factor in regalization is war. A society with strict discipline and an effective controlling of the population will have higher chances of winning a war than a more soft society. The victors are likely to force those political, ideological, and religious principles on the defeated people, that made the strong government possible, and consequently those traits will spread.

It is important to understand, however, that regalization is also possible without war. The *threat* of war is sufficient. The people will soon realize that armament, physically as well as morally, is necessary to meet the threat of war, and the public will have no problems understanding that sacrifices are necessary to defend national security. The cold war and arms race between the USA and the Soviet Union was a clear example of such a reaction. This mechanism is an example of what Campbell calls *vicarious selection* [6]. The rational reaction to the war threat reduces the risk of being attacked as well as the risk of losing a war if it should come. The

cultural result is the same as if they had passively waited for the war: regularization. The vicarious selection works in the same direction as the direct selection, but faster, more effectively, and with fewer costs. Vicarious selection is therefore a very important factor in cultural evolution. Other threats to the society as a whole may also cause regularization, such as mass immigration, economic crisis and overpopulation.

The opposite process, kalyptization, is found among people living in peaceful surroundings. In the absence of external conflicts, the internal conflicts will be the dominating factors determining the direction of cultural evolution. In a competition between alternative political systems, people will prefer the most comfortable, i.e. the one that puts the fewest demands on people and gives the highest freedom and autonomy to the individual. You may call this hedonic selection (Martindale [21]). The population cannot accept a tyrannical dictatorship, and will rebel against excessive concentrations of power. In the absence of other possibilities, the population can vote with their feet: They can simply flee from the regal society to a more kalyptic one. Such an exodus is of course most effective against a small tribe, but also bigger nation states may be influenced in the kalyptic direction by the threat of mass emigration. On the other hand, the emigrants may cause a regularization of the society they invade.

Yet another selection mechanism which may lead in the kalyptic direction is economic and technological competition. A kalyptic society is usually more tolerant towards individual economic initiatives than a regal one. This kind of liberalism provides a better breeding ground for economic growth and increasing material wealth. A k-strategy also involves higher investment in education. This investment pays off in scientific and technological progress. The result of investments in enterprises and education may be that a kalyptic society in the long term will win over a more regal society in the economic competition. During the cold war, the Soviet Union was more regal than the USA, but the latter won because the economic growth and technological progress made possible a superior military technology. The result of this selection process is that American and European culture now floods the former Soviet Union, whereas very little culture diffuses the other way.

These considerations do not, however, mean that economic competition always leads to kalyptization. Economic power and political power are strongly connected, and where economic competition favors large-scale operations the concentration of economic power will also mean a concentration of political power. Much of the *de facto* power will lie in the hands of businessmen rather than democratically elected leaders.

The difference between regal and kalyptic cultures may also be explained as a difference in the *reproductive strategies* of their memes. A regal culture is a culture which utilizes a high proportion of the energy and resources of the individual members in the interest of reproducing its memes. An obvious example is a religion which commands its adherents to proselytize. The missionary work is in the interest of the reproduction of the religion, not the missionary. The strategy of a kalyptic culture is quite different. It gambles on offering its hosts as many advantages and as few burdens as possible. Such a culture spreads by means of the egoistic choice of individuals, in contrast to the regal culture which limits freedom of choice.

The word *strategy* here does not necessarily imply conscious planning. I am using the word in the same way as when biologists talk about the reproductive strategy of a primitive animal or plant having no consciousness. The reproductive strategy of a meme complex is not the same as the strategy of the humans. A cultural pattern which is able to effectively reproduce itself may have arisen by automatic selection of random innovations, or it may be the result of the intelligent planning activity of humans. The selection mechanism works whether humans understand this mechanism or not, and whether this cultural pattern is favorable to its hosts or not.

## 7 Vicarious Psychological Mechanisms

It is a well-known psychological phenomenon that external dangers to a group strengthen the solidarity within the group and create ethnocentrism and militarism. This phenomenon has been explained as well by evolutionary biologists as by social psychologists.

The biological theories emphasize the importance of group defense, building on kin selection or group selection theory (Lorenz [18], Reynolds &c. [24]).

Within social psychology, the concept of *authoritarian personality* has traditionally been used to explain ethnocentrism and fascism. The characteristics of a person with an authoritarian personality is that he desires a strongly hierarchical power structure and is willing to submit himself to strong authorities, political, ideological and religious. He fears and hates foreigners as well as deviants within his own group, and his morals in religious and sexual matters are strict (Adorno &c. [2]).

Several investigations have demonstrated that those attitudes and behaviors which are characteristic of an authoritarian personality, are promoted by factors which endanger the social order, such as war or economic crisis (Doty &c. [10], McCann & Stewin [22], Padgett & Jorgenson [23], Rosenblatt [27], Sales [29]).

Ethologists have explained the mechanism as an infantile reaction: Just like animal young seek protection by their mother when they are afraid, so do adult humans seek protection under a strong leader in case of fear, whereby they become easily indoctrinable (Eibl & Eibesfeldt [12]). This theory has not explained, however, why collective dangers result in reactions different from dangers to the individual.

No matter which intrapsychic mechanisms may be working here, we can conclude that dangers to a society lead to a psychological tendency to solidarity and strengthening of the political organization. This mechanism is highly functional because it makes the society better prepared to meet the crisis or external threats. We may see this as a kind of vicarious selection: Crises and external dangers cause a psychological armament, enabling the society to meet the dangers and possibly winning an intergroup conflict. The psychological armament by the threat of war causes the same cultural result as the war itself would: regularization - but faster and with fewer costs. This vicarious mechanism may have been created by either genetic or cultural evolution, or most likely by a combination of several selection mechanisms.

Imagine a society in surroundings where there is peace most of the time. A regal culture would be disadvantageous in times of peace because it would spend an unnecessary amount of resources on disciplining the population and maintaining an unnecessary warrior force, and also because cultural r-selection, just like genetic r-selection, entails an uncontrolled growth in population and hence exhaustion of natural resources. In a Malthusian way this may destabilize the ecological balance and lead to famine and mass extinction (Malthus [20]). On the other hand, cultural k-selection, like genetic K-selection, would stabilize the population and ensure maintenance of the means of subsistence.

A regal culture in a peaceful environment may be inexpedient, but a kalyptic culture in bellicose surroundings would be fatal. A kalyptic group will always be easy prey to the desire of a regal neighbor for expanding its territory. A group can only survive in hostile surroundings if it is regal. There is no need to limit the population - the frequent wars take care of that. On the contrary, a fast breeding population is necessary to maintain maximal military power.

The optimal solution for a group subjected to changing external influences must be *flexibility*. A fast regularization when an external danger is threatening, and fast return to a kalyptic strategy when the danger is over. The ability for fast adaptation can only be achieved by vicarious selection. You may regard this as feed-forward control. Any mechanism that leads to such an improvement in adaptability would have such a big fitness advantage, that it would be promoted by genetic as well as cultural selection. The gene/culture coevolution is estimated to have taken place through at least two million years (Durham [11]), which is more than sufficient for a mechanism like this to become fixated in our genetic and cultural heritage. The abovementioned mechanisms may be interpreted in this way, although this is admittedly not the only possible explanation for the observed psychological reactions.

## 8 Human History in Light of Cultural r/k-Selection

When humans began to cultivate the soil and raise cattle, they started a new evolution which has since influenced every aspect of human life. Previously, humans had lived as hunters and gatherers, but now different ways of living were invented. This invention was probably not selected for until an increased population density made it necessary to produce food in a more intensive manner than simply gathering the fruits of nature (Rosenberg [26]). Another possibility is that agriculture was first introduced on the demand of a powerful chief who wanted to create a basis for increasing the population of his chiefdom for strategic reasons.

The theory of regal selection plays an important role here. A war between two tribes may lead to the result that the strongest group conquers the weaker tribe and incorporates the latter under its command, so that the two tribes become united into one bigger society under a common leadership. The biggest groups - and those which are ruled by the most despotic chiefs - will be the strongest, and thus have the potential for growing even bigger. Through this selfperpetuating process, tribes and independent villages have been united into chiefdoms, chiefdoms have become states, states have become kingdoms, and finally, through an endless series of war and cruelty, enormous empires (Carneiro [8]). Agriculture has played an important role in this regal development because it has made possible an increased population density and hence a significant military superiority.

It is difficult to tell what initially set off this self-amplifying process of political integration. Is it agriculture which has given rise to a steep population growth, or is it overpopulation and famine that has necessitated the introduction of agriculture? Is it whimsical hostilities between chiefs of different tribes that has started a series of ever bigger wars and retaliations, or is it failing hunting luck that has forced a hungry population into war? Anthropologist Robert Carneiro thinks that the evolution towards ever bigger political units has started in places where small, very fertile, areas were surrounded by less attractive areas. The population has been concentrated on the most favorable areas, which made contests over the attractive territories very likely (Carneiro [7]).

The population density in infertile areas must necessarily be low, and the big distances makes war difficult or impossible. A kalyptic equilibrium may therefore be sustained in such sparsely populated regions for millennia, whereas there are ample possibilities for regal development in densely populated fertile areas. The border areas of a fertile territory particularly invite to conflict. Outside the fertile area live hunters or nomads who are attracted by the allure of the conspicuous prosperity of the agriculturalists. The peasants, in turn, are tempted by the immense, almost unused areas outside. The two groups may attempt to conquer each other's land, only to find that the captured land is unsuited for their way of life.

It is reasonable to assume that the spiral of ever increasing regality was started by an environmental factor, namely the proximity between two areas of very different fertility. The different ecologies of the two areas led to differences in way of life, political organization, and thus also a difference in social identity. The border between the two areas inevitably invited conflicts between the two quite different groups.

The concentration of the population in towns has made possible an increased specialization and division of labor, and hence the development of trade, crafts, technology, and finally industry. This development has introduced new parameters of competition in the cultural selection: food production technology, arms technology, and communication technology. Improved food production methods have enabled a more intensive utilization of natural resources and consequently a still higher population density. Improved weapons have led to military superiority. And improved means of transport and communication have made it possible to unite bigger areas under a common government.

This continued integration and regularization has taken place in Europe, Asia, and Northern Africa with few intermissions since the end of the Stone Age. But everything has a limit. In antiquity and the Middle Ages there was a limit to how big empires could be, and the limits were, first and foremost, set by the means of communication. It was difficult to control a war that took place many days' journey from the palace of the emperor, and it was difficult to motivate people to sacrifice big resources on a war that took place so far away that it seemed totally irrelevant.

When an empire has reached the limits to its growth, then regularization stops and kalyptization commences. Only a despotic government is able to keep together such a huge empire and maintain the necessary discipline and military strength. The population can hardly see the necessity of a highly tyrannical rule, so they start to rebel. When the emperor reluctantly begins to loosen his iron hand, then the internal conflicts start to flare up. The population suddenly appears to be far less homogeneous than previously believed. All those sub-groups which, one by

one, had been incorporated into the empire have preserved some of their religious or ethnic identity, and this identity is reinforced by their urge for independence and their rebellion against the despotism of the ruler. The population becomes divided and different sub-groups fight for independence. The empire starts to disintegrate and the monarch has a hard time trying to suppress the rebellious groups and keep his empire together. In the meantime, perhaps, a new kingdom nearby has started to grow. The old empire, which now has begun to disintegrate and kalypticize, is an easy victim to the expansive efforts of the new growing kingdom. The citizens do not wholeheartedly defend their country when attacked by the army of this new empire. They cannot imagine that the new ruler could possibly be more despotic and cruel than the old one, and many capitulate to the new emperor whom they regard as their liberator. In this way a new empire grows. Part of the old empire is incorporated into the new, and the rest is split up into smaller states.

History shows numerous examples of the rise and fall of mighty empires. For example, many historians have pondered over the fall of Rome, but seen in the light of the cultural r/k-theory, it is easy to explain. When an empire has reached the limits of its growth then kalyptization sets in and the empire is weakened. After a period of beginning kalyptization the realm is either conquered by a new empire or simply split up into smaller states. The recent breakdown of the Soviet empire is a proof that this history still repeats itself.

Cultural selection has been dominated by regularization since the Stone Age, reaching its zenith around the end of the nineteenth century. By then, all the continents had been colonized and further expansion possibilities were virtually exhausted. Now, lacking other possibilities, the great powers have begun to compete in conquering the outer space, but, since outer space is not habitable, this battle has only symbolic significance.

## **9 Regal and Kalyptic Cultural Products**

The results of cultural r-selection is very evident in the area of religion. The most regal societies usually develop a monotheist religion with a powerful and punishing high god. Several meme theorists have demonstrated the suitability of religions for controlling a population (Lynch [19], Richerson & Boyd [25]), and monotheist and pantheist religions are particularly effective in this respect. The image of God as a supreme ruler is a psychological projection screen which legitimizes the undivided power of a monarch. In times of kalyptization, the god becomes imagined as more merciful, and the religion puts less emphasis on extreme punishments like purgatory and hellfire.

The regal characteristics of a society are also very conspicuous in its architecture. Powerful kings and religious leaders advertise their power by the building of ostentatious palaces, cathedrals, and monuments with the most costly and profuse ornamentation (Kempers [16]). This richness of ornamentation is repeated in painting, music, and other branches of art. The preferred music style of a culture has been found to be highly correlated with the social structure (Lomax [17]), and this correlation is particularly striking when we look at the regal/kalyptic characteristics of a society. Big orchestras and choirs led by a single conductor or master, producing highly embellished music, are characteristic of a regal culture, while kalyptic cultures prefer a music where accompaniment is equally important to the melodic voice, and where the musicians appear

to be equal (Fog [14]). The preferred styles of art, music, etc. are the ones which are psychologically most congruent with the social structure.

You may imagine different cultures, subcultures, and cultural products ordered on a continuous r/k-scale spanning from the extremely regal to the extremely kalyptic. Of course, such a scale has only intuitive value. It is hardly possible to assign absolute numbers since the r/k-value is not defined by one exact criterion, but evaluated by many different criteria, most of which are more or less subjective. The purpose of introducing such a scale is not to set culture on a mathematical formula, but to give meaning to comparative statements, such as: "*Rock music is more kalyptic than hymn singing*", and thereby provide a convenient classification scheme for cultures or cultural products. Of course, not all phenomena are comparable, but a necessary condition for a comparison to make sense is that you have a yardstick, and this is what I call the cultural r/k-scale. I have observed that many aspects of culture are connected to this cultural r/k-scale (Fog [14]), although no large scale statistics have been made yet. [Table 1] is a list of characteristics which I consider typical for regal and kalyptic cultures. The list is only intended as an aid to interpreting the r/k-scale. The reader is referred to (Fog [14]) for a more detailed discussion of how the r/k-dimension is reflected in different areas of culture.

	<b>Regal</b>	<b>Kalyptic</b>
<b>Religion</b>	Monotheism, ascetic, puritan.	Animism, polytheism, atheism, fertility cult, ancestor worship.
<b>Philosophy</b>	Individuals exist for the benefit of society. Ethnocentrism, racism, material growth, expansion.	Society exists for the benefit of the individual. Individualism, tolerance, human rights, protection of natural resources.
<b>Politics</b>	Powerful central government, imperialism, uniformity, intolerance, censorship, severe punishments, witch-hunts.	Decentralized government, democracy, tolerance, peace.
<b>Art</b>	Finical, perfectionist. Patterns with strict geometry. Perpetual repetition of small details. Portrays symbols of power such as gods, rulers, war heroes, or predators.	Unrestrained, improvised. Depicts pleasure, fantasy, colors, nature, animals, fertility, individualism, rebelliousness.
<b>Music and singing</b>	Monotonous, embellished, or by offensive regality pompous. Strict rules for rhymes and foot. Choir singing, litany. Praises gods, rulers, military superiority, true love.	Bass accompaniment dominates over melodic voice. Rhythmic, rhythm varied, imaginative, often improvised. Broad repertoire of text themes.
<b>Dance</b>	Organized, restrained.	Unorganized, hilarious.
<b>Dress</b>	Decent, tidy, uniform. Strongly sex-differentiated. Reflects social status.	Creative, individual, colorful, sexy. Reflects personal taste.
<b>Architecture</b>	Churches and government buildings are grandiose, ostentatious, rich in details, with oversized gates and towers.	Functionalistic, creative, individualistic, irregular. No stylistic demonstration of social differences.
<b>Sexual behavior</b>	Strict sexual morals. Stereotypical	Liberal sexual morals. Sex

	sex roles. Sex is only for procreation. Procreation is a duty. Children are regarded as asexual and ignorant. Contraception and abortion illegal. Early marriage. High population growth.	regarded as having several purposes. Flexible, individual, pleasureful behavior. Sexual education of children. Education comes before marriage. Contraception and abortion accepted. No population growth.
<b>Occurrence</b>	Mainland with many wars and cultural contrasts. Empires. New colonies.	Small isolated societies. Peaceful regions with low population density and no cultural contrasts.

**Table 1. Typical characteristics of regal and kalyptic cultural products**

Generally, you should expect different aspects of a culture to be in reasonable agreement with regard to cultural r/k-status, but it is unlikely that you will find perfect agreement. A complex culture may have political fractions, subcultures and institutions with very different r/k-status. For example, there may be a rebellious subculture which supports more kalyptic values than the mother-culture and is aiming at influencing the mother-culture politically, or just creating space for an alternative life-style. You may also find religious sects which are more regal than their surrounding culture. Religious belief may be a more important part of personal identity than nationality to the sect members, and the regal activities mainly take the form of relentless efforts at gaining new proselytes. Other organizations may have the r/k-ideology which best suits their function, be it economic, military, correctional, or other institutions.

Another reason for lack of consonance between different aspects of a culture is simple inertia. Old art genres which fit the r/k-status that the society had several generations ago may be kept alive side by side with more modern genres. Magnificent old buildings are not torn down just because their architectural style is out of fashion.

## 10 Conclusion

The theory of cultural r/k-selection throws new light on important aspects of cultural life, including religion, ideology, art, etc. This new model has the potential for providing causal explanations to many cultural phenomena which hitherto have been difficult to explain and for suggesting causal connections between phenomena which previously have been regarded as independent. The attractiveness of this model lies in its usefulness for explaining historical developments as well as for predicting future changes, even when the knowledge of the selection mechanisms is incomplete.

The cultural r/k-theory resembles genetic r/k-theory in many ways, but it is definitely not a perfect analogy, and it should be regarded as a theory in its own right rather than an analogy. The cultural r/k-theory is only in its infancy. The possibilities of applying the methods discussed here to a diversity of cultural phenomena opens up a whole new range of topics for future research.

## References

- [1] Abel EL. 1984. *Smoking and Reproduction: An Annotated Bibliography*. Boca Raton, Florida: CRC Press.
- [2] Adorno TW, Frenkel-Brunswik E, Levinson DJ, Sanford RN. 1950. *The Authoritarian Personality*. New York: Harper and Brothers.
- [3] Aleksic I, et al 1993. *On r- and K-selection: evidence from the bean weevil (Acanthoscelides obtectus)*. Zeitschrift für zoologische Systematik und Evolutionsforschung **31** (4): 259-268.
- [4] Association of Reproductive Health Professionals 1996. *Current Issues in Smoking and Reproductive Health*. ARHP Clinical Proceedings, Oct. 1996.
- [5] Boyd R and Richerson PJ. 1985. *Culture and the Evolutionary Process*. Chicago: University of Chicago Press.
- [6] Campbell DT. 1965. *Variation and Selective Retention in SocioCultural Evolution*. In: Barringer HR, Blanksten GI and Mack RW (Eds): *Social Change in Developing Areas: A Reinterpretation of Evolutionary Theory*. Cambridge, Massachusetts: Schenkman.
- [7] Carneiro RL. 1970. *A Theory of the Origin of the State*. Science **169** (3947): 733-169.
- [8] Carneiro RL 1991. *The Nature of the Chiefdom as Revealed by Evidence from the Cauca Valley of Colombia*. In: Rambo AT, Gillogly K (Eds). *Profiles in Cultural Evolution: Papers from a Conference in Honor of Elman R. Service*. Ann Arbor, Michigan: Museum of Anthropology. (Anthropological Papers, No. 85), pp. 167-190.
- [9] Daly M 1982. *Some Caveats about Cultural Transmission Models*. Human Ecology **10** (3): 401-408.
- [10] Doty RM, Peterson BE and Winter DG 1991. *Threat and Authoritarianism in the United States, 1978-1987*. Journal of Personality and Social Psychology **61** (4): 629-640.
- [11] Durham WH 1982. *Interactions of Genetic and Cultural Evolution: Models and Examples*. Human Ecology **10** (3): 289.
- [12] Eibl-Eibesfeldt I 1989. *Human Ethology*. New York: Aldine de Gruyter. (Original: *Die Biologie des menschlichen Verhaltens*. München: Piper 1984), p.183f.
- [13] Findlay CS, Hansell RIC and Lumsden CJ 1989. *Behavioral Evolution and Biocultural Games: Oblique and Horizontal Cultural Transmission*. Journal of Theoretical Biology **137** (3): 245-269.
- [14] Fog A 1996. *Cultural Selection*. Available at URL: <http://announce.com/agner/cultsel>. (Forthcoming printed version to be published by Rowman and Littlefield).

- [15] Hogg MA and Abrams D 1988. *Social Identifications: A Social Psychology of Intergroup Relations and Group Processes*. London: Routledge.
- [16] Kempers B 1987. *Symboliek, monumentaliteit en abstractie: Het beeld van de staat en het beroep van kunstenaar na 1789*. Sociologisch Tijdschrift **14** (1) 3-61.
- [17] Lomax A 1968. *Folk Song Style and Culture*. Washington D.C.: American Association for the Advancement of Science.
- [18] Lorenz K 1966. *On Aggression*. London: Methuen & Co. (Original: *Das sogenannte Böse: Zur Naturgeschichte der Aggression*. Vienna: Dr. G. Borotha-Schoeler Verlag 1963.)
- [19] Lynch A 1996. *Thought Contagion: How Belief Spreads Through Society*. New York: Basic Books.
- [20] Malthus TRobert 1798. *An Essay on the Principle of Population, or a View of its past and present Effects on Human Happiness*. London.
- [21] Martindale C 1986. *On Hedonic Selection, Random Variation, and the Direction of Cultural Evolution*. Current Anthropology **27** (1): 50.
- [22] McCann SJH and Stewin L 1990. *Good and Bad Years: An Index of American Social, Economic, and Political Threat (1920-1986)*. Journal of Psychology **124** (6): 601-617.
- [23] Padgett VR and Jorgenson DO 1982. *Superstition and Economic Threat: Germany, 1918-1940*. Personality and Social Psychology Bulletin **8** (4): 736-741.
- [24] Reynolds V, Falger V and Vine I (Eds) 1987. *The Sociobiology of Ethnocentrism: Evolutionary Dimensions of Xenophobia, Discrimination, Racism and Nationalism*. London: Croom Helm.
- [25] Richerson PJ and Boyd R 1989. *The Role of Evolved Predispositions in Cultural Evolution*. Ethology and Sociobiology **10** (1-3) 195-219.
- [26] Rosenberg M 1990. *The Mother of Invention: Evolutionary Theory, Territoriality, and the Origins of Agriculture*. American Anthropologist **92** (2) 399-415.
- [27] Rosenblatt PC 1964. *Origins and effects of group ethnocentrism and nationalism*. Journal of Conflict Resolution **8** (2) 131-146.
- [28] Sahlins MD 1960. *Evolution: Specific and General*. In: Sahlins MD and Service ER (Eds) *Evolution and Culture*. Ann Arbor: University of Michigan Press.
- [29] Sales SM 1973. *Threat as a Factor in Authoritarianism: An Analysis of Archival Data*. Journal of Personality and Social Psychology **28** (1) 44-57.

- [30] Schmid M 1987. *Collective Action and the Selection of Rules. Some Notes on the Evolutionary Paradigm in Social Theory*. In: Schmid M and Wuketits FM (Eds) *Evolutionary theory in social science*. Dordrecht: D. Reidel, pp. 79-100.
- [31] Takada T and Nakajima H 1992: *An Analysis of Life History Evolution in Terms of the Density-Dependent Lefkovitch Matrix Model*. Mathematical Biosciences **112** (1): 155-176.
- [32] Wilson EO 1975. *Sociobiology: The new Synthesis*. Cambridge, Massachusetts: Belknap.

# **APPLYING MEMETICS TO FINANCIAL MARKETS: DO MARKETS EVOLVE TOWARDS EFFICIENCY?**

**By**

**Joshua Frank**

Economics Department

Rensselaer Polytechnic Institute, 110 8th Street

Troy, New York 12180

frankjosh@hotmail.com

## **Abstract**

In the Finance literature, the concept of evolution has been applied loosely to justify assumptions of market efficiency. However, evolution has been invoked without applying any well-defined assumptions or rigorous analysis to determine if evolution would in fact lead to the stated results. This paper attempts to undertake that analysis and to bridge the gap between the fields of Finance and Memetics. This paper finds that most Finance literature inaccurately assumes that economic fitness would be the key variable in determining which memes prevail in the evolutionary process, when in actuality other dimensions of fitness are likely to prevail. Therefore, applying memetics to financial markets does not indicate that markets would evolve towards efficiency. In fact, evolution may actually lead to less efficient financial strategies prevailing.

**Keywords:** Memetics, Financial Market Efficiency, Market Evolution, Natural Selection

## **1 Introduction**

In the field of Finance, much of the literature depends explicitly or implicitly on the concept of market efficiency. There has been extensive research on the subject of whether or not financial markets are in fact efficient; however, there has been very little thoughtful discussion as to the mechanism of this alleged market efficiency.

The theory of efficient economic markets in general and financial markets in particular actually predates the arguments for natural selection as an efficiency-generating mechanism. Market efficiency had been assumed at times as a simplifying generalization but more often it had been assumed to be an inevitable outcome based on a series of theoretical assumptions. These assumptions have included maximizing behavior both on the part of consumers (maximizing utility) and firms (maximizing profits), with actors having perfect information and no transaction costs, and other properties including non-satiation and transitivity of preferences. Working from these assumptions, economists have shown that markets will reach a state of efficiency (for example in [21] and [26]).

However, over time it has become increasingly evident that most if not all of the underlying assumptions of market efficiency are suspect. Actors do not have perfect information, there are transaction costs, and economic agents often do not act rationally even with the information they have (for example, managers often appear to "satisfice", or search for alternatives only when

performance falls below an aspiration level and only continue searching until a satisfactory option is found rather than finding best alternative [18, 24]). In addition, although maximizing of the expected value of outcomes is possible in a stochastic environment, when outcomes are uncertain, no expected value of returns can be formed, and therefore maximizing behavior is not technically possible. New evidence has been recently surfacing that stock market returns are chaotic [19] and therefore uncertain.

The argument for evolution in financial markets has come about partially to address criticisms of the underlying assumptions of market efficiency. If the original axioms behind efficient markets are not consistent with reality, then some other mechanism must exist if market efficiency is assumed to hold. Therefore, in the absence of any alternative mechanism, the theory of financial market efficiency relies on an evolutionary mechanism that is often-cited but usually stated only vaguely. Financial theorists assume that natural selection will favor strategies that are more rational and effective in investments, capital budgeting, and other financial decisions. This eventually leads to an efficient market, according to these theorists, since only the most effective strategies can survive in the marketplace.

Friedman [6] argued that irrational agents will disappear from the marketplace since they should systematically receive lower than average returns and higher than average losses. Miller [16] used the concept of evolution to reconcile the difference between financial theory and the actual decision procedures used by financial professionals within a corporation. According to Miller, "evolutionary mechanisms are at work to give survival value to those heuristics that are compatible with rational market equilibrium, however far from rational they may appear to be when examined up close and in isolation." Miller argues that it is not necessary that these efficient investors use highly sophisticated statistical and mathematical techniques to arrive at the efficient investment strategy. Evolution insures that only the heuristics with survival value will continue to exist. Zingales [28] might have stated the general position of modern financial theory best on this matter when he states "most economic theories are either implicitly or explicitly based on an evolutionary argument".

Keynes [14] offered an alternative theory for the behavior of financial markets. Keynes argued that with the development of financial markets, owners have become increasingly removed from the enterprises they hold a stake in. Therefore, "the element of real knowledge in the valuation of investment by those who own them or contemplate purchasing them has seriously declined." Keynes believes that in the absence of real knowledge, investors value financial assets based on a convention. That is, a belief that the existing state of affairs will continue indefinitely except to the extent that investors have specific reasons to believe otherwise. But this conventional valuation can become quite volatile since it has little foundation. As a consequence, "a conventional valuation which is established as the outcome of the mass psychology of a large number of ignorant individuals is liable to change violently as the result of a sudden fluctuation of opinion due to factors which do not really make much difference to the prospective yield; since there will be no strong roots of conviction to hold it steady." Keynes further argues that even experts in the market would do better by trying to understand these psychologically-based fluctuations than they would do by trying to understand the future real income streams of the underlying asset. Therefore financial markets become a place "where we devote our intelligences to anticipating what average opinion expect the average opinion to be."

This Keynesian psychologically-based model of financial markets probably explains most of the variance in market prices much better than prevailing efficient-market hypotheses. In fact, there is empirical support for the idea that market movements are not justifiable based on changes in the stream of subsequent dividends [23]. However, little attention has been given to the Keynesian viewpoint regarding financial market behavior. Probably, Friedman, Miller, or other traditional financial theorists would counter that even if this "convention" led in the short-term to irrational pricing, in the long-run selection pressures would again work to bring market pricing close to the value of the underlying asset. However, Keynes gives several valid arguments why the investor who looks at the long-term underlying value of assets will not dominate financial markets. First, it is easier to outguess the crowd than to guess long-term changes in real earnings.

As Keynes puts it; "Investment based on genuine long-term expectation is so difficult today as to be scarcely practicable. He who attempts it must surely lead much more laborious days and run greater risks than he who tries to guess better than the crowd how the crowd will behave." In addition, time is highly discounted by investors, so there is much more desire to make large gains quickly than to invest for the long-term. Furthermore, for investments made or answerable to committees, boards, or banks, strategies that are sensible long-term but that suffer short-term losses will not be tolerated. Haugen [9] discusses this last point in more detail. Modern markets are dominated by institutional investors. These institutional investors are evaluated for performance over relatively short horizons. Therefore, Haugen argues, investment market inefficiencies can persist even when the market knows about them if these inefficiencies can only be exploited over a long-term horizon.

Without the assumption of a natural selection process, there appears to be no compelling reason to assume financial markets would behave in a manner approaching efficiency in light of Keynesian arguments, empirical evidence of irrational market behavior, widespread use in the marketplace of flawed financial valuation techniques (according to traditional theory) and a wealth of psychological evidence of human biases in irrationality (for just a couple examples, see [1] or [20]).

Yet despite its widespread use in finance as an implicit mechanism for the assumption of market efficiency, the term evolution has been used in financial markets very loosely and without the analysis and scientific rigor to verify needed its validity. A good example from the field of Finance of the potential for inconsistencies in a loosely defined theory can be seen in Cox, Ingersoll & Ross [3]. The authors examine a very commonly used construct in the financial literature, the 'expectations theory' of the term structure of interest rates and find that this 'theory' is actually several independent and mutually incompatible concepts. A similar problem exists with the assumptions made in Finance regarding evolution. When examined closely, evolution in financial markets is complex and leads to results incompatible with the assumptions of financial theorists. By neglecting to examine this basic issue, the field of Finance has also failed to assimilate extensive literature that already exists in the field of evolution as it applies to human knowledge, or memetics.

Although the main arguments in this paper have not been presented before, some economists have attempted to demonstrate that natural selection will not lead to market efficiency. Hodgson [12] for example gives seven "Problems for Pangloss" (i.e., problems with the conclusion that

evolution always leads to the best possible outcomes). The problems he lists include (among others): the traits of new entrants affect the results, not just firm survival rates; evolution can be path-dependent and frequency-dependent; and there can be multiple, shifting adaptive peaks. Others presents different arguments against natural selection leading to behavior "as if" economic agents were optimizers include Nelson & Winter [17] and Heiner [10].

This paper will start out by assuming that there is an evolutionary process taking place in financial markets, particularly investment markets, and that the stable unit of evolution is the financial idea/strategy or meme. This paper will argue that these assumptions when thought through lead to conclusions inconsistent with the assumptions of financial theorists. First, it will be shown here that the economic reproduction (dollars producing more dollars) is a relatively slow and weak transmission method that will be dominated by interpersonal communication as a method of meme transmission. Then, if we add one additional assumption, that what gets communicated and assimilated most readily will at least sometimes differ from what is most economically viable (formally, economic fitness is not perfectly correlated from interpersonal fitness), this will lead to the conclusion that evolutionary competition will not lead to market efficiency in investment markets.

## 2 Applying Memetics to Financial Markets: Some Definitions

The theory of evolution was originally presented in Charles Darwin's the Origin of Species [4]. According to Darwin's theory, when there is variance in a population, the members of the population most fit to survive will have the most progeny and will therefore become more prominent in the population in future generations. Darwin called this process "natural selection". When there are new random variations continually added to a population, over many generations natural selection will lead to ongoing refinement of biological organisms and the spontaneous creation of new species. Darwin called this process "descent with modification". It was actually Herbert Spencer who popularized the term evolution for this process [12].

Dawkins [5] came up with the term meme as the smallest unit of cultural transmission, a concept parallel to that of the gene in biology. It is important to note that Dawkins' original discussion of the meme was in the context of a larger text supporting a perspective on genetic evolution that placed the gene at the center of the evolutionary stage rather than the species, group, or even the organism. Certain genes succeed, Dawkins argued, because they are successful at producing more identical genes. From this perspective, the gene can be viewed as "selfish" while the organism can be viewed as merely a vehicle for the genetic material. The same can be said of memes: the memes that succeed will be the ones best at producing more identical memes. On the surface, this appears to be a rather technical distinction between genes or memes and their host organism, but it can have important implications. For example, in biology it implies the sustainability of altruistic behavior towards kin, and it has important implications for the existence of sexual reproduction versus asexual reproduction [25] In financial markets, it implies that the memes that eventually dominate will be the ones most likely to reproduce themselves, which may not necessarily coincide with the memes that most successfully produce dollars. This point will be discussed in more detail in later sections.

In order for evolution to occur, we need to have

- A pattern of information,
- A way to generate variation in the pattern,
- A rationale for selecting variations that are adaptive, and
- A way of replicating and transmitting selected variations [7].

All four of these conditions are satisfied by investment strategies in financial markets, however there are actually two rationales for selecting variations that are adaptive (condition 3) and three methods for investment strategies to increase their presence in the marketplace, or replicate (condition 4). Examining the variations of conditions 3 and 4 are key to understanding how evolution will occur in financial markets since these variations leads to very different conclusions.

In financial and economic theory, the key basic unit of interest is the dollar. However, dollars cannot evolve because they lack in several conditions above, starting with the fact that they lack variation. Memes work as evolving units, but memes alone do not answer the questions that are of interest in finance. Financial theorists do not care if investing based on astrological signs becomes the favored method among randomly selected United States citizens, they only care about the prevalence of an investment strategy meme to the extent that there are dollars behind that strategy in the financial market. Therefore, we must take into account the dollars backing a certain investment strategy meme in the market since this will indicate the current prevalence of that meme in dictating market behavior.

To proceed then, we will define the evolving units, investment strategies, as falling within the larger category of memes. However, the states of nature that determine the efficiency of the market are defined in terms of dollar-weighted meme frequencies. Note that the dollar-weighted frequency of a meme is defined in terms of the dollars currently being invested using that meme. In other words, let us assume a specific investor has a million dollars of total investments, but that investor only invests a hundred thousand using a specific investment strategy. If we also assume that no other investor is using that strategy and that the total market is worth a hundred million dollars, that meme's current presence in the market place can be stated as having a frequency of 0.1%, not 1.0% which is the total investor's presence in the market place.

As mentioned earlier, the last two conditions required for evolution can now be seen to be satisfied in multiple ways. A financial theorist would define a meme's adaptive value or fitness based on its ability to reproduce more dollars for the individual holding the meme since investment strategies producing more dollars will tend to increase in dollar-weighted frequency in the marketplace. A memetic theorist would more likely focus on the meme's probability of interpersonal transmission when defining fitness, since the memes that are transferred from person to person more rapidly will tend to increase in dollar-weighted frequency in the marketplace. This suggests that investment strategy memes can increase dollar-weighted frequency either by reproducing dollars, or by transferring between people. There is actually one more leading to a total of three methods of replication:

- Economic Reproduction -- where the meme is not transferred interpersonally but the number of dollars associated with that meme increases through activity in the

marketplace (note that reproductive rates less than one would indicate a decline in dollars)

- Interpersonal Reproduction -- where a meme is transferred between individuals.
- Intrapersonal Transfer -- where a particular investor changes the frequency of a meme in the market by either shifting their personal investment strategy to favor one meme at the expense of another, or by changing their total investment presence in the financial market.

Only interpersonal reproduction represents true reproduction in the sense that only this method alters the frequency of the basic unit, the meme, in the human population. However, all three methods of reproduction can legitimately alter the dollar-weighted meme frequency in the market.

The concept of intrapersonal reproduction is somewhat more difficult than the other two, so some additional explanation is warranted. If an investor uses a certain financial meme to invest only part of his dollars in the market, the meme's dollar-weighted frequency can be defined in two ways: either by the number of dollars using that strategy or by the total dollars in the market by that investor. The former definition was chosen here. However, an investor can easily move funds around, increasing or decreasing the presence of a meme in the market. This does not require communication between people in any form, nor does it require that any profits are made in the market; therefore it technically is a third distinct type of reproduction. Intrapersonal reproduction is probably minor in impact, but is added here for completeness.

In reaching their conclusions regarding market efficiency, financial theorists have only taken economic reproduction into consideration. As the following sections will demonstrate, adding the other reproductive methods into the mix can lead to evolutionary outcomes quite different from the expectations of financial theorists.

### **3 Modeling the Evolution of Investment Strategy Memes**

Although financial theorists have implicitly been assuming that economic reproduction is the dominant method of transmission of investment strategy memes, it should be noted that any reasonable model of financial markets must include interpersonal transmission. Without interpersonal transmission, the only way for a meme to reach dominance in the marketplace is for a single individual to dominate the marketplace. Of course, it is absurd to think of a single individual owning the majority of, for example, the United States stock market, and this is certainly not what Miller or other financial theorists are postulating when they discuss markets evolving towards efficiency.

It is also important to note here that a meme's interpersonal reproductive fitness may have nothing at all to do with its economic reproductive fitness. Although one could certainly argue that, for example, an investment strategy that works is more likely to be passed on interpersonally, the opposite viewpoint could also be argued as convincingly. Dawkins [5] described a meme's fitness in terms of its "psychological appeal". It can be argued that it is often

the theories with the least scientific validity that have the greatest psychological appeal. Scientifically valid ideas are often complicated, filled with nuances and caveats, and often lead to conclusions we would prefer not to believe. On the other hand, an appealing but scientifically inaccurate alternative meme can have many psychological advantages. It can be simple enough to remember and transfer easily, it can have intuitive appeal, and it can prevent cognitive dissonance by leading to conclusions that the investor would prefer to believe. An example of a psychological factor that can become important in investment strategy meme transmission is the psychological construct called 'perceived control' [20]. The finding of relevance here is that it may sometimes be psychologically healthy for people to believe they have control over a situation, even if the belief is illusory. The theory of financial market efficiency implies a relatively low level of control over financial markets by investors. Therefore, it is quite possible, for example, that theories implying markets that can be outguessed and profitably exploited using past market data may have more psychological appeal than an efficient view of financial markets.

One obvious example in financial markets is technical analysis, where investors attempt to guess future prices from past trends. With the exception of a few possible anomalies, there is considerable empirical evidence indicating that investment markets follow a random walk with future performance having no correlation with past performance (see [2, 13] for evidence regarding the stock market, and [27] for the earliest evidence regarding commodity markets). Yet this type of analysis continues to be performed not only by private investors but often by large investment firms that spend considerable amounts of money hiring 'experts' to guess what a stock's price will do tomorrow based on what the price did yesterday. Part of the reason might be the concept of perceived control mentioned earlier. There are also other psychological explanations. The idea that events known to be random can be somehow predicted from prior events is a common gambler's fallacy. There may be an inherent psychological tendency that makes humans prefer to see patterns in random events.

There are also institutional influences that will tend to encourage irrational beliefs. More specifically, it is in the interests of most industry professionals for investors to believe that simple schemes, strategies, and rules can be employed to outguess the market. Industry participants who make money by generating trades benefit when investors believe that they can outguess the market and therefore make frequent trades. Full-service investment firms and financial advisors gain when investors believe that their advice allows investors to outperform the market. Mutual funds benefit when investors believe that they will outperform the market. The investment industry has an interest in promoting certain strategies, even when evidence indicates that these strategies have no basis in reality. The end result is market volatility and inefficiency.

Also, it should be kept in mind that like its biological parallel, the market for memes can be highly competitive, with many potential contenders. If there are many possible competing financial memes and only a small percentage of those memes are economically sound, chances are that the most psychologically appealing memes are not the most economically sound memes assuming the two types of fitness exhibit some statistical independence.

The number of memes representing a given investment strategy in the population of humans (as opposed to dollars) at time  $t+1$  can be defined as a product of its interpersonal reproductive fitness and the number of memes present at time  $t$ .

#### [NOTE: Equations not available]

The above equations only include the first two methods of reproduction and do not consider the rate of intrapersonal transfer. This has been done because it simplifies the analysis without reducing the validity of the conclusions. It is possible to expand and refine equations to more precisely model reality almost without limit. However, as the next section will demonstrate, the purpose here is to illustrate the simple point that economic fitness is not the primary driver of investment strategy dollar-weighted frequencies in the marketplace. Therefore, the equations have focused only on the two primary methods of transmission. Intrapersonal transfer, in any case, can only work to reduce the importance of economic fitness in the overall equation, since the initial investment strategies of new entrants into the marketplace, the strategies adopted by investors increasing their presence in the market, and switching between investment strategies will be driven by the same factors that determine interpersonal reproductive fitness rather than economic fitness.

## 4 Applying the Dollar-Meme Model using Real World Conditions

As equation 5 shows, there are many variables to consider in determining an investment strategy meme's total reproductive rate. Variables such as assimilation rates, retention rates, expression rates, and transmission rates each will follow their own rules for determining fitness which will have little or nothing to do with economic theory and a great deal to do with psychological theory. It should be noted that two of the interpersonal fitness terms,  $\alpha$  and  $\beta$  will always be less than or equal to 1 while  $\gamma$  and  $\delta$  can be much larger than 1. In fact, in this modern age of mass communication through television, books, the internet, and other mass media, the transmission rate can easily be in the millions and become the dominant variable in defining a meme's fitness. Even in the absence of the mass media and using reasonable assumptions, it can be demonstrated that interpersonal transmission will dominate economic transmission of dollar-memes. Let us take two people each with \$1,000 to invest. One uses a high return investment strategy that produces a constant rate of return of 30% while the other uses a low return strategy with a rate of return of only 5%. This would be considered a rather extreme difference in the financial world. Let us assume the low return strategy is somehow more appealing, and therefore led to 3 people with \$1,000 each to invest being told and assimilating/retaining the strategy within a year, while the high return gets transmitted to two people with \$1,000 to invest a year rather than three, then equation 1 shows the following reproductive rates:

- High profit dollar-meme:
- Low profit dollar-meme:

In this example, natural selection would eventually lead to the domination of the low profit dollar-meme, despite the fact that the economic fitness difference would be considered extreme. Although this is a somewhat arbitrary example, it is reasonable to expect that the interpersonal reproductive fitness will dominate in most real world circumstances. Economic Reproduction is

a relatively slow process. The 1.30 fitness shown above can be taken as an extremely high fitness for economic reproduction that is unlikely to be matched by even very fit strategies in the real world. Even each investor telling one other investor about a strategy at some time during the year (assuming the second investor assimilates and uses this information) leads to an interpersonal reproductive rate more than triple the economic reproductive rate. It is difficult to get precise interpersonal reproductive fitnesses for investment strategy memes, but there is clearly a potential for interpersonal reproduction to take place extremely rapidly relative to economic reproduction. Every year there are new books containing investment strategy memes that are transmitted to millions of people. Every day, investment television shows, radio shows, and print media speak to millions of people, and thousands of people receive calls from investment industry professionals. It is easy to imagine an investment strategy meme starting with a single individual and reproducing at a rate that reaches tens of thousands of investors within five years. This would imply an interpersonal reproductive rate of 6.0 or more, a rate that would make the economic fitness of a meme insignificant. In fact investment strategy memes must reproduce at very high rates at time, because entire new markets of futures and other financial instruments have developed and reached a critical mass within only a few years. These markets could not have developed so rapidly unless the memes guiding individuals to invest in these markets had not also reproduced equally rapidly.

It appears that under real-world conditions, the interpersonal fitness of a meme will dominate the economic fitness in determining the total fitness. In fact, using economic fitness alone, it is unlikely that much evolution of financial markets could have occurred due to time constraints since the environment changes more quickly than the natural selection process. Take, for example, one meme (A) that yields a return of 20% on average and another meme (B) that yields a return of 10%. If we assume that meme A is a new variation that initially exists in 1% of the population, we can determine how long it will take this new economically superior strategy to dominate 99% of the population. If we define the frequency of meme A as  $p$ , and the frequency of meme B as  $q$ , then, using an equation developed by Haldane [8]:

(6) [Equation Not Available]

where the 0 subscript represents the starting population frequency and the subscript e represents the ending population frequency, and the fitness,  $r$ , is defined as:

(7) [Equation Not Available]

or in this case:

This amount of time is older than most current financial markets. This is also much longer than the time it takes for the market environment to change dramatically. For example, the Capital Asset Pricing Model on which much of modern portfolio theory is based has only been in existence for about 35 years [15, 22] and the amount of time betas have been used extensively by investors as a measure of risk is much less than this. Therefore, it is unlikely that economic fitness alone would lead to an efficient natural selection process since the structure of the financial market changes more quickly than investment strategies change in frequency (i.e. using only economic fitness, by the time a successful meme becomes dominant, the environment for which it is successful no longer exists).

In addition to large scale structural changes in the market that occur over time, there are also cyclical effects to consider. For example, an investor who chooses stocks that move strongly with changes in the total market (a strategy of choosing high beta stocks) would do quite well in a period of rapid economic growth, while the opposite would be true in a recession. In general, the strategies that work best in bull markets are quite different than the strategies that work in bear markets. Besides the obvious general economic and market cycles, there are other cyclical changes in the environment that will affect investment strategy success including inflation rates, interest rates, the regulatory environment, and energy prices. Strategies that are the most successful in one environment will tend to reproduce rapidly at the expense of strategies that hold a moderate level of success but that work in all environments. These temporarily successful strategies may come to dominate the market only to meet with disaster when conditions change. The result will be inefficient and over reactive financial markets.

## 5 Summary and Conclusions

If we assume evolution takes place in financial markets and that the investment meme is the unit of selection, under real-world conditions, interpersonal reproduction will dominate economic reproduction as the primary reproductive method of investment strategy memes. Fitness and natural selection in financial markets will follow the theories and models developed in the field of Memetics rather than in Finance or Economics. If we further assume that economic and interpersonal fitnesses differ, the growth rate of an investment strategy meme may have little to no correlation with its economic soundness. This leads to a conclusion quite different from that previously assumed by financial theorists. Financial evolution will probably not lead to efficient markets and more likely would result in the dominance of inefficient market strategies.

Although financial markets and particularly investment markets have been used as the primary example in this article, the same assumptions regarding evolution and efficiency have often been made in general economic theory, and much of the discussion in this article would also apply to the field of Economics.

## References

- [1] Brehmer, B. and Joyce, C. R. B. (1988). *Human Judgment: The SJT View*. North Holland, Amsterdam.
- [2] Cootner, P. H. (1964). *The Random Character of Stock Market Prices*. MIT Press, Cambridge, Massachusetts.
- [3] Cox, J., Ingersoll, J., and Ross S. (1981). A Re-examination of Traditional Hypotheses about the Term-Structure of Interest Rates, *Journal of Finance*, **36**:769-799.
- [4] Darwin, C. R. (1859). *On the Origin of Species*. Murray, London.
- [5] Dawkins, R. I. (1976). *The Selfish Gene*. Oxford University Press, New York.

- [6] Friedman, M. (1953). The methodology of positive economics, Essays in Positive Economics. University of Chicago Press, Chicago.
- [7] Gabora, L. (1996). Culture, Evolution, and Computation, 2nd On-line Workshop on Evolutionary Computation, <http://www.bioele.nuee.nagoya-u.ac.jp/wec2/papers/files/gabora.txt>.
- [8] Haldane, J. B. (1957). The cost of natural selection, Journal of Genetics, **55**:511-524.
- [9] Haugen, R. A. (1995). The New Finance: The Case Against Efficient Markets. Prentice Hall, New Jersey.
- [10] Heiner, R. A. (1983). The origin of predictable behavior, American Economic Review, **73**:560-595.
- [11] Heylighen, F. (1998). What makes a meme successful? Selection criteria for cultural evolution, 15th International Congress on Cybernetics, <http://pespmc1.vub.ac.be/Papers/MemeticsNamur.html>.
- [12] Hodgson, G. (1993). Economics and Evolution. The University of Michigan Press, Ann Arbor, MI.
- [13] Kendall, M. G. (1953). The analysis of economic time-series, Part I. Prices, Journal of the Royal Statistical Society, **96**:11-25.
- [14] Keynes, J. (1936). The General Theory of Employment, Interest and Money. Harcourt, Brace and Company, New York.
- [15] Lintner, J. (1965). The valuation of risk assets and the selection of risky investments in stock portfolios and capital budgets, Review of Economics and Statistics, **47**:13-37.
- [16] Miller, M. (1977). Debt and Taxes, Journal of Finance, **32**:261-275.
- [17] Nelson, R. & Winter, S. (1982). An Evolutionary Theory of Economic Change. Harvard University Press, Cambridge, MA.
- [18] Newell, A. & Simon, H. (1972). Human Problem Solving. Prentice-Hall, New Jersey.
- [19] Peters, E. E. (1994). Fractal Market Analysis: Applying Chaos Theory to Investment and Economics. John Wiley & Sons, New York.
- [20] Rotter, J. B. (1966). Generalized expectancies for internal versus external control of reinforcement, Psychological Monographs, **80**(1):1-28.
- [21] Samuelson, P. A. (1971). Foundations of Economic Analysis. Atheneum, New York.

- [22] Sharpe, W. F. (1964). Capital Asset Prices: A theory of market equilibrium under conditions of risk, *Journal of Finance*, **19**:425-442.
- [23] Shiller, R.J. (1981). Do stock prices move too much to be justified by subsequent changes in dividends? *American Economic Review*, **71**:421-436.
- [24] Simon, H. A. (1959). Theories of decision-making in economics and behavioral science, *American Economic Review*, **49**:253-284.
- [25] Trivers, R. (1985). Social Evolution. The Benjamin/Cummings Publishing Company, Menlo Park.
- [26] Weintraub, E. R. (1979). Microfoundations: The Compatibility of Microeconomics and Macroeconomics. Cambridge University Press, Cambridge.
- [27] Working, H. (1934) A Random Difference Series for Use in the Analysis of Time Series, *Journal of the American Statistical Association*, **29**:11-24.
- [28] Zingales, L. (1998). Survival of the Fittest or the Fattest? Exit and Financing in the Trucking Industry, *Journal of Finance*, **53**:905-938.

# MEME AND VARIATIONS: A COMPUTATIONAL MODEL OF CULTURAL EVOLUTION

By

Liane M. Gabora

Department of Biology, UCLA  
Los Angeles, CA 90024-1606  
email: liane@cs.ucla.edu

Published in (L. Nadel and D. L. Stein, eds.) 1993 *Lectures in Complex Systems*, Addison Wesley, 1995

## 1. INTRODUCTION

In order for a pattern of information to evolve, we need a way of generating variations of the pattern, and a way of selectively replicating it. In *Adaptation in Natural and Artificial Systems* [11], Holland introduced a computational model of how these processes are carried out on patterns of genetic information in biological systems. In this paper we introduce a computational model of how they are carried out on patterns of cultural information in a society of interacting individuals.

Ideas, like genomes, are patterns that evolve; however their evolution is not subject to the same constraints, and employs different mechanisms. The generation of variation is less random in cultural evolution than in biological evolution; it reflects the accumulated knowledge of individuals, and the social structure in which they are embedded. In cultural evolution, selective replication is Lamarkian; an idea can be modified through experience after it has been phenotypically expressed. Genetic information is coded in a physical sequence of nucleotides, and contains the instructions for its own replication, whereas ideas are coded in patterns of neuron activation, and do not contain the instructions for their replication; they rely on human hosts to replicate them. Ideas that satisfy our needs or drives are preferentially learned and implemented. Thus the fitness landscape for the evolution of ideas is molded by our drives. By developing a minimal model that incorporates mechanisms which differentiate cultural evolution from biological evolution, we hope to broaden our understanding of how it is that something can evolve.

In *The Selfish Gene* [8], Dawkins coined the word "meme" to refer to "ideas, catch-phrases, clothes-fashions, ways of making pots or building arches." He elaborates: "Just as genes propagate themselves in the gene pool by leaping from body to body via sperm or eggs, so memes propagate themselves in the meme pool by leaping from brain to brain." Our model is called Meme and Variations because it is based on the premise that new ideas are variations of old ones; they result from tweaking or blending existing ideas [12]. We will explore the impact of three phenomena that are unique to cultural evolution. The first is knowledge-based operators: brains detect regularity and build schemas that they use to adapt the mental equivalents of mutation and recombination to their meme substrates. The second is imitation: ideas spread when members of a society observe and copy one another. This is seen in both animal and human societies [4, 15, 16]. Imitation enables individuals to share complete or partial solutions to the

problems they face. The third is mental simulation: individuals can imagine what would happen if a meme were implemented before resources are spent on it. This provides them with a rudimentary form of selection before the phenotypic expression of a meme.

We will adopt the terminology of biology. Each component of a meme is referred to as a locus, and alternative forms of a locus are referred to as alleles. The processes that generate variation -- the cultural counterparts to mutation and recombination -- are referred to as operators. Forward mutation is mutation away from the initial or wild type allele, and backmutation is mutation from an alternative form back to wild type. Changes in the relative frequencies of different alleles due to random sampling processes in a finite population are referred to as drift [17]. The set of all existing patterns is referred to as the population. Following Braitenberg [6], an individual is referred to as a vehicle, and the set of all vehicles is referred to as the society.

## 2. THE MODEL

### 2.1 THE DOMAIN

Donald [9] has provided substantial evidence that the earliest form of culture was mimetic display. The memes in Meme and Variations represent mating displays. A meme has six loci that correspond to six body parts: left forelimb, right forelimb, left hindlimb, right hindlimb, head, and tail. Each locus has a floating point activation between -0.5 and 0.5 which determines the amount of movement (angle of rotation from rest position) of the corresponding body part when the meme is implemented. A value of 0.0 corresponds to rest position; values above 0.0 correspond to upward movement, and values below 0.0 correspond to downward movement. Floating point loci activations produce graded limb movement. However for the purpose of mutation, loci are treated as if there are only three possible alleles at each locus: stationary, up, and down. Six loci with three possible alleles each gives a total of 729 possible memes.

### 2.2 THE NEURAL NETWORK

The neural network is an autoassociator; it learns the identity function between input and output patterns. It has six input/output units numbered 1 through 6, corresponding to the six body parts. It has six hidden units numbered 7 through 12, corresponding to the general concepts, "forelimbs", "hind limbs", "left", "right", "movement", and "symmetry" (Figure 1).

*Figure 1 not available*

Figure 1. The neural network. Arrows represent connections with positive weights. For clarity, negative connections are not shown. Connections to the symmetry unit are also not shown. Hidden units are linked with positive weights to input/output units that are positive instances of the concepts they represent, and linked with negative weights to input/output units that represent negative instances of the memes they represent (thus "left forelimb" is positively linked to "left" and negatively linked to "right"). Hidden units that represent opposite concepts have negative connections between them. The hidden units enable the network to encode the semantic structure of a meme, and their activations are used to bias the generation of variation (Section 2.5).

The neural network starts with small random weights between input/output nodes. Weights between hidden nodes, and weights between hidden nodes and input/output nodes, are fixed at +/- 1.0. Memes are learned by training for 50 iterations using the generalized delta rule with a sigmoid activation function [10]. The relevant variables are:

```

aj = activation of unit j
tj = the jth component of the target pattern (the external input)
wij = weight on line from unit i to unit j
b = 0.15
q = 0.5
aj = 1 / (1 + e-[ b sum wji ai + q ] )

```

For the movement node, we use the absolute value of  $a_i$ . The error signal,  $d_j$ , is calculated as follows. For input/output units:

```
dj = (tj - aj) aj (1 - aj)
```

For hidden units:

```
di = aj (1 - aj) sum djwij
```

## 2.3 THE EMBODIMENT

The embodiment is a six-digit array that specifies the behavior of the six body parts. While the output of the neural network represents what the vehicle is thinking about, the embodiment represents what the vehicle is doing. A meme cannot be observed and imitated by other vehicles until has been copied from the neural network to the embodiment.

## 2.4 THE FITNESS FUNCTION

An optimal meme is one in which all body parts except the head are moving, and limb movement is symmetrical. (Thus if the left forelimb is moving up, the right forelimb is moving down, and vice versa.) This is implemented as follows:

```

F = fitness
c = 2.5
am = activation of movement hidden node
as = activation of symmetry hidden node
ah = activation of head node
i = } 1 if ah = 0.0; 0 otherwise
F = cam + 2cas + ci

```

This fitness function corresponds to a relatively realistic display, but it also has some interesting properties. A vehicle that develops the general rule "movement improves fitness" risks overgeneralization since head stability contributes as much to fitness as movement at every other limb. This creates a situation that is the cultural analog of overdominance in genetics; the optimal value of this locus lies midway between the two extremes. We also have a situation analogous to bidirectional selection or underdominance; the optimal value of the tail locus lies at either of the two extremes. There is epistasis; the value of what one limb is doing depends on what its left-right counterpart is doing. Finally, since there is one optimal allele for the head, two optimal alleles for the tail, two optimal forelimb combinations, and two optimal hind limb combinations, we have a total of eight different optimal memes. This enables us to perform a comparative

analysis of diversity under different ratios of creation to imitation. Note that there are nonoptimal local maxima, corresponding to memes in which limbs are moving, but in the same direction.

## 2.5 THE MEMETIC ALGORITHM

The basic idea of the memetic algorithm is to translate knowledge acquired through the process of evaluating a new meme into educated guesses about what increases meme fitness. Each locus starts out with the allele for no movement, and with an equal probability of mutating to each of the other two alleles (the alleles for upward and downward movement). A new meme is not learned unless it is fitter than the currently-implemented meme, so we use the difference between these two memes to bias the direction of mutation. Two rules of thumb are used.

The first rule is: if the fitter meme codes for more movement, increase the probability of forward mutation and decrease the probability of back mutation. Do the opposite if the fitter meme codes for less movement. This rule of thumb is based on the assumption that movement can general (regardless of which particular body part is moving) can be beneficial or detrimental. This seems like a useful generalization since movement of any body part uses energy and increases the likelihood of being detected. It is implemented as follows:

```
am1 = activation of movement unit for currently-implemented meme  
am2 = activation of movement unit for new meme  
p(fmut)i = probability of forward mutation at allele i (increased movement)  
p(bmut)i = probability of backward mutation at allele i (decreased movement)  
  
IF (am2 > am1)  
THEN p(fmut)i = MAX(1.0, p(fmut)i + 0.1)  
ELSE IF (am2 < am1)  
THEN p(fmut)i = MIN(0.0, p(fmut)i - 0.1)  
p(bmut)i = 1 - p(fmut)i
```

The second rule of thumb biases the vehicle either toward or away from symmetrical limb movement. It has two parts. First, if in the fitter meme both members of one pair of limbs are moving either up or down, increase the probability that you will do the same with the other pair of limbs. Second, if in the fitter meme, one member of a pair of limbs is moving in one direction and its counterpart is moving in the opposite direction, increase the probability that you will do the same with the other pair of limbs. This generalization is also biologically useful, since many beneficial behaviors (walking, etcetera) entail movement of limbs in opposite directions, while others (galloping, etcetera) entail movement of limbs in the same direction. Space constraints do not permit a detailed explanation of the implementation of this rule; however, it is similar to the implementation of the first rule.

In summary, each meme is associated with a measure of its effectiveness, and generalizations about what seems to work and what does not are translated into guidelines that specify the behavior of the memetic algorithm.

## 3. PROTOCOL FOR EVOLVING MEMES

Vehicles are in a two-dimensional wrap-around 10x10 grid-cell world, one vehicle per cell. Each iteration, every vehicle has the opportunity to (1) acquire a new meme through creation or imitation, (2) update the mutation operator, and (3) implement the new meme.

Vehicles have an equal probability of creating and imitating. To create a new meme, the memetic algorithm is applied to the meme currently represented on the input/output layer of the neural network. For each locus, a vehicle decides whether mutation will take place. The probability of mutation is specified globally at the beginning of a run. If it decides to mutate, the direction of mutation is stochastically determined. If the new meme has a higher fitness than the currently-implemented meme, the vehicle learns and implements the new meme.

To acquire a meme through imitation, a vehicle randomly chooses one of its eight neighbors at random and evaluates the fitness of the meme the neighbor is implementing. If its own meme is fitter than that of the neighbor, it chooses another neighbor, until it has either observed all eight neighbors or found one with a fitter meme. If no fitter meme is found, the vehicle does nothing. Otherwise, the neighbors' meme is copied to the input/output layer of the vehicle's neural network, and learned.

Since in both creation and imitation a new meme is not acquired unless it is fitter than the currently-implemented meme, the new meme provides information that is used by the memetic algorithm. For example, since we arbitrarily chose a fitness function in which movement is generally beneficial, if the new meme codes for more movement than the old meme, the probability of forward mutation will almost always increase.

In the no mental simulation condition, whether the new meme was acquired through creation or imitation, it must be implemented for at least one iteration before its fitness can be assessed. In this case mutation operators are updated the following iteration.

The new meme is copied from the neural network to the embodiment. The vehicle is now implementing the new meme.

## 4. RESULTS

The following experiments were conducted using a mutation rate of 0.17 per locus, a 1:1 creation to imitation ratio, and all cultural evolution strategies operative, unless otherwise indicated.

### 4.1 OUTLINE OF A RUN

Initially all vehicles are immobile. The immobility meme quickly mutates to a new meme that codes for movement of a body part. The new meme has a higher fitness and is preferentially implemented. Diversity increases as memes continue to mutate and spread through imitation. Diversity peaks when the first maximally-fit meme is found, and decreases as the society converges on maximally-fit memes (Figure 2). Stabilization takes longer for epistatically-linked loci than over- or underdominant loci (Figure 3). The best performance is obtained with a high mutation rate: between 0.07 and 0.22 mutations per locus, or approximately one mutation per meme (Figure 4). This is probably because mental simulation ensures that poor memes are not

implemented, and good memes are imitated by others, so they are unlikely to be lost through mutation from the society as a whole.

Since we have eight optimal memes, there are many stable configurations for the distribution of memes. Drift is observed among equally-fit alleles (Figure 3), as predicted by Cavalli-Sforza and Feldman's mathematical model [7].

## 4.2 EXPERIMENTS

The three cultural evolution strategies -- mental simulation, imitation, and knowledge-based operators -- were made inoperative one at a time to determine to their contribution to optimization speed and peak mean fitness. All three increase the rate of optimization, and mental simulation and imitation also increase peak mean fitness (Figures 5, 6, and 7).

The fitness of the fittest meme (Figure 8) increases as a function of the ratio of creation to imitation; however, the highest mean fitness is achieved when both creation and imitation are employed, in a ratio of approximately 2:1 (Figure 9). Since the vehicles with the fittest memes gain nothing by imitating others, there is a trade-off between average meme fitness and fitness of the fittest meme. Interestingly, meme diversity also varies with the ratio of creation to imitation (Figure 2), ranging from 1-2 memes when  $p(\text{create}) = 0.25$  to 10-11 memes when  $p(\text{create}) = 1.0$ . When  $p(\text{create}) = 0.75$ , the society converges on 7-8 memes; it finds all (or nearly all) of the fittest memes. A nice balance is struck between the diversifying effect of mutation and the converging effect of imitation. In future experiments in which the fitness landscape will fluctuate, maintaining diversity may prove to be more important than speed.

## 5. FUTURE PLANS

This program will soon run on the connection machine, which will allow us to increase the size of the artificial society to several thousand individuals. We will examine the effect of erecting complete or semi permeable barriers between different societies, and the effect of migration.

A salient feature of human behavior is self-tuning; individuals modify their behavior according to how well it satisfies their needs or drives. Drives amount to conceptual niches that guide the evolution of culture. In future experiments vehicles will have two drives: to mate, and to acquire territory. The implementation of a meme will produce a response in neighboring vehicles that satisfies one drive or the other. When a drive is satisfied by implementing a display, its strength decreases. Fitness functions are not built into vehicles; they vary with the relative strengths of the drives. Memes are expected to specialize for one drive or the other; evolving along different trajectories toward two different basins of attraction.

Vehicles will be able to monitor their success with creation and imitation, and adjust their creation/imitation ratio accordingly. Individual differences will be introduced. Those that have a flawed memetic algorithm might specialize in imitation, while those that can not correctly translate the behavior of a neighbor into input to their neural network might specialize in creation. When vehicles can recognize one another and associate each other with the fitness of

their memes, a hierarchical social structure could emerge in which some vehicles are ignored while others are imitated by many.

Another plan is to add recombination: vehicles will acquire new memes by combining a meme that is being imitated with a previously-learned meme. This will allow them to specialize on different parts of a meme, and then share partial solutions. Vehicles will be able to monitor the relative effectiveness of mutation and recombination throughout a run, and adjust their frequencies accordingly.

## 6. COMPARISON WITH OTHER APPROACHES

This work differs from anthropological approaches to culture in that the goal is not to put together a detailed picture of how human culture evolved, but to abstract a general model of cultural evolution. Cavalli-Sforza and Feldman [7] have developed a mathematical theory of cultural evolution which has been extended by Lumsden and Wilson [14] and Boyd and Richerson [5]. The computational approach taken here allows us to model not only what happens, but the mechanisms that make it happen. It enables us to look for patterns that arise when these mechanisms are carried out in parallel in a society of interacting individuals, and to use relatively complex memes with more than one or two alleles.

Some interesting work has been done using a genetic algorithm to investigate the evolution of cooperation [2], and the interaction between genetic evolution, learning, and culture [3]. Ackley found that Lamarckian evolution increases the efficiency of a genetic algorithm [1]. Hutchins and Hazelhurst used a computer model to explore the relationship between environment, internal representation of the environment, and cultural artifacts that mediate the learning of environmental regularity [13]. These studies model vertical (intergenerational) transmission. Meme and Variations differs from these approaches in that we model horizontal (intra-generational) transmission, and we look at the dynamics that emerges when a society of agents can each invent their own ideas and imitate others.

## ACKNOWLEDGEMENTS

I would like to thank Dave Chalmers and David Jefferson for useful comments on a draft of this manuscript, and Robert Boyd, Michael Dyer, and Charles Taylor for discussion. Thanks also to the Center for the Study of the Evolution and Origin of Life (CSEOL) at UCLA for inspiration and support.

## REFERENCES

Ackley, D. A Case for Distributed Lamarkian Evolution. Presentation at the Third Artificial Life Conference, Santa Fe NM, 1992.

Axelrod, R. Modeling the Evolution of Norms. Paper presented at the American Political Science Association Meeting, New Orleans, 1985.

- Belew, R. K. Evolution, Learning, and Culture: Computational Metaphors for Adaptive Algorithms. *Complex Systems* 4 (1990): 11-49.
- Bonner, J. T. The Evolution of Culture in Animals. Princeton: Princeton University Press, 1980.
- Boyd, R. and P. J. Richerson. Culture and the Evolutionary Process. Chicago: The University of Chicago Press, 1985.
- Braitenberg, V. Vehicles. Cambridge: MIT Press, 1984.
- Cavalli-Sforza, L. L. and M. W. Feldman. Cultural Transmission and Evolution: A Quantitative Approach. Princeton: Princeton University Press, 1981.
- Dawkins, R. The Selfish Gene. Oxford: Oxford University Press, 1976.
- Donald, M. Origins of the Modern Mind. Cambridge: Harvard University Press, 1991.
- Hinton, G. E. Implementing Semantic Networks in Parallel Hardware. In Parallel Models of Associative Memory. Hillsdale NJ: Erlbaum Press, 1981.
- Holland, J. K. Adaptation in Natural and Artificial Systems. Ann Arbor: University of Michigan Press, 1975.
- Holland, J. H., K. J. Holyoak, R. E. Nisbett, and P. R. Thagard. Induction. Cambridge: MIT Press, 1986.
- Hutchins, E. and B. Hazelhurst. Learning in the Cultural Process. In Artificial Life II. Redwood City, CA: Addison-Wesley, 1992.
- Lumsden, C. and E. O. Wilson. Genes, Mind, and Culture. Cambridge: Harvard University Press, 1981.
- Robert, M. Observational Learning in Fish, Birds, and Mammals: A Classified Bibliography Spanning over 100 Years of Research. *Psych Record*, 40 (1990): 289-311.
- Smith, W. J. The Behavior of Communicating. Cambridge: Harvard University Press, 1977.
- Wright, S. Evolution and the Genetics of Populations. Chicago: University of Chicago Press, 1969.

# A DAY IN THE LIFE OF A MEME

By

Liane Gabora

Center for the Study of the Evolution and Origin of Life  
University of California, Los Angeles  
Los Angeles, CA 90095-1567  
liane@cs.ucla.edu

To appear in *The Nature, Representation, and Evolution of Concepts*  
(Ed. Philip van Loocke) 1997, Routledge Press.

## 1. CONCEPTUAL CHANGE: A SECOND FORM OF EVOLUTION

Knowing everything there is to know about how concepts or ideas can get represented in the mind of an individual would not take us far toward an account of why individuals understand the world the way they do, or explain why their understanding differs from that of their ancestors. While some ideas fade into obscurity, others rise to prominence and spread horizontally through society and vertically from one generation to another, getting progressively refined and embellished along the way. Ideas, like the strands of DNA that encode instructions for building and maintaining living organisms, seem to undergo a process analogous to biological evolution. Understanding this process is a vital part of unraveling our cognitive make-up.

There has been a slow but steady effort to map the concept of evolution onto the dynamics of culture. Popper [1963] and Campbell [1987] alerted us to the evolutionary flavor of epistemology. Dawkins [1976] introduced the notion of a meme--a replicator of cultural information analogous to the gene. In his words: "Just as genes propagate themselves in the gene pool by leaping from body to body via sperm or eggs, so do memes propagate themselves in the meme pool by leaping from brain to brain." Others have drawn from mathematical models of population genetics and epidemiology to model the spread of ideas [Cavalli-Sforza & Feldman 1981; Lumsden & Wilson 1981; Schuster & Sigmund 1983; Boyd & Richerson 1985; Hofbauer & Sigmund 1988].

These works point toward the possibility that memetics constitutes a second form of evolution, distinct from yet intertwined with biological evolution, with the potential to provide the kind of overarching framework for the social and cognitive sciences that the first form provides for the biological sciences. However, thus far memetics has not lived up to this potential, a situation that seems unfortunate given the success of the biological precedent. Although much was known about living things before Darwin, his theory of how life evolves through natural selection united previously disparate phenomena and paved the way for further biological inquiry.

Some believe that looking to biological evolution to gain insight into cultural evolution is a waste of time. As Gould [1991] put it: "Biological evolution is a bad analog for cultural change... biological evolution is powered by natural selection, cultural evolution by a different set of principles that I understand but dimly." However at an abstract level of analysis they amount to the same thing: exploration and transformation of an information space through

variation, selection, and replication. Both present the question of what underlying mechanisms could & launch a self-perpetuating adaptive process. Thus the possibility that the two have enough in common that the former can pave the way for the latter seems at least worth pursuing.

Skeptics may wonder how we can develop a theory of cultural evolution before we understand how memes are instantiated in the brain. This situation has a precedent: Darwin came up with the theory of biological evolution through natural selection before the discovery of genes. It turned out that genes are laid out in a fairly straightforward way in physical space, which does not appear to be the case with memes. This does not mean they can't evolve, so long as there is a way of retrieving the components of a meme so they can work together as a unit. We may not yet know all the physiological details of how the information manifested in, say, a handshake between two individuals--with its unique arrangement of contact points, applied forces, and trajectory--can be traced back to these individuals' mental representations of handshakes, each other, and the situation they are in. But every day more pieces of the puzzle fall into place, and we will proceed with the confidence that the rest of the puzzle pieces exist and can be found.

This chapter outlines a theory of how memes evolve, and illustrates how the memetic perspective provides not only not only a foundation for research into the dynamics of concepts and artifacts at the societal level, but a synthetic framework for understanding how representations are generated, organized, stored, retrieved, and expressed at the level of the individual. It also sketches a tentative theory of how an infant develops a sustained train of thought and thus becomes a cog in the meme-evolving machinery.

Implications of this theory pertaining to the mechanisms underlying creativity, and why it is virtually nonexistent in other species, are explored. It concludes with discussion of how a cultural-evolution perspective can shape and inspire research in the cognitive and social sciences.

## 1.1 COMPONENTS OF AN EVOLUTIONARY SYSTEM

In order for evolution to happen there must be:

1. A pattern of information (a state within a space of possible states).
2. A way to generate variations of the pattern (explore or transform the space).
3. A rationale for selecting variations that are adaptive-- tend to give better performance than their predecessors in the context of some problem or set of constraints (a fitness landscape applied to the space).
4. A way of replicating and transmitting (or amplifying, as molecular biologists refer to it) the selected variations.

In biological evolution, the evolving patterns of information are genes encoded as sequences of nucleotides. Variations arise through mutation and recombination, and natural selection weeds out those that are maladaptive. Replication takes place at the level of the genotype. In cultural evolution, the evolving patterns of information are memes--mental representations of ideas,

behaviors, or other theoretical or imagined constructs, perhaps encoded as patterns of neuron activation. Variation arises when representations are combined, transformed, or reorganized. Replication is phenotypically mediated; it occurs when representations are transformed into action or language, transmitted through processes such as imitation, and reproduced, more or less, in another brain. Incorporation of these new information patterns into the society alters the selective pressures and constraints exerted by the social environment, which in turn leads to the generation of yet more patterns. Thus memes, like DNA, comprise a self-sustained system for the relentless exploration and transformation of a space of possible patterns.

## **1.2 COGNITION IS NOT JUST A PREDICTABLE EXTENSION OF BIOLOGICAL EVOLUTION**

The line of reasoning presented here can be succinctly conveyed in terms of information, which is related to the number of differences required to specify the state of a system [Shannon & Weaver 1963; Bateson 1972]. States have not only a structure of difference relations between them, but also a combinatorial structure--each state can itself comprise an information space--so that complex information can be built up from simple information. We can rate each state in a space of possible states against some performance measure or fitness criterion and the result is referred to as a fitness landscape, and we can move from one state to another by way of a computation. The world can be viewed as a vast network of computations wherein information is created, transformed, and destroyed. This information often exhibits pattern, or statistical regularity that can be expressed mathematically.

After seeing many shadows cast by the same object, we can develop an internal model of what that object looks like without having seen it. If more than one object is casting shadows we can learn to tell which object is casting any particular shadow.

Similarly, by viewing every pattern we encounter as a shadow or footprint of one or more broad causal principles [note 1], we can gain insight into the causal principles that manifest pattern.

If you were to go back to sometime during the first billion years of Earth's history, the only causal principle you would need to invoke to explain pattern in the information present (with the exception of yourself) would be the physical constraints and self-organizing properties of matter.

If you were to go back to sometime after the origin of life, approximately three billion years ago, this would no longer be the case. Not that life doesn't exhibit the properties of matter. But it would be virtually impossible for, say, a giraffe to appear in an information space not acted upon by natural selection. Another causal principle--biological evolution--would have to be invoked from this point on.

Today the Earth is embedded with artifacts like computer networks and circuses that cannot be accounted for by appeal to either the properties of matter or biological evolution. That is, biological evolution does not provide us with adequate explanatory power to account for the existence of computers any more than the properties of matter can explain the existence of giraffes. Computers are manifestations of yet another causal principle: the evolution of culture.

Thus pattern in the structure and dynamics of information we encounter in the everyday world can be traced to three broad causal principles the physical constraints and self-organizing properties of matter, biological evolution, and cultural evolution. This classification scheme, like all classification schemes, is somewhat arbitrary. There may be subclasses of these principles that deserve to be considered principles unto themselves [note 2], or one could argue that evolution is a self-organizing property of matter, albeit a spectacular one [note 3]. The point is: culture is the only process that has arisen since the origin of life that relentlessly exploits the combinatorial potential of information. Despite the fact that culture is grounded in biology (like biology is grounded in the physical constraints and self-organizing properties of matter), the probability of computers arising spontaneously in an information space not acted upon by cultural evolution (like the probability of giraffes arising spontaneously in an information space not acted upon by biological evolution) is vanishingly small. Thus it is inappropriate to dismiss culture as a predictable extension of biological evolution. It is different from anything else biology has produced.

Since the machinery that renders cultural evolution—the human brain--is a product of biological evolution, much of what is 'out there' cannot be cleanly traced to a biological or cultural origin. We will discuss how biology constrains culture through the preferential spread of memes that satisfy biologically-derived needs. It goes the other way too; culture not only affects biological fitness through its effect on behavior (a phenomenon known as the Baldwin Effect) but it dramatically modifies the biological world. Some of the ways in which biological information gets tainted with cultural information seem relatively inconsequential, such as the trimming of hedges, whereas others, such as dog-breeding, have a long-lasting effect. (In fact, one could view dogs as the consequence of a memetic trajectory that was launched by the need to protect property.) Nevertheless, much as it is not imperative to address the role of physical constraints like gravity in studies of, say, embryonic development or squirrel foraging behavior, much can be said about culture without addressing the role of biological constraints.

### 1.3 TAKING THE MEME-EVOLVING APPARATUS SERIOUSLY

The most salient shortcoming of existing models of cultural evolution, and quite possibly the reason they have had relatively little impact, is that although they tell us much about the transmission or characteristics of 'catchy' memes, they fail to adequately address the issue of how novelty is generated. To be of significant theoretical or predictive value, memetics must give serious consideration to the processes by which experience in the world turns into new memes in our brains, and address how memetic novelty is expanded further through creative processes.

Studies of creativity, on the other hand, suffer from the opposite problem. They tend to focus on the individual, obscuring the fact that creativity is a collective affair [e.g. Lenat 1974; Schank & Leake 1989; Mitchell 1993; Boden 1991]. The ideas and inventions an individual produces build on the ideas and inventions of others. Which memes spread and which ones die off reflects the dynamics of the entire society of individuals hosting them.

Although at a sufficiently abstract level the notion that culture evolves is obvious, we lack a theoretical framework that bridges transmission studies with studies of creativity, and spells out explicitly how the concept of evolution maps on to the case of culture. To accomplish this we

need more than a quick and dirty list of what makes a meme catchy; we have to consider the extent to which it resonates with and enriches the complex web of assumptions, beliefs, motives, and attitudes of its host--no simple matter. In short, we have to take the meme-evolving apparatus very seriously.

## **2. MEME: INFORMATION UNIT THAT EVOLVES THROUGH CULTURE**

Because the concept of the individual is so germane to the human experience we have an anthropocentric tendency to assume that the individual is the appropriate basic unit of analysis. This tendency is probably exacerbated by the fact that in biology the individual is the object of the relevant evolutionary process (or more specifically the phenotypic expression of the information undergoing evolution). The memetic approach involves relinquishing our focus on the individual and concentrating instead on the meme as the object of a second evolutionary process that makes cognition possible. This perspective can feel unnatural and disorienting but it discloses population-level phenomena that would otherwise go unnoticed because they are not readily detected through introspection.

### **2.1 DISTINCTION BETWEEN A MEME AND ITS PHENOTYPIC IMPLEMENTATION**

Durham [1991] defines a meme as "any kind, amount, and configuration of information in culture that shows both variation and coherent transmission." Problems with this definition arise because it does not distinguish between cultural information as mental representation and cultural information as implemented behavior or artifact.

The genotype-phenotype distinction is useful here. The cultural analog of a genotype is the mental representation of a meme, and the analog of a phenotype is its implementation, or the form it takes if it gets expressed or communicated, typically as action or vocalization. Implementation transforms a meme, incorporating syntactic features characteristic of the channel through which it is conveyed [Brooks 1986]. Thus, for example, a dance step looks different with each individual who performs it.

### **2.2 INTRA-MEME STRUCTURE: CORE, ENABLER, AND HITCHHIKER FEATURES**

Biologists use the term &allele to capture the notion of alternative heritable versions of a gene, and Durham [1991] accordingly adopted the term TallomemeU to refer to alternative versions of a meme. This basic concept was tailored to meet the constraints of biology; we all have the same number of genes, and two alleles of each gene (one from each parent). The cultural analog may be too clumsy to capture the subtle relationships between memes. Memes often appear to be stored in a distributed, network-like fashion, connected through webs of association [Hebb 1949; Quillian 1968; Pribram 1974]; there is not necessarily a definitive rationale for saying where one stops and another begins, in semantic space let alone physical space. For example would we consider 'My mother looks good in blue' and 'My mother looks good under a blue umbrella' to be allomemes of the same meme, or different memes?

This kind of difficulty is circumvented by avoiding the notion of alternate versions altogether and using the term feature to refer to a component of a meme. Thus related memes share features. In this paper, feature can refer to a component with any degree of granularity below that of the meme in question; thus the scope of what might be considered a feature could range from an entire array of visual information depicting every perceived quality of a particular umbrella (such as might occur early on in perception) to one bit of information indicating the presence or absence of an umbrella (such as might occur at an advanced stage of cognitive processing). There can be nonlinear (epistatic) relations amongst the features of a meme, as well as between a meme and its implementation.

A first step toward a science of memetics is to decompose memes into features or feature schemata according to how they relate to fitness. Here we will distinguish the following categories: (1) core features that contribute directly to the fitness of a meme, (2) enabler features, that enable or facilitate the implementation of core features, and (3) hitchhiker features, which exist in the meme due to arbitrary or accidental historical associations with features of the first two kinds. Core features tend to convey semantic information, and enabler features syntactic information, though one can think of situations in which some semantic information serves simply to facilitate expression of other semantic information i.e. functions as an enabler. The first two categories are vaguely analogous to the categorization of genes as structural or regulatory, and the last category is inspired by the phenomenon of genetic hitchhiking [Kojima & Schaeffer 1967; Maynard Smith & Haigh 1974; Kaplan, Hudson, & Langley 1989]. The closer together genes are on a chromosome the less likely they will be separated by crossover, so the more tightly linked they are said to be. Hitchhiker alleles confer no fitness advantage, but endure because they are linked to alleles that are important for survival. In both genetic hitchhiking and its cultural analog there is indirect selection for useless (or even detrimental) patterns through their association with beneficial ones. The concept of hitchhiking is closely related to that of exaptation -- the evolution of organs or traits not evolved through natural selection for their current use [Gould 1982].

### **2.3 MEME AS AN EXPERIENCE ENCODED IN THE FOCUS**

In his book *Thought Contagion*, Aaron Lynch claims "Much as computer viruses make up a small fraction of computer software, so too do pernicious thought contagions form a small part of our mental software." But is this conservative approach necessary, or even sensible? Where do you draw the line between potentially culturally-transmittable memes and the rest our "mental software"? The haiku-writer, for example, seems able to draw upon any sort of thought, memory, or imaginary construct.

Here we will explore the opposite approach to that taken by Lynch; we bite the bullet and consider anything that can be the subject of an instant of experience, to be a meme. The category meme now includes not only obviously transmittable ideas like 'Be good or you will go to Hell', but everything from a particular experience of vibrant redness, to a realization of a shorter route to work, to a feeling of elation associated with a friend's facial expression or posture. This may strike some readers as outrageous, but it doesn't really make things as unwieldy as it might seem to at first glance. For the price of this added complication we gain a bridge that connects memetics with phenomena like perception, body language, planning, deductive reasoning,

emotion, categorization, analogy... the stuff of the social and cognitive sciences. It may be our only viable direction. A theory of mind that can account for myth and freestyle dance, not to mention rapid personality assessment, is not easily achieved.

Our concept of meme is made more manageable by invoking Kanerva's [1988] notion of the focus--that part of the mind in which sensation (either external or internal e.g. hunger) and stored memory interact to produce a stream of experience. The states of the neurons that comprise the focus determine the content and experiential quality of an individual's awareness. One can think of a meme as a pattern of information that is or has been encoded in an individual's focus. It can be subjectively experienced as a sensation, idea, attitude, emotion, or combination of these, and it can direct implementation by the motor apparatus.

### **3. THE MEME-EVOLVING APPARATUS: CONCEPTUAL NETWORKS**

Memes are woven together in an internal model of the world or worldview, which their host calls upon to figure out what to do whenever a situation is too complicated for its hardwired instincts. The worldview orchestrates behavior such that a meme gets implemented right when it is likely to be useful, and that increases the probability that other hosts will consider it worthy of replication. This section briefly discusses some ideas about how the worldview is structured.

#### **3.1 REPRESENTATION OF LOGICAL RELATIONSHIPS**

A worldview with enough internal consistency to put together strategies and solve problems must have a way of representing not only memes themselves but the logical relationships amongst them. McCulloch and Pitts [1943] showed that networks made of neuron-like components that perform the logical operations AND, OR, and NOT are theoretically capable of computing any Turing machine-computable function. In connectionist-type systems, logical relations are represented implicitly as constraints on the possible states of a system, and computation proceeds through settling into a solution that satisfies many constraints, rather than then explicitly calculating a solution. This is accomplished through modification of the associative strengths amongst the components of the system, and the process is referred to as relaxation or simulated annealing.

#### **3.2 THE BABY AND THE BATHWATER: CONCEPTUAL LINKAGE DISEQUILIBRIUM**

Arguments against a theory of cultural evolution generally consist of a series of statements as to how the cultural situation differs from that of biology [e.g. Gould 1991; Thagard 1980]. These arguments, however, do not constitute a viable reason to discard the idea that culture is an evolutionary process. Imagine that 100 years before Darwin proposed the theory of biological evolution through natural selection, another scientist had discovered another system whereby information patterns evolved, say in a test tube. Given this scenario, would it have made sense for Darwin to dismiss biological evolution simply because it works differently from the test tube form? This would obviously have been foolish. It would have robbed humanity of not only the unifying power of a theory of biological evolution, but the opportunity to use knowledge of how evolution works under one set of constraints and affordances as a scaffold to direct the study of

how it works under different constraints and affordances. But time and again it is argued that a theory of cultural evolution is doomed simply because it would have to work through different mechanisms from those of biological evolution.

Ironically this situation in itself provides us with a nice example of how knowledge of evolution acquired in the realm of biology can help unravel analogous situations in the realm of culture. The biasing effect of historical association is an important theme in population genetics. Alleles of linked genes, such as those that code for red hair and freckles, continue to co-occur more often than chance even after individuals in the lineage from which these alleles originated begin mating randomly with individuals from other lineages that did not have these alleles. One can theoretically measure the number of generations necessary for these genes to achieve a state of random association or linkage equilibrium, and this process can be modeled computationally. Similarly, people often have difficulty applying an idea or problem-solving technique to situations other than the one in which it was originally encountered, and conversely, exposure to one problem-solving technique interferes with the ability to solve a problem using another technique [e.g. Luchins 1942]. Psychologists refer to this as mental set (though it is more commonly known as 'throwing the baby out with the bathwater'). We could view mental set as a state of conceptual linkage disequilibrium. In the present example, conceptual linkage disequilibrium hinders our ability to abstract the basic concept of evolution from its biological manifestation so that it can be applied with ease to the case of culture. One could argue that it would make sense for cultural evolution to be the default form of evolution in disciplines outside of biology, much as in tropical climates the default form of skiing is water-skiing rather than snow-skiing.

Conceptual linkage equilibrium is achieved when all instances of hitchhiking have been obliterated. In an influential paper on the relationship between DNA polymorphism and recombination rates, Begun and Aquadro [1992] suggested that genetic hitchhiking may have significant evolutionary impact:

This correlation suggests that levels of neutral variation in many of the gene regions for which variation has been measured have been reduced by one or more hitchhiking events. Provided that a new selectively favored mutation goes to fixation before another advantageous mutation arises close to it, each fixation will be surrounded by a 'window' of reduced polymorphism, the relative size of which is proportional to the rate of recombination for that region of the genome.

The general idea here translates nicely to cognition: if a meme goes to fixation in a society due to selective advantage conferred by one or more core features, its enabler and hitchhiker features will also exhibit reduced polymorphism, and the size of the window will vary with the extent to which hitchhiker features are conceptually bound to that meme. (For example, the basic concept of a typewriter/computer keyboard was historically enabled by the QWERTY keyboard design, which has now gone to fixation despite poor performance relative to other possible keyboard designs.)

### **3.3 INDIVIDUAL VERSUS SOCIETAL WORLDVIEWS**

Our worldviews overlap to the extent that similar experiences and genetic make-ups cause our brains to store the same memes and styles of computation. But they don't overlap perfectly. Each host's train of thought traces out a unique trajectory through conceptual space. It can be useful to think in terms of not only the worldview of an individual, but also the worldview of a group or even human society at large, wherein all frontiers of human endeavor are incorporated.

## **4. SELECTION AND MEME FITNESS**

The next few sections examine in some detail how each of the three phases of evolution--selection, variation, and transmission--map onto the case of culture. Though these phases are discussed one at a time, it is worthwhile to keep in mind that in culture they are less spatiotemporally distinct than in biology. Selection can be coupled to either the generation of variation, or replication, or all three can occur simultaneously (for example when paraphrasing).

### **4.1 MEMES RELY ON BRAINS TO SELECT, VARY AND REPLICATE THEM**

Von Neumann [1966] postulated that any self-replicating system consists of two parts: (1) uninterpreted information—a self-description that is passively copied to offspring and (2) interpreted information instructions for how to construct offspring. This turned out to be true of the genetic code; there are genes that provide instructions to the body for how to sustain itself, and genes that provide instructions for how, with the help of someone of the opposite sex, to create a child. But unlike genes, memes do not come prepackaged with instructions for their reproduction. They rely on the pattern-evolving machinery of their hosts' brains to create, select, and replicate them. Since we preferentially spread ideas that satisfy needs, our needs define viable niches for memes to evolve toward. As infants we might cry and kick no matter what need is most pressing, but as children we acquire and continually refine a repertoire of memes that, when implemented, satisfy various needs. We learn that reaching into the cookie jar satisfies one need, shouting help satisfies another, et cetera. Our memes, and the behavior they elicit, slide into need-defined attractors (regions of stability) in the memetic fitness landscape.

The fact that memes are not independently self-replicating does not prevent them from achieving reproductive success. In fact it may ironically work in their favor, because the cognitive machinery they depend upon not only (1) organizes them into a worldview which orchestrates their implementation in a way that makes their usefulness apparent, but (2) transforms them to produce 'offspring-memes'. Both tendencies--to reorganize memes, and create new ones--stem from the need to be able to evaluate and carry out plans of action, and seem to resurface when needs more directly relevant to survival have been met. We now turn to how these more pressing survival needs affect meme evolution.

## **4.2 BRAINS SELECT MEMES THAT SATISFY BIOLOGICAL AND CULTURAL NEEDS**

Since many of our needs have a biological basis--e.g. the need for food, shelter, et cetera--meme generation is largely constrained by our heritage as products of biological evolution. Thus the topology of the memetic fitness landscape largely echoes that of the biological fitness landscape. In the short term, the biological fitness landscape, and thus the memetic fitness landscape, fluctuates continuously as one need is satisfied and others take precedence [Hull 1943; McFarland & Sibly 1975; Gabora & Colgan 1988; Maes 1991]. For example, after eating, ideas that pertain to finding food are less likely. However over the lifetime of an individual the set of biologically-based needs remains relatively constant. The trajectory of survival-motivated thought can be described as a limit cycle (periodic attractor) that moves through the set of stable memes whose implementations satisfy the various biological needs.

Variation-inducing operations (which will be discussed in the next section) restructure conceptual space and thus affect the memetic fitness landscape. Much as the evolution of rabbits created ecological niches for species that eat them and parasitize them, the invention of cars created cultural niches for gas stations, seat belts, and garage door openers. As one progresses from infancy to maturity, and simple needs give way to increasingly complex needs, the stream of thought acquires the properties of a chaotic or strange attractor, which can be viewed as the formation of crevices in the original limit cycle. The landscape is fractal (i.e. there is statistical similarity under change of scale) in that the satisfaction of one need creates other needs every crevice when examined closely reveals more crevices. This is analogous to the fractal distributions of species and vegetation patterns described by ecologists [Mandelbrot 1982; Palmer 1992; Scheuring & Riedl 1994]. An endpoint of a cultural evolution trajectory turns out to be not just a point in multidimensional space, but a set of points with their own fitness metric--a micro-landscape in its own right. So although the memetic fitness landscape loosely follows the biological fitness landscape, there are places where it deviates, and this effect undoubtedly becomes more pronounced throughout an individual's lifetime. This means that the potential for meme diversity is open-ended. Fueled by need and constrained by association we carve out trajectories through meme space, and because the fitness landscape that guides this process is fractal, every time that landscape steers the production of a new meme (or even just a slight variant of a preceding meme), the new meme in turn redefines the landscape, and so on, recursively. Thus there is a continuous co-evolutionary interplay between meme and landscape, and this would seem to contribute to the oft-noted rapidity with which culture evolves.

## **4.3 HARD-WIRED SELECTION**

To the extent that the memetic fitness landscape echoes the shape of the biological fitness landscape, to which we have been adapting since life began, cultural selection is built right into our architecture. Our perceptual and cognitive systems are wired up such that they are primed to focus on and highlight those aspects of external reality that are relevant to our survival (or were in the past). The mental representations we form reflect that bias [e.g. Hubel & Wiesel 1979; Marr 1982].

Second, the associative organization of memory both permits and constrains meme storage and variation- generating operations. This is discussed in detail extensively elsewhere [e.g. Van Loocke1991 and this volume] and will not be addressed here.

#### **4.4 MALLEABLE FORMS OF SELECTION**

In order to create, or even just understand, a new meme, there has to be a conceptual framework from within which it will make sense, and a need, or niche, for it. Therefore, any relevant precursor or proto-memes must first be assimilated [Wallas 1926]. This constraint amounts to a malleable, or plastic, form of selection on new memes.

When one host exhibits a meme that another observing host values, the observing host often displays reinforcing body language and emits words of encouragement. Likewise, if the meme is threatening or inconsistent with valued memes, the observing host's words and behavior tend to be discouraging. Our need for social acceptance makes us more likely to exhibit memes that have been reinforced and less likely to exhibit those that have been discouraged. From a meme's-eye perspective this looks like a subtle strategy whereby a meme acting through one host selectively shapes the probability of its implementation in other hosts.

Selection can also occur after a representation has been internalized but prior to being phenotypically expressed. For example, mentally &simulating what would happen if an idea were implemented can weed out unworthy ideas [Nersessian 1993]. The success of mental simulation varies with the accuracy of one's' internalized model of the world, but it provides at least a rudimentary form of selection.

Finally, selection can operate through biased transmission; that is we choose to imitate certain individuals and not others [Boyd & Richerson 1985].

### **5. CREATIVITY: WELLSPRING OF CULTURAL VARIATION**

#### **5.1 STRATEGY GUIDES TRAJECTORIES THROUGH MEMETIC FITNESS LANDSCAPE**

The existence of an open niche does not guarantee that the niche will ever be found. In biology the process by which this happens is largely random. Though most mutations and recombinations are detrimental, so many variants are generated that it is not necessary to be clever about how they are generated. We could say that biological evolution is a more breadth-first search algorithm than cultural evolution because it relies primarily on massive parallelism rather than strategy.

In culture, on the other hand, variants are generated strategically. We could say that cultural evolution is a more depth-first approach to searching a space of possibilities. The trajectory of a stream of thought is constrained by connections between representations that are similar or spatiotemporally related [Schank 1983], which increases the probability that an advantageous variant is found. For example, when considering the problem of having to get out of your car every day to open the garage door, you would not think about doilies or existentialism, but

concepts related to the problem electricity, human laziness, and various openers you have encountered before. If you were to spend several months on this problem, ideas that pertain to openers of various kinds would for you become a highly active region of conceptual space, analogous to the uncharacteristically high level of activity (and polymorphism) in a small portion of the human genome known as the major histocompatibility complex (MHC) which deals with immune response [Hughes & Nei 1988].

## **5.2 WORLDVIEW REORGANIZATION: CHUNKING, CATEGORIZING, AND REDESCRIBING**

Frequently many memes get integrated into one through a process referred to in the psychological literature as chunking [Miller 1956]. Chunking involves forming associations amongst previously-learned memes and establishing this constellation of associations as a new meme in long term memory; it is analogous to the formation of co-adapted genes, or schemata [Holland 1975].

Whereas chunking generally refers to the binding of semantically unrelated memes (as in the memorization of an arbitrary string of numbers), categorization involves the recognition of semantic relationships. Categorization and the resultant hierarchical structure of knowledge is discussed in other chapters and will not addressed here in any depth, though it is of relevance to point out that it seems reasonable to expect that the more extensively memes have been chunked or categorized, the greater the complexity of what can be held in the focus at once. Thus what constitutes a meme (and thus a feature) will differ amongst individuals and within an individual over time.

During creative thought, memes potentially relevant to a solution would evoke or activate one another, altering or strategically (though not necessarily consciously) manipulating them, a process that is said to involve pattern completion, constraint satisfaction [Rumelhart & McClelland 1986], and the tweaking, blending, redescription, abstraction, and recoding of representations [Hofstadter 1985; Holland et al. 1986; Karmiloff-Smith 1986; 1992; Ram 1993; Clark & Thornton 1995]. Neurophysiological evidence suggests that creating new contexts for representations, that is manipulating them, involves hippocampal binding or linking [Squire 1992], and synchronization [Klimesch 1995], of features encoded by distributed cortical cell assemblies.

One could argue that recreation is the re-creation of information patterns in different domains from the ones in which they were originally encountered, thereby filtering out conceptual prejudices that reflect nothing more than mechanical constraints or historical legacies of the original domain. Play, intellectual pursuits, and other creative endeavors are then algorithms for achieving a state of conceptual linkage equilibrium through mental operations that, like genetic recombination, increase polymorphism by reducing fixation through hitchhiking.

### 5.3 SPARSE DISTRIBUTED MEMORY AS A PLATFORM FOR GENERATING NOVELTY

If the worldview is like a patchwork quilt of memes, thought is like shining a flashlight onto the quilt such that the light is not restricted to one patch but can overlap several, distilling what is relevant from each. In a train of consecutive thoughts we spontaneously weave external stimuli with concepts or experiences that are related, sometimes superficially, and other times through abstract or metaphorical resemblances. That is, experience has a thread of continuity. Sparse, distributed & memory [Kanerva 1988], or SDM, is a mathematical model of the mechanics underlying the storage and retrieval of memories that nicely captures these aspects of cognition.

Kanerva draws an analogy between the focus and a combined address-datum register in a computer; they both contain data and serve as a pointer to memory, and can both read from and write to memory. An instant of experience is encoded in the focus by a high-dimensional vector of difference relations, or bits, that represent the presence or absence of some feature, and the mathematics generalizes such that a pattern of bits can represent a value along some dimension. The Hamming distance between two memes is the number of bits that differ. (So the Hamming distance between 11111 and 11100 is two.) Since each meme has an antipode (for example, the antipode of 11111 is 00000), the space of all possible memes can be visualized as a sphere. The address of a meme is the information pattern that specifies where the meme is stored.

If  $L$  is the number of possible features in a meme, the number of possible memes is  $2^L$ . Assuming  $L$  is large the size this space is enormous, so the memory is sparse in that it stores only a small fraction of the set of all possible memes. For example, to construct a SDM with  $L=1,000$ , then out of the  $2^{1,000}$  possible addresses, a workable number of them, say 1,000,000, are chosen at random to be actual storage locations. The number of memes at Hamming distance  $k$  away from any given meme is equal to the binomial coefficient of  $L$  and  $k$ , which is well approximated by a Gaussian or normal curve. If meme X is 111...1 and its antipode 000...0, and we consider meme X and its antipode to be the poles of the hyper sphere, then approximately 68% of the other memes lie within one standard deviation ( $\sqrt{L}$ ) of the equator region between these two extremes. As we move through Hamming space away from the equator toward either Meme X or its antipode, the probability of encountering a meme falls off sharply by the proportion  $\sqrt{L}/L$ . In our example, the median distance from one location to another is 424 bits, and 99.8% of stored memes lie between 451 and 549 bits of any given location.

A computer reads from memory by simply looking at the address in the address register and retrieving the item at the location specified by that address. The sparseness of the SDM prohibits this kind of one-to-one correspondence, but it has two tricks up its sleeve for getting around this problem.

First, it feigns content addressability, as follows. The particular pattern of 1s and 0s that constitutes a meme causes some of the synapses leading out from the focus to be excited and others inhibited. The locations where memes get stored are memory neurons, and the address of a neuron amounts to the pattern of excitatory and inhibitory synapses from focus to memory that makes that neuron fire.

Activation of a memory neuron causes the meme to get written into it. Thus there is a systematic relationship between the memes' information content and the locations they activate.

Second, since the probability that the ideal address for storing a meme corresponds to an actual location in memory is vanishingly small, storage of the meme is &distributed across those locations whose addresses lie within a sphere (or more accurately, hypersphere) of possible addresses surrounding the ideal address. The radius (in Hamming metric) of this sphere is determined by the neuron activation threshold. Each location participates in the storage of many memes. In this example we assume that 10,000 memes have been stored in memory. Each meme is stored in 1,000 (of the 1,000,000 possible) locations, so there are approximately 10 memes per location. The storage process works by updating each of the L counters in each location; to store a 1 the counter is incremented by 1, and to store a 0 it is decremented by 1. These nearly one million operations occur in parallel.

If after a meme, say meme X, is stored, the individual's attention is directed toward external stimuli, then nothing is retrieved from memory. But to the extent that memory contributes to the next instant of awareness, the storage of X activates retrieval of not only X itself but all the other memes that have been stored in the same locations. The next meme to be encoded in the focus, X', is found by determining the best match; that is, by averaging the contributions of all retrieved memes feature-by-feature. Whereas the 1,000 retrieved copies of X (and memes similar to X) reinforce one another, the roughly 10,000 other retrieved memes are statistically likely to cancel one another out, so that XU ends up being similar to X. Though X' is a reconstructed blend of many memes it can still be said to have been retrieved from memory. X' can now be used to address the memory, and this process can be reiterated until it converges on meme Y that satisfies a current need. This is how the SDM accomplishes depth-first search. The closer Y is to X, the faster the convergence. In our example, assuming  $r = 425$ , if X and Y are more than 200 bits apart Y is unlikely to be retrieved, but if they are 170 bits apart Y will be retrieved in about four iterations.

Keeler [1988] has shown that SDM is a superset of Hopfield-type and connectionist models of auto associative or hetero associative memory. The SDM formulation is used here because it lends itself to an understanding of the mechanics of phenomena we are interested in. Because of how the dynamics emerges from the statistics, rather than from a central executive, it can cope with creative and seemingly unmechanical cognitive phenomena such as wordplay or slips of the tongue. Moreover it is ideally suited to handle the problem of sequential access, which will become relevant when we look at how an infant establishes a train of thought. To model the recollection of a sequence, meme X is simply used as the address to write Y, Y as the address to write Z, and so on. Working memory can be viewed as the memes that lie within a given Hamming distance of the meme in the focus such that they are retrievable within a certain number of iterations. Categorization could involve the identification of a pattern, and readdressing memes that contain this pattern so that their new addresses put them within working memory reach of one another. Kanerva shows that the architecture of common neural components and circuits in the brain are ideally suited to implement a SDM.

In SDM, associations between memes are not explicitly represented as connection strengths but as proximity in multidimensional space. However in the end they amount to the same thing.

The smaller the Hamming distance between two memes, the higher the probability that they will be retrieved simultaneously and blended together in the focus (or one after the other in a chain of related thoughts). What allows them to be retrieved simultaneously, however, is that they are either stored in the same neurons or in neurons with nearby addresses, which in turn reflects the neurons' connectivity. Thus factors that affect the storage of a meme will also affect retrieval of that meme; the two processes are intimately connected.

## 6. MEME REPLICATION AND TRANSMISSION

The memetic approach to cognition is not incompatible with approaches that stress the role of innate mechanisms [e.g. Pinker 1995]. Rather, as Lumsden and Wilson [1981] point out, it builds on this framework, adding that the study of cognition will flounder until we admit that the role of transmission is equally undeniable. Transmission links the memetic processing within an individual to not only memetic processing in other directly-encountered individuals, but processing in individuals they encounter, and so on. The ideas and inventions any one individual produces build on the ideas and inventions of others. This phenomenon is known as the ratchet effect, and its impact is demonstrated in the following example. If you were suddenly dropped into the Australian desert, you probably would not survive for long. However if you were to run into an aborigine who grew up learning desert survival skills from her family and community that had been passed on and improved upon for generations (such as how to find water in obscure places) you might survive for some time [note 4].

### 6.1 INTRA-HOST MEME REPLICATION VIA IMPLICIT POINTERS TO MEMORY

We saw how, unlike genetic material, memes do not contain instructions for how to make copies of themselves; they replicate when their hosts teach or imitate one another. The memes in a SDM-like memory, however, have a self- replication capacity in the following sense.

The pattern of information that constitutes a meme determines which of the synapses leading out from the focus are excited, and which are inhibited--it determines how activation flows through the memory network--which in turn determines the neurons where the meme is stored and from which the next meme is retrieved. Thus embedded in the neural environment that supports their replication, memes act as implicit pointers to memory. These pointers prompt the dynamic reconstruction of the next meme to be subjectively experienced, which is a variation of (statistically similar to) the one that prompted it. It is in that sense that they self-replicate.

### 6.2 TRANSMISSION IS LAMARKIAN AND PHENOTYPICALLY MEDIATED

Internal replication (with variation) makes cultural transmission Lamarckian—modifications acquired since the acquisition of a meme can be passed on to others [Dawkins 1976]. The related point that transmission is phenotypically mediated, as Dennett [1995] points out, makes a science of memetics less daunting. It means that, unlike biologists, we don't have to fully understand the nature of mental representation to study transmission.

### **6.3 ALL STIMULI SCULPT MEMES**

While biological needs affect the focus from the inside, environmental stimuli impact it from the outside. The information-based orientation supports a broader conceptualization of the transmission process than is generally taken. Transmission often occurs through imitation of conspecifics [Smith 1977; Bonner 1980; Robert 1990], or guided instruction [Vygotsky 1978; Tomasello et al. 1993], but not necessarily. For example, does it matter whether a child learns to peel a banana by watching her mother, or a monkey, or a cartoon character on TV? What matters is that the child has a mental representation of how to peel a banana. All interaction between an organism and its environment sculpts the worldview, even if just to increase the organisms' assessment of the probability of encountering a certain stimulus. Therefore all stimuli can potentially affect the interplay of ideas and emotions that evolve through culture.

## **7. A SCENARIO FOR THE ORIGIN OF COGNITIVE EVOLUTION**

We have discussed how memes evolve through selection, variation, replication and transmission. We now address the issue of how this second form of evolution might have begun in the first place.

### **7.1 THE ORIGIN OF LIFE AND ITS COGNITIVE ANALOG**

The origin of life poses the following paradox: how could something as complex as a self-replicating molecule arise spontaneously? Traditional attempts to explain this entail the synchronization of a large number of vastly-improbable events. Proponents argue that the improbability of the mechanism they propose does not invalidate it because it only had to happen once; as soon as there was one self-replicating molecule, the rest could be copied from this template. However Kauffman [1993] proposes an alternative scenario that does not entail the synchronization of numerous improbable events. He suggests that life arose through the self-organization of a set of autocatalytic polymers. When catalytic polymers interact with one another their average length increases. As their length increases, the number of reactions by which polymers can interconvert increases faster than the number of polymers. Therefore a set of interacting molecules under conditions such as are likely to have existed at the time life began would inevitably reach a critical point where there is a catalytic pathway to every polymer present. Jointly they form a self-reproducing metabolism.

We now ask: What is the cultural analog to the origin of life? One could say it is the point in history when organisms acquired the capacity for social transmission, but as many authors [e.g. Darwin 1871; Plotkin 1988] have pointed out, although transmission is wide-spread throughout the animal kingdom, no other species has anything remotely approaching the complexity of human culture. Donald [1991] argues convincingly that the bottleneck in cultural evolution is the capacity for innovation. Innovation requires more than a kind of awareness that integrates survival needs with environmental affordances, and draws upon memory only to interpret stimuli, or consult a mental map, or recall how some drive was satisfied in the past. It requires an ongoing train of creative representational redescription and counterfactual 'what if...' type thinking. This suggests that the cultural analog to the origin of life was the origin of the first self-perpetuated, potentially-creative stream of thought in an individual's brain.

When an infant has its first experience, there is nothing in memory to draw upon to contribute to that experience; the first meme to occupy its focus does not remind it of anything. Therefore experience is initially driven only by external or internal stimuli, not by memory. Thus the evolution of culture poses a paradox analogous to that of the origin of the self-replicating molecule how does an infant develop the capacity for a self-sustained train of thought that creatively integrates new experiences with previous ones? Consistent with Kauffman's assertion that the bootstrapping of an evolutionary process is not an inherently improbable event, the 'it only had to happen once' argument does not hold water here because the cultural analog to the origin of life takes place in the brain of every infant.

## 7.2 ESTABLISHING AN AUTOCATALYTIC SET OF SPARSE, DISTRIBUTED MEMORIES

This section outlines how a SDM-like stream of thought might get established. Let us say that the first meme to occupy an infant's focus and then get stored in memory is a visual experience of its mother in a blue coat. The next is the sound of a dog barking. The Hamming distance between these memes exceeds the maximum for one meme to evoke the memory of another, so the barking does not remind the infant of its mother. Later the infant sees its mother in a red coat. This meme evokes or catalyzes the memory of its mother in a blue coat. To avoid getting stuck in an endless loop wherein 'mother in blue coat' then evokes 'mother in red coat' et cetera, it may form the category 'mother'. However that meme does not remind it of anything, so this stream of thought dies off quickly.

As the infant accumulates memes, the statistical probability that a meme in the focus will activate a meme from storage increases, so the streams of remindings get longer. Eventually the memory becomes so densely packed that any meme that comes to occupy the focus is bound to be close enough in Hamming distance to &some previously- stored meme(s) to activate a variant of itself. This marks a phase transition to a state in which, just as with the origin of life, the sequential activation of self-similar patterns is self-propelled; the memes now form an autocatalytic set. The focus is no longer just a spot for coordinating stimuli with action, but a forum for the variation-producing operations that emerge naturally through the dynamics of iterative retrieval. The resultant memes evolve along different trajectories toward different basins of attraction, specializing in the fulfillment of one need or another. Those that satisfy the same need compete until one becomes habitual, while those that fulfill different needs are able to coexist within the same host. As with biological speciation, small differences are amplified through positive feedback leading to transformation of the space of viable niches for the evolution of information patterns.

Note that in this example the 'mother' meme is the infant's first category. A simple way of describing this situation is: if the 'mother in blue coat' meme is represented as 111, and the 'mother in red coat' meme is represented as 110, the 'mother' meme can be represented as 11\*, where \* means either 1 or 0. It is also the infant's first derived meme. That is, it is the first information pattern to enter the focus not purely by way of external stimuli but through the necessity of a logical operation on previously-stored memes--in this case an OR gate--which could be realized in the brain via adjustment of connection strengths. The act of categorization projects the original information space, which had n relevant dimensions, onto a new space that

has  $n-1$  dimensions (for example, here coat color is no longer relevant). It effectively makes the space denser, and increases susceptibility to the autocatalytic state. On the other hand, creating new memes by combining previously-stored memes could interfere with the establishment of a sustained stream of thought by increasing the dimensionality of the space, thereby decreasing density. If indeed cross-category blending disrupts conceptual autocatalysis, one might expect it to be less evident in young children than in older children, and this expectation is born out experimentally [Karmiloff-Smith 1990].

Note also that the density of memes necessary to reach and maintain this autocatalytic state will depend on the neuron activation threshold. If the threshold is too high (the hypersphere of potentially activated memes is too small) even very similar memes cannot evoke one another, so a stream of reminders, if it happens at all, dies off readily. If the threshold is too low (the hypersphere too large), then any meme will evoke a multitude of others not necessarily meaningfully related to it. Successive patterns in the focus will have little or no resemblance to one another; the system may be catalytic but it is not autocatalytic. The free-association of the schizophrenic [see Weisberg 1986] seems to correspond to what one might expect of a system like this. For memory to produce a steady stream of meaningfully-related yet potentially creative reminders, the threshold must fall within a narrow intermediate range. This is consistent with Langton's [1992] finding that the information-carrying capacity of a system is maximized when its interconnectedness falls within a narrow regime between order and chaos. The situation may turn out to be slightly more complicated; sustaining a creative train of thought may involve not only keeping the activation threshold within a narrowly-prescribed range but dynamically tuning it in response to the situation at hand. This is particularly likely if the memory is not uniformly dense (i.e. clusters of highly-correlated memes) or if different kinds or stages of thought require different degrees of conceptual blending. For example, finding unusual associations may depend on the preconscious ability to temporarily increase hypersphere radius.

Thus we have a plausible scenario for how cultural evolution, like biological evolution, could have originated in a phase transition to a self-organized web of catalytic relations between patterns.

## **8. WHY IS CULTURE UNIQUE TO HUMANS? A SPECULATIVE ANSWER**

Recall that in order for a network of memes to reach an autocatalytic state, the activation threshold must be calibrated to fall within a narrow range to achieve a delicate balance between the capacity for semantic continuity on the one hand and creative association on the other. The penalty for having too low a threshold would be very high; successive thoughts would not necessarily be meaningfully related to one another, and thinking would be so muddy that survival tasks are not accomplished. Too high a threshold, on the other hand, would not be life-threatening. The focus would virtually always be impacted with external stimuli or internal drives such as hunger; memory would be pretty much reserved for recalling how some goal was accomplished in the past. A stream of experience that involved the iterative reorganization of stored memes would likely die out before it produced something creative. This may be the situation present in most brains on this planet, and the reason that apes are limited to episodic memory [e.g. Donald 1991].

The advantages of a sustained train of thought would be largely lost on animals because they have neither the vocal apparatus nor the manual dexterity and freedom of upper limbs to implement complex ideas. No matter how brilliant their thoughts were it would be difficult to do something useful with them. Moreover, in an evolutionary line there is individual variation, so the lower the average activation threshold, the higher the fraction of individuals for which it is so low that they do not survive.

These ideas are of course very speculative. However it seems reasonable to suggest that apes are not a priori prohibited from evolving complex cognition, but that there is insufficient evolutionary pressure to tinker with the activation threshold until it achieves the requisite delicate balance to sustain a stream of thought, or to establish and refine the necessary feedback mechanisms to dynamically tune it to match to the degree of conceptual fluidity needed at any given instant. It may be that humans are the only species for which the benefits of this tinkering process have outweighed the risks.

## 9. A COMPUTATIONAL APPROACH

Meme and Variations, or MAV [Gabora 1995] is a computer model of a society composed of interacting neural network- based agents. Unlike other such models that combine biological and cultural evolution [e.g. Ackley 1994; Spector & Luke 1996] these agents don't have genomes, and neither die nor have offspring, but they can invent, implement, and imitate memes. MAV successfully evolves patterns of information through cultural implementations of variation, selection, and replication, and exhibits phenomena observed in biological evolution such as: (1) drift--random statistical bias due to sampling error [Wright 1951; Cavalli-Sforza & Feldman 1981] (2) epistasis increases time to reach equilibrium, (3) increasing frequency of variation-generating operations increases diversity, and (4) although in the absence of variation-generating operations meme evolution does not occur, increasing variation-generation much beyond the minimum necessary for evolution causes average fitness to &decrease. MAV also addresses issues specific to cultural evolution, such as the effects of mental simulation, imitation, and strategy. Perhaps the most interesting finding it yielded was that although for the society as a whole the optimal creation-to-imitation ratio was approximately 2:1, for the agent with the fittest memes, the less it imitated (i.e. the more computational effort reserved for creation) the better.

MAV will hopefully serve as a stepping stone to more advanced models of memetic evolution. Of particular interest will be models that: (1) like Tierra, a model of biological evolution [Ray 1991], harness the power of evolution to explore and transform an &open-ended space of possible patterns, but (2) explore the space strategically on the basis of accumulated knowledge rather than at random<sup>5</sup>, (3) have fitness landscapes that emerge through the needs of the agents within the constraints of their environment, and (4) have agents that must learn for themselves which memes, when implemented, best satisfy each of their various needs.

Mathematical models of culture are too minimal to cope with the open-ended diversity of culturally-derived information (variation is generally restricted to trial and error learning or transmission error) let alone address the numerous intra-individual factors that undoubtedly have emergent inter-individual consequences, such as how representations are grounded in experience

and how they are stored, retrieved, and implemented. Models of individual intelligence and creativity, on the other hand, lack transmission and replication.

Although this research may not explicitly attempt to address group processes it typically focuses not on the sorts of simple inferences and creative acts that a person raised alone in the wild would be capable of, but on complex acts such as story comprehension, that might be unlikely to develop in isolation. With the advent of massively parallel computers it is becoming increasingly feasible to place computational models of individual creativity and problem-solving in a cultural context. This approach could provide insight into not only problems pertaining to representation and culture, but evolution in general, through comparison with biology. For example, the question of why there is so much redundancy in the genetic code has generated much discussion which may also apply to the question of why there are redundant mental maps in the brain; both may reflect constraints on the nature of an information-evolving code.

## 10. A MEMETIC PERSPECTIVE ON PSYCHOLOGICAL CONCEPTS

We turn now to how the cultural evolution perspective can shed light on some aspects of the human experience.

### 10.1 MENTAL CENSORSHIP AND INDUCTION

Initially an infant is unselective about meme acquisition, since (1) it doesn't know much about the world yet, so it has no basis for choosing, and (2) its parents have lived long enough to reproduce, so they must be doing something right. However just as importing foreign plants can bring ecological disaster, acquisition of a foreign meme can disrupt the established network of relationships amongst existing memes. Therefore the infant develops mental censors that ward off internalization of potentially disruptive memes. Censors might also be erected when a meme is found to be embarrassing or disturbing or threatening to the self-image [Minsky, 1975]. In a SDM-type architecture this could be accomplished by increasing the activation threshold so as to prematurely terminate the relaxation process and prevent the content of the focus from assimilating with stored memes. Much as erecting a fence keeps people on one side or the other, erecting a censor around a meme would warp the statistical probability that it will partake in the ongoing process of associative recall such that its host either avoids the censored meme or dwells on it excessively. This seems to be consistent with our bipolar attitude toward highly censored subjects such as aggression and sexuality.

On the other hand, when the cost of the disruption is outweighed by the potential benefit accrued by a world model that can accommodate the new meme, the threshold would be lowered. Most thoughts seem to have little effect on our understanding of the world at large, but once in a while we experience a meme that significantly modifies our worldview. The situation is reminiscent of superconductivity; lowered resistance increases correlation distance, and thus a perturbation to any one pattern can percolate through the system and affect even distantly-related patterns. It would be interesting to determine experimentally whether the inductiveness of our memes, like other self-organizing systems, exhibits the ubiquitous inverse power law [Bak, Tang & Weisenfeld 1988]. Just as in a sand pile perched at the proverbial edge of chaos once in a while a collision between two grains will lead to another in just the right chain reaction to generate a

large avalanche, occasionally one thought will trigger a chain reaction of others in a way that reconfigures the conceptual network.

## 10.2 THE UNCONSCIOUS

The concept of the unconscious has been influential and useful despite the obvious incongruity: how is it that we can consciously discuss something that is unconscious? What we may be referring to is the fleeting experience of memes that are dynamically reconstructed as in a SDM but which do not readily assimilate with other memes and so get discarded from the focus. In other words, the need for worldview consistency prohibits further computational resources from being spent on trying to integrate what appears to be a nonsensical construction into the memory. Of course there is no reason why a meme that is not immediately integrated into the memory might nevertheless affect the memory; the very process of determining whether it can be assimilated or not might itself have effects that infiltrate the system. This possibility is supported by the finding that subjects' behavior can be affected by priming material of which they have no recall [e.g. Dunbar & Schunn 1990; Fehrer & Raab 1962]. Subconscious processing of this sort could, in fact, resculpt the memetic fitness landscape in such a way that a previously- discarded meme is more readily assimilated the next time it is encountered.

## 10.3 CULTURAL MOMENTUM

Despite being derived, directly or indirectly, from human need, memes do not always promote our survival [Greene 1978; Alexander 1980]. As Dawkins [1982] points out, "It is true that the relative survival success of a meme will depend critically on the social and biological climate in which it finds itself, and this climate will certainly be influenced by the genetic make-up of the population. But it will also depend on the memes that are already numerous in the meme-pool." Much like runaway selection in biology, once a meme can replicate with variation on the basis of some selection criterion, it can evolve out of the orbit of the need that originated it. We can't help but engage in a stream of thought, spontaneously generating new memes like 'if only such and such had been different...', any more than biological evolution can help but generate new species. This cultural momentum could explain why, despite the intuition that individuals control their streams of thought, creators often express surprise at the sudden appearance of an idea, and deny active effort in its immediate creation [Bowers et al 1990; Guilford 1979; Kubose et al 1980; Wallas 1926]. We seem to control the birth of our ideas only to the extent that we provide a fertile ground for them to be fruitful and multiply by internalizing relevant background knowledge, identifying new needs, and exposing ourselves to stimuli that help trigger ideas that fulfill those needs. (So if you don't like this idea, don't blame me.)

Spurious basins of attraction sometimes arise in recurrent neural networks through the compositional interaction of explicitly-trained attractors [Hopfield 1982]. Cultural momentum may boil down to a phenomenon of this sort. Just because the memetic fitness landscape largely echoes the biological fitness landscape, that doesn't mean that behavior elicited by memes in spurious basins of attraction arising through representational redescription need always be conducive to survival. Nevertheless a stream of thought could be censored before it elicits harmful behavior. Streams of thought probably get blocked or sidetracked on a regular basis, not just by censors, but by minute-to-minute undulations in the hyper-dimensional fitness landscape,

that is, change in the relative urgency of the multitude of survival-related or derived needs impacting the focus.

The concept of cultural momentum sheds light on the issue of free will. Those who argue for the existence of a central executive in memory may come to be viewed as the creationists of philosophy and cognitive science. Human will can instead be viewed as the emergent orchestration of needs, stimuli, and retrieved memories impacting the focus, which is subject to cultural momentum and therefore, in a sense, beyond our control.

## 10.4 THE BIRTH OF IDEAS

The biologically-inspired model developed here supports a variant of the combination theory of creativity that new ideas arise through combinations and transformations of old ones [Boden 1991; Koestler 1964]. The aspect of this theory that does not ring true is that it neglects the role of emotions. Here we consider emotions, as well as ideas, to be encoded as information in memes; some components of a meme are simply interpreted by parts of the mind that experience them as emotion, whereas others are interpreted by parts of the mind that experience them as ideas. Much research on analogy deals with how the structure or conceptual skeleton underlying one idea gets abstracted and applied to another [Gentner 1983; Gick & Holyoak 1983]. We can expand on this general idea by suggesting that many forms of creative expression begin with the (unvoiced) question: What would the pattern of information that encoded the emotion I experienced during this particular event look like if expressed through the constraints of that medium? The existence of inherent limitations on how a pattern could translate from one domain to another is consistent with the frequent observation that creativity involves both freedom and aesthetic constraint. Thus all creativity is directly or indirectly derived from experience in the world, and since the mathematics underlying this world, the set of all natural functions, is a small subset of all &possible functions, the constraints that guide creation are not arbitrary but objective and familiar; for example the drum beat of a song might echo a heartbeat, when the rhythm and chord progression are reminiscent of the sound of someone sobbing we feel sad, and we hear the wrong note even if we have never heard the song before.

It makes sense to expect that a meme or meme complex that has been censored (for example because it evoked unbearable sadness or anger) would be vulnerable to being targeted as an area where worldview cohesion could be increased. Since at the time the censored material was experienced it was prohibited from forming associations to obviously-related memes, it in turn can not be retrieved through these expected or straightforward associations. It can only be retrieved via backdoor entrances, which are associations that reflect structural congruity at an abstract level. Thus a musician may come to habitually funnel patterns encountered in a variety of domains—and particularly censored material—through modules that filter out hitchhiker and enabler features, and adapt the core features (or feature schemata) to the constraints of music. It is in this repackaged format that memes originally deemed dangerous can be integrated into the memory at large without harm, and it is through this process that the creator establishes a sense of control over memes that were previously off-limits.

The account proposed here may seem too simple to explain the seemingly limitless human potential for creativity, but it may seem less far-fetched when we consider the variety of species

produced by biological evolution, which operates without the benefit of strategy. Furthermore, raw materials for the creative process may be acquired in exceedingly subtle ways. It is conceivable that you might watch a stream flow and without your consciously thinking, 'It flows... things can flow... I could even, in some sense, adopt a more flowing approach to life', the experience might be reconfiguring your memetic infrastructure in a way that makes you more easygoing. I am not making any claims about the extent to which experiences of this sort affect us or even whether they occur at all, but rather suggesting that we not prematurely place a lid on the kinds of processes that could affect a network of representations and thereby affect creation and transmission.

## 10.5 MEMETICS AS MISSING LINK BETWEEN SCIENCE AND SPIRIT

Recall that to implement a meme is to express it, so that it crystallizes from the world of ideas into words or body language or objects in the physical world. Memes fool potential hosts into wanting or identifying with them, by aligning themselves with memes we already hold dear (as advertisers are well aware). The more we value ourselves in terms of the memes and implemented artifacts we possess or lack, the more vulnerable we are to ever-more-seductive forms of persuasion and advertising which tie up time, energy, and resources that could be applied toward other goals.

One way to defend oneself against painful or manipulative memes is to construct what Dennett [1995] refers to as a meme-immunological system; that is, formulate new memes specifically to deflect memetic antigens. However constructing memetic antibodies of this sort is time-consuming, and like any immunological response it has to be repeated every time the outside agent evolves a counter-response.

Perhaps this explains the purported benefits of transcending the ego [e.g. Walsh & Vaughan 1993], which can be taken to mean getting in touch with who we were before our minds were colonized by memes, through practices such as meditation. These practices may work by giving the brain time to anneal material that was never fully assimilated because of distraction or censorship--mend flaws in the fabric of the individual's worldview. Release from the restrictive power of censors may then enhance the individual's raw 'sense of self' and produce a feeling of unity or one-ness.

It is almost impossible to discuss creative insight or spiritual unfoldment without drawing upon metaphors involving light. For example: her face lit up, he glowed with enthusiasm, moment of illumination, reflect on an idea, dim-witted person, and dark night of the soul, to mention a few. Even nonverbal communication uses these metaphors; everyone knows that when a light bulb appears above Charlie Brown's head he has had a 'bright' idea. Experiences of inner light have been documented since the dawn of civilization. Some cultures have a name for it; Eskimo shamen call it qaumaneq, while Vedanta Hindus refer to it as Atma. Lao-tzu advised his people "Use the light that is within you". The Gnostics described flashes of insight or discovery as "light sparks" or scintillae, and they conceived of the soul as a radiant sphere that is prone to contamination, in which case it darkens and becomes more material. Indra's Net is a Buddhist allegory about a net of threads stretching horizontally through space and vertically through time, and at every crossing of the threads is a crystal bead that represents an individual. Buckminster

Fuller spoke of the soul as a bright sphere of awareness that moves wherever the individual moves. Jung suggested we think of ego-consciousness as "surrounded by a multitude of little luminosities".

Since we know the physical world largely through how light of various frequencies reflects off objects, it is perhaps natural that we use something analogous to also make sense out of the world of abstract ideas and emotions. In diagrams of SDM or neural networks, when a meme or meme feature becomes active it is generally represented as lit up. These diagrams may be derived from the same kind of intuitive understanding as the mystical experiences mentioned above. They may turn out to be early prototypes for a portrayal of spiritual knowledge that is real in the same way as a pie chart of business expenses represents something real. We may be on the doorstep of an era when spiritual experience not only rings true, but makes logical sense as well.

## **10.6 MEMETICS AND WOMENUS STUDIES**

One of the things I hope the study of memetics will one day shed light on is the cumulative effect on women's self esteem of encountering the default 'he' or 'his', or phrases like 'man and his world'. Each time it happens, it reinforces, however subtly, the thought "I'm not one of the important ones", or "that doesn't pertain to me so I don't really count". The effect of this may be insignificant the first few, or even hundred times, it happens. But by the time a young girl reaches adulthood, experiences of this kind probably number in the millions, and that may be when the effects start to kick in.

## **11. A MEMEUS-EYE VIEW**

In this section we examine a hypothetical and admittedly speculative scenario for how evolutionary concepts borrowed from biology might apply to the dynamics of a specific meme.

### **11.1 THE ONTOGENY OF A MEME**

One day a classroom bully named Tony put his arm around a girl named Memela. Memela felt threatened by Tony's advance. Her first impulse was to get angry but she censored this reaction. The need for worldview cohesion motivated a desire to escape the restrictive power of this censor and find a backdoor vent for her anger. Eventually she whispered to a classmate: "Tony Testosteroni made a pass at me." Thus began the era of Memela's meme.

Memela's meme can be traced back to a number of precursor or protomemes, many of which originated in the minds of family and friends and were subsequently transmitted to Memela. These protomemes provided an environmental niche in which the joke could flourish. But Memela was more prone to this kind of humor than her siblings. A sort of conceptual speciation had taken place in their brains, wherein small initial differences were amplified through positive feedback, and Memela had become known as the jokester.

Perhaps Memela was predisposed to produce the joke because her brain spontaneously exerts a high degree of control over its activation threshold. Her ability to find unusual associations by

increasing hypersphere radius, and subsequently refine a train of thought by decreasing the radius, facilitates word play such as the establishment of epistatic relationships between semantic and syntactic features of a meme. The semantic applicability of the word testosterone to Tony's aggressive behavior, and the alteration of this word to make it sound Italian and echo the syntax of (i.e. rhyme with) the Italian first name, contribute to the humor of Memela's meme. Its relation to the highly censored subjects of aggression and sexuality may also have added to its appeal.

## 11.2 A DAY IN THE LIFE OF MEMELA'S MEME

This kind of conceptual epistasis provokes laughter, which draws attention to a meme. Memela's meme took full advantage of this. When Memela told the joke to one of her classmates the ensuing laughter attracted a small crowd of other potential hosts, and within an hour Memela's meme had reached most of the students in the classroom. Their willingness to invest time acquiring this meme was a smart move; the joke not only provided amusement, but it proved to be a useful precursor to the formulation and understanding of subsequent jokes in this social circle. Some were direct descendants of Memela's meme, such as jokes along the lines of "What's for lunch--rigatoni a la testosteroni?" Others were more distantly related, such as nicknames for other classmates that arose because Memela's meme had activated the general concept of nicknaming. Across the classroom, ideas that pertained to nicknaming came to constitute a highly active region of conceptual space.

By the time recess ended it had migrated extensively through the school population. There were certain subpopulations of individuals it failed to penetrate, such as social outcasts who were excluded from much of the memetic exchange. These individuals' minds exhibited a cultural version of the Founder Effect [Holgate 1966]--reduced variation in a small population due to genetic drift. Memela's meme also failed to reach students who engaged in projects that took them away from the playground at recess. However since these individuals had had less opportunity to witness Tony's behavior, they had less need to diffuse their fear of him. Thus even if they had heard Memela's meme they did not possess the necessary precursor memes to fully appreciate it; the prerequisite memetic niches were not in place. At any rate despite its failure to reach these subgroups, Memela's meme experienced a high degree of memetic fitness. It migrated far beyond the classroom in which it was originally formulated, reconfiguring networks of representations in ways that affected the subsequent thought and behavior of a number of individuals. The telling of this meme and its various incarnations constituted an act of memetic altruism between like-minded individuals analogous to the biological altruism that occurs between genetically-similar individuals. It played a small role in an ongoing network of positive reciprocal interactions through which there emerged a memetically-derived social structure wherein individuals that regularly generated pleasurable or powerful memes came to be observed carefully and imitated frequently, while other individuals were ignored. Thus the fate of Memela's meme and its descendants reflected the social and psychological dynamics of an entire society of interacting individuals.

Like the other students in the classroom, Tony was affected by the sound of laughter advertising the presence and amusement value of Memela's meme. However it wasn't until someone told him that it was a joke was about him that he felt willing to do almost anything to hear it. Upon

hearing the nickname Tony felt ridiculed. In the case of Memela's meme a meme-immunological system could be something along the lines of "That nickname is silly and stupid."

## **12. CAN MEMETIC EVOLUTION PROVIDE A SYNTHETIC FRAMEWORK FOR THE COGNITIVE AND SOCIAL SCIENCES?**

This chapter gives us a taste of how the concept of evolution may be able to provide the kind of overarching framework for the humanities that it provides for the biological sciences. It sketches out a theoretical framework for cultural evolution, adopting an approach analogous to that of the population geneticist in that it emphasizes meme evolution through social interaction and de-emphasize individual hosts. Although the cultural evolution of memes operates through very different mechanisms from those of biology, culture is the only system comparable to biology, because it is the only other system to exhibit the imperative features of evolution adaptive exploration and transformation of an information space through variation, selection, and transmission. All patterns in the information we encounter can be traced to either (1) the physical constraints and self-organizing properties of matter, (2) biological evolution, (3) cultural evolution, or (4) interactions between these causal principles.

One important difference between the two forms of evolution is that culture is less random--new patterns have a greater-than-chance probability of being more fit than their predecessors. The reason for this is interesting. Since memes (unlike genomes) do not come packaged with instructions for their replication, they must rely on the pattern-evolving machinery of our brains to do it for them. Ironically, this state of dependence enhances their proliferative potential, because the machinery they depend upon constructs and continually updates mental models of its world--that is, weaves the memes into a worldview--and uses this worldview to enhance the assimilation and implementation of the memes and their offspring.

Cultural evolution presents a puzzle analogous to the origin of life: the origin of a self-sustained stream of potentially-creative thought in an infant's brain. The idea that life originated with the self-organization of a set of autocatalytic polymers suggests a possible mechanism for how this comes about. Once a threshold density of assimilated memes is surpassed, any meme that occupies the focus is close enough in Hamming distance to evoke or 'catalyze' the spontaneous retrieval or creative reconstruction of a statistically similar meme, thus the memes form an autocatalytic set. Note that this macroscopic account suggests an explanation for only that aspect of human consciousness that differentiates us from other 'experiencers'; it does not address the mystery of 'raw awareness' that some say characterizes not only our experience but that of a cow or a mosquito or even a thermostat [e.g. Chalmers 1996]. Whether or not this specific theory turns out to be correct, it illustrates how the analogy to biology can focus our study of culture by providing a scaffold around which explanatory theories can be built.

The cultural framework for cognition developed here suggests tentative explanations for psychological phenomena such as censorship and creativity, and why they are virtually unique to humans. It also suggests explanations for various social phenomena.

There is good reason to view the mind as a memetic host—a generator, selector, and replicator of cultural information—and society as a continually-regenerated constellation of such hosts.

Interactions in the world affect our memetic structure, which in turn affects future interactions, and these interactions allow cultural information to evolve, recursively and in parallel in its many hosts. With each new generation of hosts, the conceptual infrastructure from which memes are born and grow reflects how memes were encoded and implemented by the previous generation. I cannot imagine a more exciting role to play in this process than partaking in the exploration of memes about memes.

## ACKNOWLEDGMENTS

I would like to thank David Chalmers, Anthony Francis, Peter Greene, Pentti Kanerva, Ashwin Ram, Bruce Sawhill, Roger Schank, Patrick Tufts, and Philip Van Loocke for thoughtful comments on the manuscript.

## REFERENCES

- Alexander, R. D. (1980) Darwinism and human affairs, Pitman.
- Ackley, D. (1994) A Case for Distributed Lamarkian Evolution. In Artificial Life III, Addison-Wesley.
- Bak, P., Tang, C., & Weisenfeld, K. (1988) Self-organized criticality. Phys. Rev. A 38: 364.
- Bateson, G. (1972) Steps to an ecology of mind, Ballantine Books.
- Begin, D.J. & Aquadro, C.F. (1992) Levels of naturally occurring DNA polymorphism correlate with recombination rates in *D. melanogaster*. Nature, 356, p. 519- 520.
- Boden, M. A. (1991) The creative mind: Myths and mechanisms, Cambridge University Press.
- Bonner, J. T. (1980) The evolution of culture in animals, Princeton University Press.
- Bowers, K.S., Regehr, G., Balthazard, C. & Parker, K. (1990) Intuition in the context of discovery. Cognitive Psychology, 22:72-110.
- Boyd, R. & Richerson, P. J. (1985) Culture and the evolutionary process, The University of Chicago Press.
- Brooks, V. (1986) The neural basis of motor control, Oxford University Press.
- Campbell, D. T. (1987) Evolutionary epistemology. In Evolutionary epistemology, rationality, and the sociology of knowledge, eds. G. Radnitzky and W. W. Bartley, 4 7- 89, Open Court.
- Cavalli-Sforza, L. L. & Feldman, M. W. (1981) Cultural transmission and evolution: A quantitative approach, Princeton University Press.

- Chalmers, D.J. (1996) *The conscious mind: In search of a fundamental theory*, Oxford University Press.
- Clark, A. & Thornton, C. Trading spaces: Computation, representation, and the limits of uninformed learning. Forthcoming in *Brain and Behavioral Sciences*.
- Darwin, C. (1871) *The descent of man*, John Murray Publications.
- Dawkins, R. (1976) *The selfish gene*, Oxford University Press.
- Dawkins, R. (1982) *The extended phenotype*, W. H. Freeman and Company.
- Dennett, D. C. (1995) *Darwin's dangerous idea*, Little, Brown and Company.
- Donald, M. (1991) *Origins of the modern mind*, Harvard University Press.
- Dunbar, K. & Schunn, C.D. (1990) The temporal nature of scientific discovery: The roles of priming and analogy. In: *Proceedings of the Twelfth Annual Meeting of the Cognitive Science Society*, 93-100.
- Durham, W.H. (1991) *Coevolution: genes, culture, and human diversity*. Stanford University Press.
- Fehrer, E. & Raab D. (1962) Reaction time to stimuli masked by metacontrast. *Journal of Experimental Psychology*, 63, 143-147.
- Gabora, L. M. & Colgan, P. W. (1990) A model of the mechanisms underlying exploratory behaviour. In *The simulation of adaptive behavior*, eds. S. Wilson and J. A. Mayer, MIT Press.
- Gabora, L. M. (1995) Meme and variations: A computational model of cultural evolution. In *1993 Lectures in Complex Systems*, Addison-Wessley.
- Gentner, D. (1983) Structure-mapping: A theoretical framework for analogy. *Cognitive Science*, 7(2).
- Gick, M. L. & Holyoak, K. J. (1983) Schema induction and analogical transfer. *Cognitive Psychology*, 15, 1-38.
- Gould, S. J. (1991) *Bully for brontosaurus: reflections in natural history*. W. W. Norton & Company. p. 63-66.
- Gould, S. J. & Vrba, E. S. (1982) Exaptation--A missing term in the science of form. *Paleobiology*, 8(1), 4-15.
- Greene, P. J. (1978) From genes to memes? *Contemporary Sociology* 7, 06-709.

- Guilford, J. P. (1979) Some incubated thoughts on incubation. *Journal of Creative Behavior*, 13, 1-8.
- Hebb, D. (1949) *The organization of behavior*, Wiley and Sons.
- Hofbauer, J. & Sigmund, K. (1988) *The theory of evolution and dynamical systems*, Cambridge University Press.
- Hofstadter, D. R. (1985) Variations on a theme as the crux of creativity. In *Metamagical Themes*, 232-259, Basic Books.
- Holgate, P. (1966) A mathematical study of the founder principle of evolutionary genetic. *J. Appl. Probl.*, 3, 115-128.
- Holland, J. K. (1975) *Adaptation in natural and artificial systems*, University of Michigan Press.
- Holland, J. H., Holyoak, K. J., Nisbett, R. E. & Thagard, P. R. (1986) *Induction*, MIT Press.
- Hopfield, J. J. (1982) Neural networks and physical systems with emergent collective computational abilities. *Proceedings of the National Academy of Sciences (Biophysics)*. 79(8), 2554-2558.
- Hubel, D. H. & Wiesel, T. N. (1979) Brain mechanisms and vision. *Scientific American*, 241(3), 150-62.
- Hughes, D.M. & Nei, M. (1988) Patterns of nucleotide substitution at major histocompatibility complex class I loci reveals overdominant selection, *Nature*, 335, 167-170.
- Hull, C. L. (1943) *Principles of behavior*, Appleton-Century- Crofts.
- Kanerva, P. (1988) *Sparse distributed memory*, MIT Press.
- Kaplan, N. L., Hudson, R. R., & Langley, C. H. (1989) The hitchhiking effect revisited. *Genetics*, 123, 887-899.
- Karmiloff-Smith, A. (1986) From meta-processes to conscious access: Evidence from children's metalinguistic and repair data. *Cognition* 23, 95-147.
- Karmiloff-Smith, A. (1990) Constraints on representational change: Evidence from children's drawing. *Cognition* 34, 57-83.
- Karmiloff-Smith, A. (1992) Beyond modularity: A developmental perspective on cognitive science, MIT Press.
- Kauffman, S. (1993) *Origins of order*, Oxford University Press.

- Keeler, J. D. (1988) Comparison between KanervaUs SDM and Hopfield-type neural networks. *Cognitive Science* 12, 299-329.
- Klimesch, W. (1995) Memory processes described as brain oscillations. *Psycoloquy*, 6.06.
- Koestler, A. (1964) *The act of creation*, Picador.
- Kojima, K. & Schaeffer, H. E. (1967) Survival processes of linked mutant genes. *Evolution*, 21, 518-531.
- Kubose, S. K. & Umenoto, T. (1980) Creativity and the Zen koan. *Psychologica*, 23(1), 1-9.
- Langton, C. G. (1992) Life at the edge of chaos. In *Artificial life II*. eds. C. G. Langton, J. D. Farmer & S. Rasmussen, Addison-Wesley.
- Lenat, D.B. (1974) The ubiquity of discovery. *Artificial Intelligence Journal*, 9, 257-86.
- Luchins, A.S. (1942) Mechanization in problem solving. *Psychological Monographs*, 54, No. 248.
- Lumsden, C. & Wilson, E. O. (1981) *Genes, mind, and culture*, Harvard University Press.
- Lynch, A. (1996) *Thought Contagion: How Belief Spreads Through Society*, Basic Books.
- Mandelbrot, B. B. (1982) *The fractal nature of geometry*. W. H. Freedman and Company.
- Marr, D. (1982) *Vision*, Freeman.
- Maynard Smith, J. & Haigh, J. (1974) The hitchhiking effect of a favourable gene. *Genetic Research*, 23, 23-35.
- McCulloch, W. S. & Pitts, W. (1943) A logical calculus of the ideas immanent in nervous activity. *Bulletin of Mathematical Biophysics*, 5, 115-133.
- McFarland, D. J. & Sibly, R. M. (1975) The behavioural final common path. *Philosophical transactions of the London Royal Society* 270B, 265-93.
- Miller, G. A. (1956) The magic number seven, plus or minus two: Some limits on our capacity for processing information. *Psychological Review*, 63, 81-97.
- Minsky, M. (1985) *The Society of Mind*. Simon and Schuster.
- Nersessian, N. (1993) In the theoretician's laboratory: Thought experimenting as mental modeling. In PSA 1992, vol. 2, eds. D. Hull, M. Forbes, & K Okrulik. PSA.

Palmer, M. W. (1992) The coexistence of species in fractal landscapes. *American Naturalist*, 139, 375-397.

Pinker, S. (1995) *The language instinct*. HarperPerennial.

Plotkin, H. C. (1988) *The role of behavior in evolution*. MIT Press.

Popper, K. R. (1963) *Conjectures and refutations*, Routledge & Kagen.

Pribram, K. H. (1974) How is it that sensing so much we can do so little? In *The neurosciences third study program*, eds. F.O. Schmidt & F.G. Worden, MIT Press.

Quillian, M. R. (1968) Semantic memory. In *Semantic information processing*, ed. M. Minsky, MIT Press.

Ram, A. (1993) Creative conceptual change. *Proceedings of the Fifteenth Annual Conference of the Cognitive Science Society*, 17-26.

Ray, T. (1991) An approach to the synthesis of life. In *Artificial life II*. eds. C.G. Langton, C. Taylor, J.D. Farmer & S. Rasmussen, Addison-Wesley.

Robert, M. (1990) Observational learning in fish, birds, and mammals: A classified bibliography spanning over 100 years of research. *Psych Record* 40, 289-311.

Rumelhart, D. E. & McClelland, J. L. eds. (1986) *Parallel distributed processing*, Bradford/MIT Press.

Schank, R.C. (1983) *Dynamic memory*, Cambridge University Press.

Schank, R.C. & Leake, D.B. (1989) Creativity and learning in a case-based explainer. *Artificial Intelligence*, 40(1-3): 353-385.

Scheuring, I. & Riedi, R. H. (1994) Application of multifractals to the analysis of vegetation pattern. *Journal of Vegetation Science*, 5, 489-496.

Schuster, P. & Sigmund, K. (1983) Replicator dynamics. *Journal of Theoretical Biology* 100, 533-38.

Shannon, C.E. & Weaver, W. (1963) *The mathematical theory of communication*. University of Illinois Press.

Smith, W. J. (1977) *The behavior of communicating*, Harvard UniversityPress.

Spector, L. & Luke, S. (1996) Culture enhances the evolvability of cognition. In *Proceedings of the 1996 Cognitive Science Society Meeting*.

- Squire, L. R. (1992) Memory and the hippocampus: A synthesis from findings with rats, monkeys, and humans. *Psychological Review*, 99, 195-231.
- Thagard, P. (1980) Against evolutionary epistemology. *PSA 1980*, ed. P. D. Asquith & R. N. Giere 187-96.
- Tomasello, M. Kruger, A. C., & Ratner, H. H. (1993) Cultural learning. *Behavioral and Brain Sciences*, 16, 495-552.
- Van Loocke, P. (1991) Study of a neural network with a meta- layer. *Connection Science*, Vol. 3, No. 54.
- Von Neumann, J. (1966) *Theory of self-reproducing automata*, University of Illinois Press.
- Vygotsky, L. S. (1978) *Mind in society: The development of higher psychological processes*, eds. M. Cole, V. John-Steiner, S. Scribner & E. Souberman, Harvard University Press.
- Wallas, G. (1926) *The art of thought*, Harcourt, Brace & World.
- Walsh, R. & Vaughan, F. (1993) *Paths beyond ego*, Jeremy P. Tarcher/Perigee.
- Weisberg, R.W. (1986) *Creativity: Genius and other myths*, Freeman.
- Wright, S. (1969) *Evolution and the genetics of populations*, University of Chicago Press.

## **FOOTNOTES**

1. Note that by "causal principle" I mean something that generates useful descriptions, rather than a law.
2. Though viruses are unique in the biological world in that they rely on hosts to replicate, we will consider viral evolution an anomalous offshoot of biological evolution, because: (1) the evolving patterns of information are encoded as sequences of nucleotides, (2) variation is through mutation and recombination, and (3) transmission and selection are mediated through genotype.
3. Or one could argue that the 'selection' of matter over antimatter, and its subsequent amplification and variation, constitutes yet another form of evolution.
4. This example is a variation of one transmitted by R. Boyd [pers. com.]

## **MEMES: THE CREATIVE SPARK**

**By**  
**Liane M. Gabora**

Memes, unlike genes, do not come packaged with instructions for their replication; they rely on the pattern-evolving machinery of our brains. We tend to replicate memes that satisfy needs - biological needs like food, shelter, and sex, or needs that contribute less directly to survival of self and progeny. These include the need for love, and the need for a coherent internal model of the world, something we can call upon whenever a situation is too complicated for our hardwired instincts.

This worldview weaves each new instant of experience into its tapestry of associations. Mental censors, however, can disrupt the assimilation of memes that threaten the individual's ego or belief structure. This prevents the brain from forging associations between the new meme and previously stored memes. Conversely, insightful memes sometimes trigger a chain reaction that modifies their host's entire worldview - a conceptual phase shift. Since brains are wired so that related memes trigger one another domino style, cultural evolution, like biological evolution, has built-in momentum. We control the birth of "our" memes only to the extent that we provide a fertile ground for them to grow and multiply.

The brain's tendency to play with its memes, continually reorganizing and ironing out arbitrary associations to forge more meaningful ones, often creates unanticipated solutions to pressing problems. This suggests that creativity, the fountainhead of cultural variation, is strategic - not random, as are the variation-generating processes in biology.

Another consequence of meme play is that knowledge or emotion encoded in one kind of experience can be translated into another kind of experience. The tension produced by censored sexual material, for example, might be diffused in a joke. Or a musician may come to habitually funnel memes encountered in all types of situations - particularly censored material - through brain modules that filter out domain-dependent elements and adapt the core components to the constraints of music. It is in this repackaged format that memes are more fully integrated into their host's memory, and it is through this process that the creator establishes a sense of control over memes that were previously off-limits. Creativity is thus directly or indirectly derived from experience in the world, and since the mathematics underlying this world - the set of all natural functions - is a small subset of all possible functions, the constraints that guide creation are not arbitrary. The drumbeat of a song might echo a heartbeat, and when rhythm and chord progression suggest the sound of someone sobbing, we feel sad.

Memes, as advertisers are well aware, can fool potential hosts into believing they are needed by associating with memes we already identify as necessary. It takes time to degrade these unwarranted associations and assimilate memes that were previously censored - in other words, to mend flaws in the fabric of the individual's worldview. Thus the power of meditative practices: release from the isolating restriction of censors creates a feeling of oneness. Transcending the ego can be taken to mean getting in touch with that part of ourselves that

existed before our minds were colonized by memes, often visualized as a spark, halo, or sphere of light.

If each field of knowledge is a pile of sand, the overlapping sandpiles of our era have begun to fuse into one big block. Perhaps the 21st century will see the entire slab of human knowledge stood on its side, and with the help of a new breed of sciences that are not just interdisciplinary but transdisciplinary, we will carve slices that run perpendicular to all the traditionally defined disciplines, encompassing part of each. Memetics appears not only to put us on the road to understanding the pervasiveness, diversity, and adaptive complexity of the cultural debris that surrounds and infests us. It also yields unexpected insight into creativity and spiritual matters that have mystified us since the first fledgling memes appeared in our ancestors' brains.

*Liane M. Gabora (liane@cs.ucla.edu), a research fellow at UCLA's Center for the Study of the Evolution and Origin of Life, serves on the editorial board of the Journal of Memetics ([www.fmb.mmu.ac.uk/jom-emit/](http://www.fmb.mmu.ac.uk/jom-emit/)). She is working toward putting together an animation/artificial life/motion capture/dance laboratory for applied research into the oldest form of memetic exchange: body language.*

# THE ORIGIN AND EVOLUTION OF CULTURE AND CREATIVITY

By

Liane Gabora

Center for the Study of the Evolution and Origin of Life,  
University of California,  
Los Angeles Los Angeles, CA 90095-1567, USA.  
[lgabora@vub.ac.be](mailto:lgabora@vub.ac.be)

## Abstract

Like the information patterns that evolve through biological processes, mental representations, or memes, evolve through adaptive exploration and transformation of an information space through variation, selection, and transmission. Since unlike genes, memes do not come packaged with instructions for their replication, our brains do it for them, strategically, guided by a fitness landscape that reflects both internal drives and a worldview that is continually updated through meme assimilation. This paper presents a model for how an individual becomes a meme-evolving agent via the emergence of an autocatalytic network of sparse, distributed memories, and discusses implications for complex, creative thought processes and why they are unique to humans. Memetics can do more than account for the spread of catchy tunes; it can pave the way for the kind of overarching framework for the humanities that the first form of evolution has provided for the biological sciences.

**Keywords:** autocatalysis, creativity, culture, evolution, imitation, induction, information, meme, memory, mental representation, natural selection, social learning.

## 1 Culture as a Second Form of Evolution

While some ideas instantly fade into obscurity, others spread horizontally through society, and vertically from one generation to another, getting progressively refined and embellished along the way. Thus ideas, like the strands of DNA that encode instructions for building and maintaining living organisms, seem to undergo a process analogous to biological evolution. Accordingly there has been a slow but steady effort to map the concept of evolution onto the dynamics of culture. Popper [72] and Campbell [11] alerted us to the evolutionary flavor of epistemology. Dawkins [17] introduced the notion of a meme - a replicator of cultural information analogous to the gene. In his words: "Just as genes propagate themselves in the gene pool by leaping from body to body via sperm or eggs, so do memes propagate themselves in the meme pool by leaping from brain to brain." Others have drawn from mathematical models of population genetics and epidemiology to model the spread of ideas (Cavalli-Sforza & Feldman [12], Lumsden & Wilson [56], Schuster & Sigmund [84], Boyd & Richerson [9], Hofbauer & Sigmund [33]).

These works point toward the possibility that memetics constitutes a second form of evolution, distinct from yet intertwined with biological evolution, with the potential to provide the kind of overarching framework for the social and cognitive sciences that the first form provides for the

biological sciences. However thus far memetics has not lived up to this potential, a situation that seems unfortunate given the success of the biological precedent. Although much was known about living things before Darwin, his theory of how life evolves through natural selection united previously disparate phenomena and paved the way for further biological inquiry.

Some believe that looking to biological evolution to gain insight into cultural evolution is a waste of time. As Gould [28] put it: "Biological evolution is a bad analog for cultural change... biological evolution is powered by natural selection, cultural evolution by a different set of principles that I understand but dimly." However at an abstract level of analysis they amount to the same thing: exploration and transformation of an information space through variation, selection, and replication. Both present the question of what underlying mechanisms could *launch* a self-perpetuating adaptive process. Thus the possibility that the two have enough in common that the former can pave the way for the latter seems at least worth pursuing. Skeptics may wonder how we can develop a theory of cultural evolution before we understand how memes are instantiated in the brain. This situation has a precedent: Darwin came up with the theory of biological evolution through natural selection before the discovery of genes. It turned out that genes are laid out in a fairly straightforward way in physical space, which does not appear to be the case with memes. This does not mean they can't evolve, so long as there is a way of retrieving the components of a meme so they can work together as a unit. We may not yet know all the physiological details of how the information manifested in, say, a handshake between two individuals - with its unique arrangement of contact points, applied forces, and trajectory - can be traced back to these individuals' mental representations of handshakes, each other, and the situation they are in. But let us proceed with the confidence that a solution exists and can be found.

This paper outlines a theory of how memes evolve, and illustrates how the memetic perspective provides not only not only a foundation for research into the dynamics of concepts and artifacts at the societal level, but a synthetic framework for understanding how mental representations are generated, organized, stored, retrieved, and expressed at the level of the individual. It also sketches a tentative theory of how an infant develops a sustained train of potentially-creative thought and thus becomes a cog in the meme-evolving machinery. Implications of this theory pertaining to the mechanisms underlying creativity, and why it is virtually nonexistent in other species, are explored. It concludes with discussion of how a cultural-evolution perspective can shape and inspire research in the cognitive and social sciences.

## 1.1 Components of an Evolutionary System

In order for evolution to happen there must be:

- A *pattern of information* (a state within a space of possible states)
- A way to generate *variations* of the pattern (explore or transform the space)
- A rationale for *selecting* variations that are adaptive - tend to give better performance than their predecessors in the context of some problem or set of constraints (a fitness landscape applied to the space)
- A way of *replicating and transmitting* (or amplifying, as molecular biologists refer to it) the selected variations

In biological evolution, the evolving patterns of information are genes encoded as sequences of nucleotides. Variations arise through mutation and recombination, and natural selection weeds out those that are maladaptive. Replication takes place at the level of the genotype. In cultural evolution, the evolving patterns of information are memes - mental representations of ideas, behaviors, or other theoretical or imagined constructs, perhaps encoded as patterns of neuron activation. Variations are created by combining, transforming, and reorganizing representations, consciously or unconsciously, or through errors in transmission. Replication is phenotypically mediated; it occurs when representations are transformed into action or language, transmitted through processes such as imitation, and reproduced, more or less, in another brain.

Incorporation of these new information patterns into the society alters the selective pressures and constraints exerted by the social environment, which in turn leads to the generation of yet more patterns. Thus memes, like DNA, comprise a self-sustained system for the relentless exploration and transformation of a space of possible patterns.

## 1.2 Culture is Not Just a Predictable Extension of Biological Evolution

The line of reasoning presented here can be succinctly conveyed in terms of *information*, which is related to the number of differences required to specify the state of a system (Shannon & Weaver [85], Bateson [4]). States have not only a structure of difference relations between them, but also a combinatorial structure - each state can itself comprise an information space, so that complex information can be built up from simple information. We can rate each state in a space of possible states against some performance measure or fitness criterion and the result is referred to as a *fitness landscape*, and we can move from one state to another by way of a *computation*. The world can be viewed as a vast network of computations wherein information is created, transformed, and destroyed. This information often exhibits *pattern*, or statistical regularity that can be expressed mathematically.

After seeing many shadows cast by the same object we can develop an internal model of what that object looks like without having seen it, and if more than one object is casting shadows we can learn to tell which object is casting any particular shadow. Similarly, by viewing every pattern we encounter as a shadow or footprint of one or more broad causal principles [note 1], we can gain insight into the causal principles that manifest that pattern.

If you were to go back to sometime during the first billion years of Earth's history, the only causal principle you would need to invoke to explain pattern in the information present (with the exception of yourself) would be the *physical constraints and self-organizing properties of matter*.

If you were to go back to sometime *after* the origin of life, approximately three billion years ago, this would no longer be the case. Not that life doesn't exhibit the properties of matter. But it would be virtually impossible for, say, a giraffe to appear in an information space not acted upon by natural selection. Another causal principle - *biological evolution* - would have to be invoked from this point on.

Today the Earth is embedded with artifacts like computer networks and circuses that cannot be accounted for by appeal to either the properties of matter or biological evolution. That is, biological evolution does not provide us with adequate explanatory power to account for the existence of computers any more than the properties of matter can explain the existence of giraffes. Computers are manifestations of yet another causal principle: *the evolution of culture*. Thus pattern in the structure and dynamics of information we encounter in the everyday world can be traced to three broad causal principles - the physical constraints and self-organizing properties of matter, biological evolution, and cultural evolution. This classification scheme, like all classification schemes, is somewhat arbitrary. There may be subclasses of these principles that deserve to be considered principles unto themselves [note 2], or one could argue that evolution is a self-organizing property of matter, albeit a spectacular one [note 3]. The point is: culture is the only process that has arisen since the origin of life that relentlessly exploits the combinatorial potential of information. Despite the fact that culture is grounded in biology (like biology is grounded in the physical constraints and self-organizing properties of matter), the probability of computers arising spontaneously in an information space not acted upon by cultural evolution (like the probability of giraffes arising spontaneously in an information space not acted upon by biological evolution) is vanishingly small. Thus it is inappropriate to dismiss culture as a predictable extension of biological evolution. It is qualitatively different from anything else biology has produced.

Since the machinery that renders cultural evolution - the human brain - is a product of biological evolution, much of what is 'out there' cannot be cleanly traced to a biological or cultural origin. We will discuss how biology constrains culture through the preferential spread of memes that satisfy biologically-derived needs. It goes the other way too; culture not only affects biological fitness through its effect on behavior (a phenomenon known as the Baldwin Effect) but it dramatically modifies the biological world. Some of the ways in which biological information gets tainted with cultural information seem relatively inconsequential, such as the trimming of hedges, whereas others, such as dog-breeding, have a long-lasting effect. (In fact, one could view dogs as the consequence of a memetic trajectory that was launched by the need to protect property.) Nevertheless, much as it is not imperative to address the role of physical constraints like gravity in studies of, say, embryonic development or squirrel foraging behavior, much can be said about culture without addressing the role of biological constraints.

### 1.3 Taking the Meme-Evolving Apparatus Seriously

The most salient shortcoming of existing models of cultural evolution, and quite possibly the reason they have had relatively little impact, is that although they tell us much about the transmission or characteristics of 'catchy' memes, they fail to adequately address the issue of how novelty is generated. To be of significant theoretical or predictive value memetics must give serious consideration to the processes by which experience in the world turns into new memes in our brains, and address how memetic novelty is expanded further through creative processes. Studies of creativity, on the other hand, have focused on the *individual*, obscuring the fact that creativity is a collective affair e.g. (Lenat [54], Schank & Leake [82], Mitchell [67], Boden [6]). The ideas and inventions an individual produces build on the ideas and inventions of others (the ratchet effect). Which memes spread and which ones die off reflects the dynamics of the entire society of individuals hosting them.

Thus although at a sufficiently abstract level the notion that culture evolves is obvious, we lack a theoretical framework that bridges transmission studies with studies of creativity, and spells out explicitly how the concept of evolution maps on to the case of culture. To accomplish this we need more than a quick and dirty list of what makes a meme 'catchy'; we have to consider the extent to which it resonates with and enriches the complex web of assumptions, beliefs, motives, and attitudes of its host - no simple matter. In short, we have to take the meme-evolving apparatus very seriously.

## **2 Meme: the Unit of Information that Evolves Through Culture**

Because the concept of the individual is so germane to the human experience we have an anthropocentric tendency to assume that the individual is the appropriate basic unit of analysis. This tendency is probably exacerbated by the fact that in biology the individual *is* the object of the relevant evolutionary process (or more specifically the phenotypic expression of the information undergoing evolution). The memetic approach involves relinquishing our focus on the individual, and concentrating instead on the meme as the object of a second evolutionary process that makes cognition possible. This perspective can feel unnatural and disorienting but it discloses population-level phenomena that would otherwise go unnoticed because they are not readily detected through introspection.

### **2.1 The Distinction between a Meme and its Phenotypic Implementation**

Durham [22] defines a meme as "any kind, amount, and configuration of information in culture that shows both variation and coherent transmission." Problems with this definition arise because it does not distinguish between cultural information as mental representation and cultural information as implemented behavior or artifact.

The genotype-phenotype distinction is useful here. The cultural analog of a genotype is the mental representation of a meme, and the analog of a phenotype is its *implementation*, or the form it takes if it gets expressed or communicated, typically as action or vocalization. Implementation transforms a meme, incorporating syntactic features characteristic of the channel through which it is conveyed (Brooks [10]). Thus, for example, a dance step looks different with each individual who performs it.

### **2.2 The Interconnectedness of Memes**

Biologists use the term 'allele' to capture the notion of alternative heritable versions of a gene, and Durham [22] accordingly adopted the term 'allomeme' to refer to alternative versions of a meme. This basic concept was tailored to meet the constraints of biology; we all have the same number of genes, and two alleles of each gene (one from each parent). The cultural analog may be too clumsy to capture the subtle relationships between memes. Memes often appear to be stored in a distributed, network-like fashion, connected through webs of association (Hebb [32], Quillian [74], Pribram [73]) there is not necessarily a definitive rationale for saying where one stops and another begins, in semantic space let alone physical space. For example would we consider 'My mother looks good in blue' and 'My mother looks good under a blue umbrella' to be allomemes of the same meme, or different memes?

This kind of difficulty is circumvented by avoiding the notion of alternate versions altogether and using the term 'feature' to refer to a component of a meme. Thus related memes share features. In this paper, 'feature' can refer to a component with any degree of granularity below that of the meme in question; thus the scope of what might be considered a feature could range from an entire array of visual information depicting every perceived quality of a particular umbrella (such as might occur early on in perception) to one bit of information indicating the presence or absence of an umbrella (such as might occur at an advanced stage of cognitive processing). There can be nonlinear (epistatic) relations amongst the features of a meme, as well as between a meme and its implementation.

### **2.3 Meme as Pattern of Information Encoded in the Focus**

In his book *Thought Contagion* [57], Aaron Lynch claims "Much as computer viruses make up a small fraction of computer software, so too do pernicious thought contagions form a small part of our mental software." But is this conservative approach necessary, or even sensible? Where do you draw the line between potentially culturally-transmittable memes and the rest our "mental software"? The haiku-writer, for example, seems able to draw upon any sort of thought, memory, or imaginary construct. Therefore here we explore the opposite approach to that taken by Lynch; we bite the bullet and consider anything that can be the subject of an instant of experience, to be a meme. The category 'meme' now includes not only obviously transmittable ideas like 'Be good or you will go to Hell', but everything from a particular experience of vibrant 'redness', to a realization of a shorter route to work, to a feeling of dread associated with a teacher's posture or facial expression. This may strike some readers as outrageous, but it doesn't really make things as unwieldy as it might seem to at first glance. For the price of this added complication we gain a bridge that connects memetics with phenomena like perception, body language, planning, deductive reasoning, emotion, categorization, analogy... the stuff of the social and cognitive sciences. It may be our only viable direction. A theory of mind that can account for myth and freestyle dance, not to mention rapid personality assessment, is not easily achieved.

Our concept of meme can be clarified by invoking Kanerva's [42] notion of the *focus* - that part of the mind in which sensation (either external or internal e.g. hunger) and stored memory interact to produce a stream of experience. The states of the neurons that comprise the focus determine the content and experiential quality of an individual's awareness. One can think of a meme as a pattern of information that is or has been encoded in an individual's focus. It can be subjectively experienced as a sensation, idea, attitude, emotion, or combination of these, and it can direct implementation by the motor apparatus.

### **2.4 Chunking and Categorization of Memes**

Frequently many memes get integrated into one through a process referred to in the psychological literature as 'chunking' (Miller [65]). Chunking involves forming associations amongst previously-learned memes and establishing this constellation of associations as a new meme in long term memory; it is analogous to the formation of co-adapted genes, or *schemata* (Holland [36]). Whereas chunking generally refers to the binding of semantically unrelated memes (as in the memorization of an arbitrary string of numbers), categorization involves the recognition of semantic relationships. Categorization and the resultant hierarchical structure of

knowledge will not be addressed here in any depth, though it is of relevance to point out that it seems reasonable to expect that the more extensively memes have been chunked or categorized, the greater the complexity of what can be held in the focus at once. Thus what constitutes a meme (and thus a feature) will differ amongst individuals, and within an individual over time.

## 2.5 Core, Enabler, and Hitchhiker Features

A first step toward a 'science of memetics' is to decompose memes into features or feature schemata according to how they relate to fitness. Here we will distinguish the following categories: (1) *core features* that contribute directly to the fitness of a meme, (2) *enabler features*, that enable or facilitate the implementation of core features, and (3) *hitchhiker features*, which exist in the meme due to arbitrary or accidental historical associations to features of the first two kinds. Core features tend to convey semantic information, and enabler features syntactic information, though one can think of situations in which some semantic information serves simply to facilitate expression of other semantic information i.e. functions as an enabler. The first two categories are vaguely analogous to the categorization of genes as structural or regulatory, and the last category is inspired by the phenomenon of genetic hitchhiking (Kojima & Schaeffer [51], Maynard Smith & Haigh [62], Kaplan, Hudson & Langley [43]). The closer together genes are on a chromosome the less likely they will be separated by crossover, so the more tightly linked they are said to be. Hitchhiker alleles confer no fitness advantage, but endure because they are linked to alleles that are important for survival. In both genetic hitchhiking and its cultural analog there is indirect selection for useless (or even detrimental) patterns through their association with beneficial ones. The concept of hitchhiking is closely related to that of exaptation - the evolution of organs or traits not evolved through natural selection for their current use (Gould [29]).

## 3 Selection and the Memetic Fitness Landscape

The next few sections examine in some detail how each of the three phases of evolution - selection, variation, and transmission - map onto the case of culture. Though these phases are discussed one at a time, it is worthwhile to keep in mind that in culture they are less spatiotemporally distinct than in biology. Selection can be coupled to either the generation of variation, or replication, or all three can occur simultaneously (for example when paraphrasing).

### 3.1 Memes 'Rely On' Brains to Select, Vary and Replicate Them

Von Neumann [91] postulated that any self-replicating system consists of two parts: (1) *uninterpreted information* - a self-description that is passively copied to offspring, and (2) *interpreted information* - instructions for how to construct offspring. This turned out to be true of the genetic code; there are genes that provide instructions to the body for how to sustain itself, and genes that provide instructions for how, with the help of someone of the opposite sex, to create a child. But unlike genes, memes do not come prepackaged with instructions for their reproduction. They rely on the pattern-evolving machinery of their hosts' brains to create, select, and replicate them. Since we preferentially spread ideas that satisfy needs, our needs define viable niches for memes to evolve toward. As infants we might cry and kick no matter what need is most pressing, but as children we acquire and continually refine a repertoire of memes that,

when implemented, satisfy various needs. We learn that reaching into the cookie jar satisfies one need, shouting 'help' satisfies another, et cetera. Our memes, and the behavior they elicit, slide into need-defined attractors (regions of stability) in the memetic fitness landscape.

The fact that memes are not independently self-replicating does not prevent them from achieving reproductive success. In fact it may ironically work in their favor, because the cognitive machinery they depend upon not only actively manipulates them to produce 'offspring-memes', but organizes them into a model of the world, or worldview, which it can use to figure out what to do whenever a situation is too complicated for its hardwired instincts. The worldview orchestrates behavior such that a meme gets implemented right when it is likely to be useful, and that increases the probability that other hosts will consider it worthy of replication. This also means that there is a continuous co-evolutionary interplay between pattern and landscape, which contributes to the oft-noted rapidity with which culture evolves.

### 3.2 Brains Select Memes that Satisfy Biological and Cultural Needs

Since many of our needs have a biological basis - e.g. the need for food, shelter, et cetera - meme generation is largely constrained by our heritage as products of biological evolution. Thus the topology of the memetic fitness landscape largely echoes that of the biological fitness landscape. In the short term, the biological fitness landscape, and thus the memetic fitness landscape, fluctuates continuously as one need is satisfied and others take precedence (Hull [41], McFarland & Sibly [64], Gabora & Colgan [24], Maes [59]). For example, after eating, ideas that pertain to finding food are less likely. However over the lifetime of an individual the set of biologically-based needs remains relatively constant. The trajectory of survival-motivated thought can be described as a limit cycle (periodic attractor) that moves through the set of stable memes whose implementations satisfy the various biological needs.

Variation-inducing operations restructure conceptual space and thus affect the memetic fitness landscape. Much as the evolution of rabbits created ecological niches for species that eat them and parasitize them, the invention of cars created cultural niches for gas stations, seat belts, and garage door openers. As one progresses from infancy to maturity, and simple needs give way to increasingly complex needs, the stream of thought acquires the properties of a chaotic or strange attractor, which can be viewed as the formation of crevices in the original limit cycle. The landscape is fractal (i.e. there is statistical similarity under change of scale) in that the satisfaction of one need creates other needs - every crevice when examined closely reveals more crevices. This is analogous to the fractal distributions of species and vegetation patterns described by ecologists (Mandelbrot [60], Palmer [69], Scheuring & Riedi [83]). An endpoint of a cultural evolution trajectory turns out to be not just a point in multidimensional space, but a set of points with their own fitness metric - a 'micro-landscape' in its own right. So although the memetic fitness landscape *loosely* follows the biological fitness landscape, there are places where it deviates, and this effect undoubtedly becomes more pronounced throughout an individual's lifetime. This means that the potential for meme diversity, though constrained by host need, is open-ended.

### **3.3 Personal versus Societal Worldviews**

Our worldviews overlap to the extent that similar experiences and genetic make-ups cause our brains to store the same memes and styles of computation. But they don't overlap perfectly. Each host's train of thought traces out a unique trajectory through conceptual space. It can be useful to think in terms of not only the worldview of an individual, but also the worldview of a group or even human society at large, wherein all frontiers of human endeavor are incorporated.

### **3.4 The Landscape is Sculpted by the Need for Worldview Cohesion**

A need that seems to surface to the forefront (have a large impact on the focus) when other needs are not pressing is the need to connect fragmented representations of the world into a logically-consistent worldview. Since our ability to make predictions and evaluate possible plans of action hangs on the accuracy of this worldview, the survival value of such a tendency is clear.

McCulloch and Pitts [63] showed that networks made of neuron-like components that perform the logical operations AND, OR, and NOT are theoretically capable of computing any Turing machine-computable function. In connectionist-type systems, logical relations are represented implicitly as constraints on the possible states of a system, and computation proceeds through settling into a solution that satisfies many constraints, rather than then explicitly calculating a solution. This is accomplished through modification of the associative strengths amongst the components of the system, and the process is referred to as relaxation or simulated annealing.

### **3.5 Hard-wired Selection**

To the extent that the memetic fitness landscape echoes the shape of the biological fitness landscape, to which we have been adapting since life began, cultural selection is built right into our architecture. Our perceptual and cognitive systems are wired up such that they are primed to focus on and highlight those aspects of external reality that are relevant to our survival (or were in the past). The mental representations we form reflect that bias e.g. (Hubel & Wiesel [39], Marr [61]). Second, the associative organization of memory constrains variation-generating operations. So selection is built right into our hardware.

### **3.6 Malleable Forms of Selection**

In order to create, or even just understand, a new meme, there has to be a conceptual framework from within which it will make sense, and a need, or niche, for it. Therefore, any relevant precursor or "*proto-memes*" must first be assimilated (Wallas [93]). This constraint amounts to a malleable, or plastic, form of selection on new memes.

When one host exhibits a meme that another observing host values, the observing host often displays reinforcing body language and emits words of encouragement. Likewise, if the meme is threatening or inconsistent with valued memes, the observing host's words and behavior tend to be discouraging. Our need for social acceptance makes us more likely to exhibit memes that have been reinforced and less likely to exhibit those that have been discouraged. From a meme's-eye perspective this looks like a subtle strategy by which a meme in one host selectively shapes the probability of its implementation in other hosts.

Selection can also occur after a representation has been internalized but prior to being phenotypically expressed. For example, *mentally simulating* what would happen if an idea were implemented can weed out unworthy ideas (Nersessian [68]). The success of mental simulation varies with the accuracy of one's internalized model of the world, but it provides at least a rudimentary form of selection.

Finally, selection can operate through *biased transmission*; that is we choose to imitate certain individuals and not others (Boyd & Richerson [9]).

## 4 Creativity: the Source of Cultural Variation

### 4.1 Strategy Guides Trajectories through the Memetic Fitness Landscape

The existence of an open niche does not guarantee that the niche will ever be found. In biology the process by which this happens is largely random. Though most mutations and recombinations are detrimental, so many variants are generated that it is not necessary to be clever about *how* they are generated. We could say that biological evolution is a more breadth-first search algorithm than cultural evolution because it relies primarily on massive parallelism rather than strategy.

In culture, on the other hand, variants are generated strategically. We could say that cultural evolution is a more depth-first approach to searching a space of possibilities. The trajectory of a stream of thought is constrained by connections between representations that are similar or spatiotemporally related (Schank [81]), which increases the probability that an advantageous variant is found. For example, when considering the problem of having to get out of your car every day to open the garage door, you would not think about doilies or existentialism, but concepts related to the problem - electricity, human laziness, and various openers you have encountered before. If you were to spend several months on this problem, ideas that pertain to openers of various kinds would for you become a highly active region of conceptual space, analogous to the uncharacteristically high level of activity (and polymorphism) in a small portion of the human genome known as the major histocompatibility complex (MHC) which deals with immune response (Hughes & Nei [40]). During creative thought, memes potentially relevant to a solution would evoke or activate one another, altering or strategically (though not necessarily consciously) manipulating them, a process that is said to involve pattern completion, constraint satisfaction (Rumelhart & McClelland [80]), and the tweaking, blending, redescription, abstraction, and recoding of representations (Hofstadter [34], Holland &c. [37], Karmiloff-Smith [44], [46], Ram [75], Clark & Thornton [14]). Neurophysiological evidence suggests that creating new contexts for representations, that is manipulating them, involves hippocampal binding or linking (Squire [88]), and synchronization (Klimesch [49]), of features encoded by distributed cortical cell assemblies.

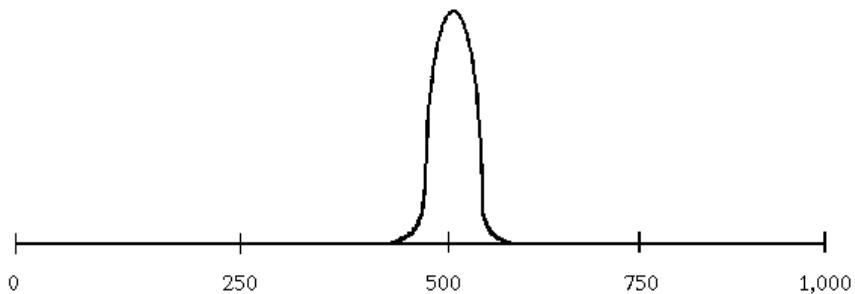
To sum up: fueled by need and constrained by association we carve out trajectories through meme space, and because the fitness landscape that guides this process is fractal, every time that landscape steers the production of a new meme (or even just a slight variant of a preceding meme), the new meme in turn redefines the landscape, and so on, recursively.

## 4.2 Sparse Distributed Memory as a Platform for Generating Variation

SPARSE, DISTRIBUTED MEMORY (Kanerva [42]), or SDM, is a mathematical model of the mechanics underlying the storage and retrieval of memories. It was motivated by the desire to understand how memory provides conscious experience with a thread of continuity via the spontaneous sequential activation of concepts or experiences that are related to one another, sometimes superficially, and other times through resemblances that are highly abstract or metaphorical.

Kanerva draws an analogy between the focus and a combined address-datum register in a computer; they both contain data and serve as a pointer to memory, and can both read from and write to memory. An instant of experience is encoded in the focus by a high-dimensional vector of difference relations, or bits, that represent the presence or absence of some feature, and the mathematics generalizes such that a pattern of bits can represent a value along some dimension. The Hamming distance between two memes is the number of bits that differ. (So the Hamming distance between 11111 and 11100 is two.) Since each meme has an antipode (for example, the antipode of 11111 is 00000), the space of all possible memes can be visualized as a sphere. The address of a meme is the information pattern that specifies where the meme is stored.

If  $L$  is the number of possible features in a meme, the number of possible memes is  $2^L$ . Assuming  $L$  is large the size this space is enormous, so the memory is sparse in that it stores only a small fraction of the set of all possible memes. For example, to construct a SDM with  $L=1,000$ , then out of the  $2^{1,000}$  possible addresses, a workable number of them, say 1,000,000, are chosen at random to be actual storage locations. The number of memes at Hamming distance  $k$  away from any given meme is equal to the binomial coefficient of  $L$  and  $k$ , which is well approximated by a Gaussian or normal curve. If meme  $X$  is 111...1 and its antipode 000...0, and we consider meme  $X$  and its antipode to be the 'poles' of the hypersphere, then approximately 68% of the other memes lie within one standard deviation ( $\sigma$ ) of the 'equator' region between these two extremes (figure 1). As we move through Hamming space away from the equator toward either Meme  $X$  or its antipode, the probability of encountering a meme falls off sharply by the proportion  $e^{-k^2/2\sigma^2}$ . In our example, the median distance from one location to another is 424 bits, and 99.8% of stored memes lie between 451 and 549 bits of any given location.



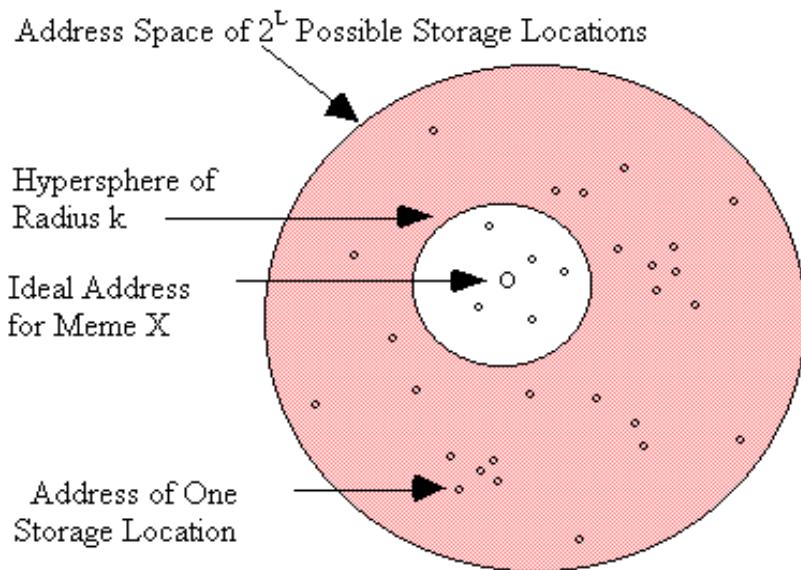
**Figure 1:** Schematic distribution of Hamming distances between addresses of meme  $X$  and addresses of storage locations

A computer reads from memory by simply looking at the address in the address register and retrieving the item at the location specified by that address. The sparseness of the SDM prohibits

this kind of one-to-one correspondence, but it has two tricks up its sleeve for getting around this problem.

First, it feigns content addressability, as follows. The particular pattern of 1s and 0s that constitutes a meme causes some of the synapses leading out from the focus to be excited and others to be inhibited. The locations where memes get stored are memory neurons, and the address of a neuron amounts to the pattern of excitatory and inhibitory synapses from focus to memory that make that neuron fire. Activation of a memory neuron causes the meme to get written into it. Thus there is a systematic relationship between the memes' information content and the locations they activate.

Second, since the probability that the ideal address for storing a meme corresponds to an actual location in memory is vanishingly small, storage of the meme is *distributed* across those locations whose addresses lie within a sphere (or more accurately, hypersphere) of possible addresses surrounding the ideal address (figure 2). The radius (in Hamming metric) of this sphere is determined by the neuron activation threshold. Each location participates in the storage of many memes. In this example we assume that 10,000 memes have been stored in memory. Each meme is stored in 1,000 (of the 1,000,000 possible) locations, so there are approximately 10 memes per location. The storage process works by updating each of the  $L$  counters in each location; to store a 1 the counter is incremented by 1, and to store a 0 it is decremented by 1. These nearly one million operations occur in parallel.



**Figure 2:** Meme  $X$  is stored in all locations within hypersphere of radius  $k$  surrounding its ideal address

If after a meme, say meme  $X$ , is stored, the individual's attention is directed toward external stimuli, then nothing is retrieved from memory. But to the extent that memory contributes to the next instant of awareness, the storage of  $X$  activates retrieval of not only  $X$  itself but all the other memes that have been stored in the same locations. The next meme to be encoded in the focus, , is found by determining the best match; that is, by averaging the contributions of all retrieved

memes feature-by-feature. Whereas the 1,000 retrieved copies of  $X$  (and memes similar to  $X$ ) reinforce one another, the roughly 10,000 other retrieved memes are statistically likely to cancel one another out, so that ends up being similar to  $X$ . Though is a reconstructed blend of many memes it can still be said to have been retrieved from memory. can now be used to address the memory, and this process can be reiterated until it converges on meme  $Y$  that satisfies a current need. This is how the SDM accomplishes depth-first search. The closer  $Y$  is to  $X$ , the faster the convergence. In our example, assuming  $r=425$ , if  $X$  and  $Y$  are more than 200 bits apart  $Y$  is unlikely to be retrieved, but if they are 170 bits apart  $Y$  will be retrieved in about four iterations. Keeler [48] has shown that SDM is a superset of Hopfield-type and connectionist models of autoassociative or heteroassociative memory. The SDM formulation is used here because it lends itself to an understanding of the mechanics of phenomena we are interested in. Since the dynamics emerges from the statistics, rather than from a central executive, it can cope with creative and seemingly unmechanical cognitive phenomena such as wordplay or slips of the tongue. Moreover it is ideally suited to handle the problem of sequential access, which will become relevant when we look at how an infant establishes a train of thought. To model the recollection of a sequence, meme  $X$  is simply used as the address to write  $Y$ ,  $Y$  as the address to write  $Z$ , and so on. Working memory can be viewed as the memes that lie within a given Hamming distance of the meme in the focus such that they are retrievable within a certain number of iterations. Categorization could involve the identification of a feature pattern, and readdressing memes that contain this pattern so that their new addresses put them within working memory reach of one another. Kanerva shows that the architecture of common neural components and circuits in the brain are ideally suited to implement a SDM.

In SDM, associations between memes are not explicitly represented as connection strengths but as proximity in multidimensional space. However in the end they amount to the same thing. The smaller the Hamming distance between two memes, the higher the probability that they will be retrieved simultaneously and blended together in the focus (or one after the other in a chain of related thoughts). What allows the memes to be retrieved simultaneously, however, is that they are either stored in the same neurons or in neurons with nearby addresses, which in turn reflects the neurons' connectivity. Thus factors that affect the storage of a meme will also affect retrieval of that meme; the two processes are intimately connected.

## 5 The Replication and Transmission of Memes

The memetic approach to cognition is not incompatible with approaches that stress the role of innate mechanisms e.g. (Pinker [70]). Rather, as Lumsden & Wilson [56] point out, it builds on this framework, adding that the study of cognition will flounder until we admit that the role of transmission is equally undeniable. Transmission links the memetic processing within an individual to not only memetic processing in other directly-encountered individuals, but processing in individuals *they* encounter, and so on. The ideas and inventions any one individual produces build on the ideas and inventions of others. This phenomenon is known as the *ratchet effect*, and its impact is demonstrated in the following example. If you were suddenly dropped into the Australian desert, you probably would not survive for long. However if you were to run into an aborigine who grew up learning desert survival skills from her family and community that had been passed on and improved upon for generations (such as how to find water in obscure places) you might survive for some time [note 4].

## **5.1 Intra-Individual Meme Replication via Implicit Pointers to Memory**

We saw how, unlike genetic material, memes do not contain instructions for how to make copies of themselves; they replicate when their hosts teach or imitate one another. The memes in a SDM-like memory, however, have a self-replication capacity in the following sense. The pattern of information that constitutes a meme determines which of the synapses leading out from the focus are excited, and which are inhibited - it determines how activation flows through the memory network - which in turn determines the neurons where the meme is stored and from which the next meme is retrieved. Thus embedded in the neural environment that supports their replication, memes act as implicit pointers to memory. These pointers prompt the dynamic reconstruction of the next meme to be subjectively experienced, which is a variation of (statistically similar to) the one that prompted it. It is in that sense that they self-replicate.

## **5.2 Transmission is Lamarckian and Phenotypically Mediated**

Internal replication (with variation) makes cultural transmission Lamarckian - modifications acquired since the acquisition of a meme can be passed on to others (Dawkins [17]). The related point that transmission is phenotypically mediated, as Dennett [19] points out, makes a "science of memetics" less daunting. It means that, unlike biologists, we don't have to fully understand the nature of mental representation to study transmission.

## **5.3 Any Experience can Affect Transmission**

While biological needs affect the focus from the inside, environmental stimuli impact it from the outside. The information-based orientation supports a broader conceptualization of the transmission process than is generally taken. For the purpose of understanding the evolutionary mechanics underlying culture, any interaction between an organism and its environment that impacts the focus is part of this process. It often occurs through imitation of conspecifics (Smith [86], Bonner [7], Robert [77]), or guided instruction (Vygotsky [92], Tomasello &c. [90]), but not necessarily. For example, does it matter whether a child learns to peel a banana by watching her mother, or a monkey, or a cartoon character on TV? What matters is that the child has a mental representation of how to peel a banana. All kinds of interaction with the environment provide us with new representations or alter existing ones, and therefore have the potential to affect the interplay of ideas and emotions that are culturally transmitted.

## **5.4 Transmission Studies in the Social Sciences**

Steps toward an evolutionary perspective on cultural transmission have been taken in the social sciences. This research falls under various names: social diffusion e.g. (Rogers [79]), social epidemics e.g. (Mackintosh & Stewart [58]), coevolution (Durham [22]), social learning e.g. (Tomasello &c. [90]), and social contagion (Rodgers & Rowe [78]). There has also been work done on group creativity e.g. (Czikszentmihalyi [15]).

The evolutionary context could be made much more explicit in this work, capitalizing on what we have learned about evolution in the biological realm by looking for phenomena such as epistasis and drift (random statistical bias due to sampling error) (Wright [96], Cavalli-Sforza &

Feldman [12]) or by viewing reciprocal interactions between like-minded individuals as a cultural analog of biological altruism. It may be through ongoing reciprocal interactions of this kind that a memetically-derived social structure emerges, wherein individuals who regularly generate pleasurable or powerful memes come to be observed carefully and imitated frequently, while other individuals are ignored. These outcasts may be excluded from memetic exchange and come to exhibit a cultural version of the Founder Effect (Holgate [35]) - reduced variation due to drift.

Our knowledge of biological speciation could be applied to study individuation and division of labor in a family or society. In both cases small differences are amplified through positive feedback leading to transformation of the space of viable niches for the evolution of information patterns. This approach could help us address questions such as why siblings are often so different from one another.

## 6 A Scenario for the Origin of Cultural Evolution

We have discussed how memes evolve through selection, variation, replication and transmission. We now address the issue of how cultural evolution got started in the first place.

### 6.1 The Origin of Life and its Cultural Analog

The origin of life poses the following paradox: how could something as complex as a self-replicating molecule arise spontaneously? Traditional attempts to explain this entail the synchronization of a large number of vastly-improbable events. Proponents argue that the improbability of the mechanism they propose does not invalidate it because it only had to happen once; as soon as there was *one* self-replicating molecule, the rest could be copied from this template. However Kauffman [47] proposes an alternative scenario that does not entail the synchronization of numerous improbable events. He suggests that life arose through the self-organization of a set of autocatalytic polymers. When catalytic polymers interact with one another their average length increases. As their length increases, the number of reactions by which polymers can interconvert increases faster than the number of polymers. Therefore a set of interacting molecules under conditions such as are likely to have existed at the time life began would inevitably reach a critical point where there is a catalytic pathway to every polymer present. Jointly they form a self-reproducing metabolism.

We now ask: What is the cultural analog to the origin of life? One could say it is the point in history when organisms acquired the capacity for social transmission, but as many authors e.g. (Darwin [16], Plotkin [71]) have pointed out, although transmission is wide-spread throughout the animal kingdom, no other species has anything remotely approaching the complexity of human culture. Donald [20] argues convincingly that the bottleneck in cultural evolution is the capacity for innovation. Innovation requires more than a kind of awareness that integrates survival needs with environmental affordances, and draws upon memory only to interpret stimuli, or consult a mental map, or recall how some drive was satisfied in the past. It requires an ongoing train of creative representational redescription and counterfactual 'what if...' type thinking. This suggests that the cultural analog to the origin of life was the origin of the first self-perpetuated, potentially-creative stream of thought in an individual's brain.

When an infant has its first experience, there is nothing in memory to draw upon to contribute to that experience; the first meme to occupy its focus does not remind it of anything. Therefore experience is initially driven only by external or internal stimuli, not by memory. Thus the evolution of culture poses a paradox analogous to that of the origin of the self-replicating molecule - how does an infant develop the capacity for a self-sustained train of thought that creatively integrates new experiences with previous ones? Consistent with Kauffman's assertion that the bootstrapping of an evolutionary process is not an inherently improbable event, the "it only had to happen once" argument does not hold water here because the cultural analog to the origin of life takes place in the brain of every infant.

## 6.2 Establishing an Autocatalytic Set of Sparse, Distributed Memories

This section outlines how a SDM-like stream of thought might get established. Let us say that the first meme to occupy an infant's focus and then get stored in memory is a visual experience of its mother in a blue coat. The next is the sound of a dog barking. The Hamming distance between these memes exceeds the maximum for one meme to evoke the memory of another, so the barking does not remind the infant of its mother. Later the infant sees its mother in a red coat. This meme evokes or 'catalyzes' the memory of its mother in a blue coat. To avoid getting stuck in an endless loop wherein 'mother in blue coat' then evokes 'mother in red coat' et cetera, it may form the category 'mother'. However *that* meme does not remind it of anything, so this stream of thought dies off quickly.

As the infant accumulates memes, the statistical probability that a meme in the focus will activate a meme from storage increases, so the streams of remindings get longer. Eventually the memory becomes so densely packed that any meme that comes to occupy the focus is bound to be close enough in Hamming distance to *some* previously-stored meme(s) to activate a variant of itself. This marks a phase transition to a state in which, just as with the origin of life, the sequential activation of self-similar patterns is self-propelled; the memes now form an autocatalytic set. The focus is no longer just a spot for coordinating stimuli with action, but a forum for the variation-producing operations that emerge naturally through the dynamics of iterative retrieval. The resultant memes evolve along different trajectories toward different basins of attraction, 'specializing' in the fulfillment of one need or another. Those that satisfy the same need compete until one becomes habitual, while those that fulfill different needs are able to coexist within the same host. As with biological speciation, small differences are amplified through positive feedback leading to transformation of the space of viable niches for the evolution of information patterns.

Note that in this example the 'mother' meme is the infant's first category. A simple way of describing this situation is: if the 'mother in blue coat' meme is represented as 111, and the 'mother in red coat' meme is represented as 110, the 'mother' meme can be represented as 11\*, where \* means either 1 or 0. It is also the infant's first *derived* meme. That is, it is the first information pattern to enter the focus not purely by way of external stimuli but through the necessity of a logical operation on previously-stored memes - in this case an OR gate - which could be realized in the brain via adjustment of connection strengths. The act of categorization projects the original information space, which had n relevant dimensions, onto a new space that has n-1 dimensions (for example, here coat color is no longer relevant). It effectively makes the

space denser, and increases susceptibility to the autocatalytic state. On the other hand, creating new memes by combining previously-stored memes could interfere with the establishment of a sustained stream of thought by increasing the dimensionality of the space, thereby decreasing density. If indeed cross-category blending disrupts conceptual autocatalysis, one might expect it to be less evident in young children than in older children, and this expectation is born out experimentally (Karmiloff-Smith [45]).

Note also that the density of memes necessary to reach and maintain this autocatalytic state will depend on the neuron activation threshold. If the threshold is too high (the hypersphere of potentially activated memes is too small) even very similar memes cannot evoke one another, so a stream of remindings, if it happens at all, dies off readily. If the threshold is too low (the hypersphere too large), then any meme will evoke a multitude of others not necessarily meaningfully related to it. Successive patterns in the focus will have little or no resemblance to one another; the system may be catalytic but it is not autocatalytic. The free-association of the schizophrenic (see Weisberg [95]) seems to correspond to what one might expect of a system like this. For memory to produce a steady stream of meaningfully-related yet potentially creative remindings, the threshold must fall within a narrow intermediate range. This is consistent with Langton's [53] finding that the information-carrying capacity of a system is maximized when its interconnectedness falls within a narrow regime between order and chaos. The situation may turn out to be slightly more complicated; sustaining a creative train of thought may involve not only keeping the activation threshold within a narrowly-prescribed range but dynamically tuning it in response to the situation at hand. This is particularly likely if the memory is not uniformly dense (i.e. clusters of highly-correlated memes) or if different kinds or stages of thought require different degrees of conceptual blending. For example, finding unusual associations, such as puns that employ epistatic relationships between semantic and syntactic components of a meme, may depend on the preconscious ability to voluntarily increase hypersphere radius. Refinement of the pun may then require a decrease in radius.

Thus we have a plausible scenario for how cultural evolution, like biological evolution, could have originated in a phase transition to a self-organized web of catalytic relations between patterns.

## 7 Why is Culture Unique to Humans? - A Speculative Answer

Recall that in order for a network of memes to reach an autocatalytic state, the activation threshold must be calibrated to fall within a narrow range to achieve a delicate balance between the capacity for semantic continuity on the one hand and creative association on the other. The penalty for having too low a threshold would be very high; successive thoughts would not necessarily be meaningfully related to one another, and thinking would be so muddy that survival tasks are not accomplished. Too high a threshold, on the other hand, would not be life-threatening. The focus would virtually always be impacted with external stimuli or internal drives such as hunger; memory would be pretty much reserved for recalling how some goal was accomplished in the past. A stream of experience that involved the iterative reorganization of stored memes would likely die out before it produced something creative. This may be the situation present in most brains on this planet, and the reason that apes are limited to episodic memory e.g. (Donald [20]).

The advantages of a sustained train of thought would be largely lost on animals because they have neither the vocal apparatus nor the manual dexterity and freedom of upper limbs to *implement* complex ideas. No matter how brilliant their thoughts were it would be difficult to do something useful with them. Moreover, in an evolutionary line there is individual variation, so the lower the average activation threshold, the higher the fraction of individuals for which it is so low that they do not survive. It seems reasonable to suggest that apes are not *a priori* prohibited from evolving complex cognition, but that there is insufficient evolutionary pressure to tinker with the activation threshold until it achieves the requisite delicate balance to sustain a stream of thought, or to establish and refine the necessary feedback mechanisms to dynamically tune it to match to the degree of conceptual fluidity needed at any given instant. It may be that humans are the only species for which the benefits of this tinkering process have outweighed the risks.

## 8 Computational Approaches to Cultural Evolution

Meme and Variations, or MAV (Gabora [25]) is a computer model of a society composed of interacting neural network-based agents. Unlike other such models that combine biological and cultural evolution e.g. (Ackley [2], Spector & Luke [87]) these agents don't have genomes, and neither die nor have offspring, but they can invent, implement, and imitate memes. MAV successfully evolves patterns of information through cultural implementations of variation, selection, and replication, and exhibits phenomena observed in biological evolution such as: (1) drift (2) epistasis increases time to reach equilibrium, (3) increasing frequency of variation-generating operations increases diversity, and (4) although in the absence of variation-generating operations meme evolution does not occur, increasing variation-generation much beyond the minimum necessary for evolution causes average fitness to *decrease*. MAV also addresses issues specific to cultural evolution, such as the effects of mental simulation, imitation, and strategy. Perhaps the most interesting finding it yielded was that although for the society as a whole the optimal creation-to-imitation ratio was approximately 2:1, for the agent with the fittest memes, the less it imitated (i.e. the more computational effort reserved for creation) the better.

MAV will hopefully serve as a stepping stone to more advanced models of memetic evolution. Of particular interest will be models that: (1) like Tierra, a model of biological evolution (Ray [76]), harness the power of evolution to explore and transform an *open-ended* space of possible patterns, but (2) explore the space strategically on the basis of accumulated knowledge rather than at random [note 5], (3) have fitness landscapes that emerge through the needs of the agents within the constraints of their environment as in (Maes [59]), and (4) have agents that must learn for themselves which memes, when implemented, best satisfy each of their various needs.

Mathematical models of culture are too minimal to cope with the open-ended diversity of culturally-derived information (variation is generally restricted to trial and error learning or transmission error) l&c. one address the numerous intra-individual factors that undoubtedly have emergent inter-individual consequences, such as how representations are grounded in experience and how they are stored, retrieved, and implemented. Models of individual intelligence and creativity, on the other hand, lack transmission and replication. Although this research may not explicitly attempt to address group processes it typically focuses not on the sorts of simple inferences and creative acts that a person raised alone in the wild would be capable of, but on

complex acts such as story comprehension, that might be unlikely to develop in isolation. With the advent of massively parallel computers it is becoming increasingly feasible to place computational models of individual creativity and problem-solving in a cultural context. This approach could provide insight into not only problems pertaining to representation and culture, but evolution in general, through comparison with biology. For example, the question of why there is so much redundancy in the genetic code has generated much discussion which may also apply to the question of why there are redundant mental maps in the brain; both may reflect constraints on the nature of an information-evolving code.

## **9 A Memetic Perspective on Induction, Censors, and the Unconscious**

We turn now to how the cultural evolution perspective can shed light on some aspects of how people think and interact.

### **9.1 Mental Censorship and Induction**

Initially an infant is unselective about meme acquisition, since (1) it doesn't know much about the world yet, so it has no basis for choosing, and (2) its parents have lived long enough to reproduce, so they must be doing something right. However just as importing foreign plants can bring ecological disaster, acquisition of a foreign meme can disrupt the established network of relationships amongst existing memes. Therefore the infant develops mental censors that ward off internalization of potentially disruptive memes. Censors might also be erected when a meme is found to be embarrassing or disturbing or threatening to the self-image (Minsky [66]). In a SDM-type architecture this could be accomplished by increasing the activation threshold so as to prematurely terminate the relaxation process and prevent the content of the focus from assimilating with stored memes. Much as erecting a fence increases the probability that people will stay on either one side or the other, this would warp the statistical probabilities involved in this meme's partaking in the ongoing process of associative recall, such that the individual either avoids the censored meme or dwells on it excessively. This seems to be consistent with our bipolar attitude toward highly censored subjects such as aggression and sexuality.

On the other hand, when the cost of the disruption is outweighed by the potential benefit accrued by a world model that can accommodate the new meme, the threshold would be lowered. Most thoughts seem to have little effect on our understanding of the world at large, but once in a while we experience a meme that significantly modifies our world view. The situation is reminiscent of superconductivity; lowered resistance increases correlation distance, and thus a perturbation to any one pattern can percolate through the system and affect even distantly-related patterns. It would be interesting to determine experimentally whether the "inductiveness" of our memes, like other self-organizing systems, exhibits the ubiquitous inverse power law (Bak, Tang & Weisenfeld [3]). Just as in a sand pile perched at the proverbial Tedge of chaos' once in a while a collision between two grains will lead to another in just the right chain reaction to generate a large avalanche, occasionally one thought will trigger a chain reaction of others in a way that reconfigures the conceptual network.

## 9.2 The Unconscious

The concept of the unconscious has been influential and useful despite the obvious incongruity: how is it that we can consciously discuss something that is unconscious? What we may be referring to is the fleeting experience of memes that are dynamically reconstructed as in a SDM but which do not readily assimilate with other memes and so get discarded from the focus. In other words, the need for worldview consistency prohibits further computational resources from being spent on trying to integrate what appears to be a nonsensical construction into the memory. Of course there is no reason why a meme that is not immediately integrated into the memory might nevertheless affect the memory; the very process of determining whether it can be assimilated or not might itself have effects that infiltrate the system. This possibility is supported by the finding that subjects' behavior can be affected by priming material of which they have no recall e.g. (Dunbar & Schunn [21], Fehrer & Raab [23]). Subconscious processing of this sort could, in fact, resculpt the memetic fitness landscape in such a way that a previously-discarded meme is more readily assimilated the next time it is encountered.

## 9.3 Cultural Momentum

Despite being derived, directly or indirectly, from human need, memes do not always promote our survival (Greene [30], Alexander [1]). As Dawkins [18] points out, "It is true that the relative survival success of a meme will depend critically on the social and biological climate in which it finds itself, and this climate will certainly be influenced by the genetic make-up of the population. But it will also depend on the memes that are already numerous in the meme-pool." Much like runaway selection in biology, once a meme can replicate with variation on the basis of some selection criterion, it can evolve out of the orbit of the need that originated it. We can't *help* but engage in a stream of thought, spontaneously generating new memes like "if only such and such had been different...". Any more than biological evolution can help but generate new species. This *cultural momentum* could explain why, despite the intuition that individuals control their streams of thought, creators often express surprise at the sudden appearance of an idea, and deny active effort in its immediate creation (Bowers &c. [8], Guilford [31], Kubose &c. [52], Wallas [93]). We seem to control the birth of 'our' ideas only to the extent that we provide a fertile ground for them to be fruitful and multiply - by internalizing relevant background knowledge, identifying new needs, and exposing ourselves to stimuli that help trigger ideas that fulfill those needs. (So if you don't like this idea, don't blame me.)

Spurious basins of attraction sometimes arise in recurrent neural networks through the compositional interaction of explicitly-trained attractors (Hopfield [38]). Cultural momentum may boil down to a phenomenon of this sort. Just because the memetic fitness landscape largely echoes the biological fitness landscape, that doesn't mean that behavior elicited by memes in spurious basins of attraction arising through representational redescription need always be conducive to survival. Nevertheless a stream of thought could be censored before it elicits harmful behavior. Streams of thought probably get blocked or sidetracked on a regular basis, not just by censors, but by minute-to-minute undulations in the hyperdimensional fitness landscape, that is, change in the relative urgency of the multitude of survival-related or derived needs impacting the focus.

The concept of cultural momentum sheds light on the issue of free will. Those who argue for the existence of a central executive in memory may come to be viewed as the creationists of philosophy and cognitive science. Human will can instead be viewed as the emergent orchestration of needs, stimuli, and retrieved memories impacting the focus, which is subject to cultural momentum and therefore, in a sense, beyond our control.

#### 9.4 The Birth of Creative Ideas

The biologically-inspired model developed here supports a variant of the combination theory of creativity - that new ideas arise through combinations and transformations of old ones (Boden [6], Koestler [50]). The aspect of this theory that does not ring true is that it neglects the role of emotions. Here we consider emotions, as well as ideas, to be encoded as information in memes; some components of a meme are simply interpreted by parts of the mind that experience them as emotion, whereas others are interpreted by parts of the mind that experience them as ideas. Much research on analogy deals with how the structure or 'conceptual skeleton' underlying one idea gets abstracted and applied to another (Gentner [26], Gick & Holyoak [27]). We can expand on this general idea by suggesting that many forms of creative expression begin with the (unvoiced) question: What would the pattern of information that encoded the emotion I experienced during *this* particular event look like if expressed through the constraints of *that* medium? The existence of inherent limitations on how a pattern could be translated from one domain to another is consistent with the frequent observation that creativity involves both freedom and aesthetic constraint. Thus all creativity is directly or indirectly derived from experience in the world, and since the mathematics underlying this world, the set of all natural functions, is a small subset of all *possible* functions, the constraints that guide creation are not arbitrary but objective and familiar; for example the drum beat of a song might echo a heartbeat, when the rhythm and chord progression are reminiscent of the sound of someone sobbing we feel sad, and we hear the wrong note even if we have never heard the song before.

It makes sense to expect that a meme or meme complex that has been censored (for example because it evoked unbearable sadness or anger) would be vulnerable to being targeted as an area where worldview cohesion could be increased. Since at the time the censored material was experienced it was prohibited from forming associations to obviously-related memes, it in turn cannot be *retrieved* through these expected or straightforward associations. It can only be retrieved via 'backdoor entrances', that is associations that reflect structural congruity at an abstract level. Thus a musician may come to habitually funnel patterns encountered in a variety of domains - and particularly censored material - through modules that filter out hitchhiker and enabler features, and adapt the core features (or feature schemata) to the constraints of music. It is in this repackaged format that memes originally deemed dangerous can be integrated into the memory at large without harm, and it is through this process that the creator establishes a sense of control over memes that were previously 'off-limits'.

#### 9.5 Conceptual Linkage Disequilibrium and Hitchhiking

Arguments against a theory of cultural evolution generally consist of a series of statements as to how the cultural situation differs from that of biology e.g. (Gould [28], Thagard [89]). These arguments, however, do not constitute a viable reason to discard the idea that culture is an

evolutionary process. Imagine that 100 years before Darwin proposed the theory of biological evolution through natural selection, another scientist had discovered another system whereby patterns of information evolved, say in a test tube. Given this scenario would it have made sense for Darwin to dismiss the importance of biological evolution simply because it proceeds through different mechanisms from the originally-discovered test tube form of evolution? This would obviously have been foolish. It would have robbed humanity of not only the unifying power of a theory of biological evolution, but the opportunity to use knowledge of how evolution works under one set of constraints and affordances as a scaffold to direct the study of how it works under a different set of constraints and affordances. But time and again it is argued that a theory of cultural evolution is doomed simply because it would have to work through different mechanisms from those of biological evolution.

Ironically this situation in itself provides us with a nice example of how knowledge of evolution acquired in the realm of biology can help unravel analogous situations in the realm of culture. The biasing effect of historical association is an important theme in population genetics. Alleles of linked genes, such as those that code for red hair and freckles, continue to co-occur more often than chance even after individuals in the lineage from which these alleles originated begin mating randomly with individuals from other lineages that did not have these alleles. One can theoretically measure the number of generations necessary for these genes to achieve a state of random association or *linkage equilibrium*, and this process can be modeled computationally. Similarly, people often have difficulty applying an idea or problem-solving technique to situations other than the one in which it was originally encountered, and conversely, exposure to one problem-solving technique interferes with the ability to solve a problem using another technique e.g. (Luchins [55]). Psychologists refer to this as *mental set* (though it is more commonly known as 'throwing the baby out with the bathwater'). We could view mental set as a state of conceptual linkage *disequilibrium*. In the present example, conceptual linkage disequilibrium hinders our ability to abstract the basic concept of evolution from its biological manifestation so that it can be applied with ease to the case of culture. One could argue that it would make sense for cultural evolution to be the default form of evolution in disciplines outside of biology, much as in tropical climates the default form of skiing is water-skiing rather than snow-skiing.

Conceptual linkage equilibrium is achieved when all instances of hitchhiking have been obliterated. In an influential paper on the relationship between DNA polymorphism and recombination rates, Begun and Aquadro [5] suggested that genetic hitchhiking may have significant evolutionary impact:

"This correlation suggests that levels of neutral variation in many of the gene regions for which variation has been measured have been reduced by one or more hitchhiking events. Provided that a new selectively favored mutation goes to fixation before another advantageous mutation arises close to it, each fixation will be surrounded by a 'window' of reduced polymorphism, the relative size of which is proportional to the rate of recombination for that region of the genome." The general idea here translates nicely to cognition: if a meme goes to fixation in a society due to selective advantage conferred by one or more core features, its enabler and hitchhiker features will also exhibit reduced polymorphism, and the size of the 'window' will vary with the extent to which hitchhiker features are conceptually bound to that meme. (For example, the basic concept

of a typewriter/computer keyboard was historically enabled by the QWERTY keyboard design, which has now gone to fixation despite poor performance relative to other possible keyboard designs.) One could argue that recreation is the re-creation of information patterns in different domains from the ones in which they were originally encountered, thereby filtering out conceptual prejudices that reflect nothing more than mechanical constraints or historical legacies of the original domain. Play, intellectual pursuits, and other creative endeavors are then algorithms for achieving a state of conceptual linkage equilibrium through mental operations that, like genetic recombination, increase polymorphism by reducing fixation through hitchhiking.

The account of creativity proposed here may seem too simple to explain the seemingly limitless human potential for creativity, but it may seem less far-fetched when we consider the variety of species produced by biological evolution, which operates without the benefit of strategy. Furthermore, raw materials for the creative process may be acquired in exceedingly subtle ways. It is conceivable that you might watch a stream flow and without your consciously thinking, "It flows... things can flow... I could even, in some sense, adopt a more flowing approach to life", the experience might be reconfiguring your memetic infrastructure in a way that makes you more easygoing. I am not making any claims about the extent to which experiences of this sort affect us or even whether they occur at all, but rather suggesting that we not prematurely place a lid on the kinds of processes that could affect a network of representations and thereby affect creation and transmission.

## **10 Memetics as the Missing Link between Science, Spiritual Notions, and Feminism**

Recall that to implement a meme is to express it, so that it crystallizes from the world of ideas into words or body language or objects in the physical world. Memes fool potential hosts into believing they want or need or identify with them by attaching themselves to supporting memes that we already identify with, or that represent things we need or want (as advertisers are well aware). Thus the greater the extent to which we identify with or value ourselves in terms of the memes (including those that pertain to the self) and implemented artifacts we possess or lack, the more vulnerable we are to ever-more-seductive forms of persuasion and advertising which tie up time, energy, and resources that could be applied toward other goals.

One way to defend oneself against painful or manipulative memes is to construct what Dennett [19] refers to as a 'meme-immunological system'; that is, formulate new memes specifically to deflect 'memetic antigens'. However constructing 'memetic antibodies' of this sort is time-consuming, and like any immunological response it has to be repeated every time the outside agent evolves a counter-response. Perhaps this explains the purported benefits of 'transcending the ego' e.g. (Walsh & Vaughan [94]), which can be taken to mean getting in touch with who we were before our minds were colonized by memes, through practices such as meditation. These practices may also give the brain time to anneal material that was never fully assimilated because of distraction or censorship - mend flaws in the fabric of the individual's worldview - so that the censor-ridden personal worldview comes to more closely approximate unbiased conceptual space. Release from the restrictive power of censors may produce a feeling of unity or one-ness.

One of the things I hope the study of memetics will one day shed light on is the cumulative effect on women's' self esteem of encountering the default "he" or "his", or phrases like "man and his world". Each time it happens, it reinforces, however subtly, the thought 'I'm not one of the important ones', or 'that doesn't pertain to me so I don't really count'. The effect of this may be insignificant the first few, or even hundred, times it happens. But by the time a young girl reaches adulthood, experiences of this kind probably number in the millions, and that may be when the effects start to kick in.

## 11 Can Cultural Evolution Provide a Synthetic Framework for the Cognitive and Social Sciences?

This paper presents a theoretical framework for cultural evolution, adopting an approach analogous to that of the population geneticist in that it emphasizes meme evolution through social interaction and de-emphasize individual 'hosts'. Although the cultural evolution of memes operates through very different mechanisms from those of biology, culture is the only system comparable to biology, because it is the only other system to exhibit the imperative features of evolution - adaptive exploration and transformation of an information space through variation, selection, and transmission. All pattern in the information we encounter can be traced to either (1) the physical constraints and self-organizing properties of matter, (2) biological evolution, (3) cultural evolution, or (4) interactions between these causal principles.

One important difference between the two forms of evolution is that culture is less random - new patterns have a greater-than-chance probability of being more fit than their predecessors. The reason for this is interesting. Since memes (unlike genomes) do not come packaged with instructions for their replication, they must rely on the pattern-evolving machinery of our brains to do it for them. Ironically, this state of dependence enhances their proliferative potential, because the machinery they depend upon constructs and continually updates mental models of its world - that is, weaves the memes into a worldview - and uses this worldview to enhance the assimilation and implementation of the memes and their offspring.

Cultural evolution presents a puzzle analogous to the origin of life: the origin of a self-sustained stream of potentially-creative thought in an infant's brain. The idea that life originated with the self-organization of a set of autocatalytic polymers suggests a possible mechanism for how this comes about. Once a threshold density of assimilated memes is surpassed, any meme that occupies the focus is close enough in Hamming distance to evoke or 'catalyze' the spontaneous retrieval or creative reconstruction of a statistically similar meme, thus the memes form an autocatalytic set. Note that this macroscopic account suggests an explanation for only that aspect of human consciousness that differentiates us from other 'experiencers'; it does not address the mystery of 'raw awareness' that some say characterizes not only our experience but that of a cow or a mosquito or even a thermostat e.g. (Chalmers [13]). Whether or not this specific theory turns out to be correct, it illustrates how the analogy to biology can focus our study of culture by providing a scaffold around which explanatory theories can be built.

The cultural framework for cognition developed here suggests tentative explanations for psychological phenomena such as censorship and creativity, and why they are virtually unique to humans. It also suggests explanations for various social phenomena. For example the telling of a

story, rumor, or joke can be viewed as an act of memetic altruism between like-minded individuals analogous to the biological altruism that occurs between genetically-similar individuals. Ongoing reciprocal interactions of this kind may result in the emergence of a memetically-derived social structure.

If we are to take seriously the idea that culture is an evolutionary process, we can look to evolution to provide the kind of overarching framework for the humanities that it provides for the biological sciences. This approach may put us on the road to understanding the pervasiveness, diversity, and adaptive complexity of the cultural debris that surrounds and infests us.

## Acknowledgments

I would like to thank David Chalmers, Anthony Francis, Peter Greene, Pentti Kanerva, Ashwin Ram, Bruce Sawhill, Roger Schank, and Patrick Tufts for thoughtful comments on the manuscript.

## Notes

1. Note that by 'causal principle' I mean something that generates useful descriptions, rather than a 'law'.
2. Though viruses are unique in the biological world in that they rely on hosts to replicate, we will consider viral evolution an anomalous offshoot of biological evolution, because: (1) the evolving patterns of information are encoded as sequences of nucleotides, (2) variation is through mutation and recombination, and (3) transmission and selection are mediated through genotype.
3. Or one could argue that the 'selection' of matter over antimatter, and its subsequent amplification and variation, constitutes yet another form of evolution.
4. This example is a variation of one transmitted by R. Boyd pers. com.
5. MAV has this property to some extent; a more sophisticated example is Copycat, a model of analogy-building, (Mitchell [67]).

## References

- [1] Alexander, R.D. (1980) Darwinism and human affairs, Pitman.
- [2] Ackley, D. (1994) A Case for Distributed Lamarkian Evolution. In Artificial Life III, Addison-Wesley.
- [3] Bak, P., Tang, C., & Weisenfeld, K. (1988) Self-organized criticality. Phys. Rev. A 38: 364.
- [4] Bateson, G. (1972) Steps to an ecology of mind, Ballantine Books.

- [5] Begun, D.J. & Aquadro, C.F. (1992) Levels of naturally occurring DNA polymorphism correlate with recombination rates in *D. melanogaster*. *Nature*, 356, p. 519-520.
- [6] Boden, M. A. (1991) *The creative mind: Myths and mechanisms*, Cambridge University Press.
- [7] Bonner, J. T. (1980) *The evolution of culture in animals*, Princeton University Press.
- [8] Bowers, K.S., Regehr, G., Balthazard, C. & Parker, K. (1990) Intuition in the context of discovery. *Cognitive Psychology*, 22:72-110.
- [9] Boyd, R. & Richerson, P. J. (1985) *Culture and the evolutionary process*, The University of Chicago Press.
- [10] Brooks, V. (1986) *The neural basis of motor control*, Oxford University Press.
- [11] Campbell, D. T. (1987) Evolutionary epistemology. In *Evolutionary epistemology, rationality, and the sociology of knowledge*, eds. G. Radnitzky and W. W. Bartley, 47-89, Open Court.
- [12] Cavalli-Sforza, L. L. & Feldman, M. W. (1981) *Cultural transmission and evolution: A quantitative approach*, Princeton University Press.
- [13] Chalmers, D.J. (1996) *The conscious mind: In search of a fundamental theory*, Oxford University Press.
- [14] Clark, A. & Thornton, C. Trading spaces: Computation, representation, and the limits of uninformed learning. Forthcoming in *Brain and Behavioral Sciences*.
- [15] Czikszentmihalyi, M. (1990) *Flow: The Psychology of Optimal Experience*. Harper & Row.
- [16] Darwin, C. (1871) *The descent of man*, John Murray Publications.
- [17] Dawkins, R. (1976) *The selfish gene*, Oxford University Press.
- [18] Dawkins, R. (1982) *The extended phenotype*, W. H. Freeman and Company.
- [19] Dennett, D. C. (1995) *Darwin's dangerous idea*, Little, Brown and Company.
- [20] Donald, M. (1991) *Origins of the modern mind*, Harvard University Press.
- [21] Dunbar, K. & Schunn, C.D. (1990) The temporal nature of scientific discovery: The roles of priming and analogy. In: *Proceedings of the Twelfth Annual Meeting of the Cognitive Science Society*, 93-100.

- [22] Durham, W.H. (1991) Coevolution: genes, culture, and human diversity. Stanford University Press.
- [23] Fehrer, E. & Raab D. (1962) Reaction time to stimuli masked by metacontrast. *Journal of Experimental Psychology*, 63, 143-147.
- [24] Gabora, L. M. & Colgan, P. W. (1990) A model of the mechanisms underlying exploratory behaviour. In *The simulation of adaptive behavior*, eds. S. Wilson and J. A. Mayer, MIT Press.
- [25] Gabora, L.M. (1995) Meme and variations: A computational model of cultural evolution. In 1993 Lectures in Complex Systems, Addison-Wessley.
- [26] Gentner, D. (1983) Structure-mapping: A theoretical framework for analogy. *Cognitive Science*, 7(2).
- [27] Gick, M.L. & Holyoak, K.J. (1983) Schema induction and analogical transfer. *Cognitive Psychology*, 15, 1-38.
- [28] Gould, S. J. (1991) *Bully for brontosaurus: reflections in natural history*. W. W. Norton & Company. p. 63-66.
- [29] Gould, S.J. & Vrba, E.S. (1982) Exaptation - a missing term in the science of form. *Paleobiology*, 8(1), 4-15.
- [30] Greene, P.J. (1978) From genes to memes? *Contemporary Sociology* 7, 6-709.
- [31] Guilford, J.P. (1979) Some incubated thoughts on incubation. *Journal of Creative Behavior*, 13, 1-8.
- [32] Hebb, D. (1949) *The organization of behavior*, Wiley and Sons.
- [33] Hofbauer, J. & Sigmund, K. (1988) *The theory of evolution and dynamical systems*, Cambridge University Press.
- [34] Hofstadter, D. R. (1985) Variations on a theme as the crux of creativity. In *Metamagical Themes*, 232-259, Basic Books.
- [35] Holgate, P. (1966) A mathematical study of the founder principle of evolutionary genetic. *J. Appl. Probl.*, 3, 115-128.
- [36] Holland, J. K. (1975) *Adaptation in natural and artificial systems*, University of Michigan Press.
- [37] Holland, J. H., Holyoak, K. J., Nisbett, R. E. & Thagard, P. R. (1986) *Induction*, MIT Press.

- [38] Hopfield, J.J. (1982) Neural networks and physical systems with emergent collective computational abilities. *Proceedings of the National Academy of Sciences (Biophysics)*. 79(8), 2554-2558.
- [39] Hubel, D.H. & Wiesel, T.N. (1979) Brain mechanisms and vision. *Scientific American*, 241(3), 150-62.
- [40] Hughes, D.M. & Nei, M. (1988) Patterns of nucleotide substitution at major histocompatibility complex class I loci reveals overdominant selection, *Nature*, 335, 167-170.
- [41] Hull, C. L. (1943) Principles of behavior, Appleton-Century-Crofts.
- [42] Kanerva, P. (1988) Sparse distributed memory, MIT Press.
- [43] Kaplan, N.L., Hudson, R.R., & Langley, C.H. (1989) The Rhitchhiking effect revisited. *Genetics*, 123, 887-899.
- [44] Karmiloff-Smith, A. (1986) From meta-processes to conscious access: Evidence from children's metalinguistic and repair data. *Cognition* 23, 95-147.
- [45] Karmiloff-Smith, A. (1990) Constraints on representational change: Evidence from children's drawing. *Cognition* 34, 57-83.
- [46] Karmiloff-Smith, A. (1992) Beyond modularity: A developmental perspective on cognitive science, MIT Press.
- [47] Kauffman, S. (1993) Origins of order, Oxford University Press.
- [48] Keeler, J.D. (1988) Comparison between Kanerva's SDM and Hopfield-type neural networks. *Cognitive Science* 12, 299-329.
- [49] Klimesch, W. (1995) Memory processes described as brain oscillations. *Psycholoquy*, 6.06.
- [50] Koestler, A. (1964) The act of creation, Picador.
- [51] Kojima, K. & Schaeffer, H.E. (1967) Survival processes of linked mutant genes. *Evolution*, 21, 518-531.
- [52] Kubose, S.K. & Umenoto, T. (1980) Creativity and the Zen koan. *Psychologia*, 23(1), 1-9.
- [53] Langton, C.G. (1992) Life at the edge of chaos. In *Artificial life II*. eds. C.G. Langton, C. Taylor, J.D. Farmer & S. Rasmussen, Addison-Wesley.
- [54] Lenat, D.B. (1974) The ubiquity of discovery. *Artificial Intelligence Journal*, 9, 257-86.

- [55] Luchins, A.S. (1942) Mechanization in problem solving. *Psychological Monographs*, 54, No. 248.
- [56] Lumsden, C. & Wilson, E. O. (1981) Genes, mind, and culture, Harvard University Press.
- [57] Lynch, A. (1996) Thought Contagion: How Belief Spreads Through Society, Basic Books.
- [58] Mackintosh, D.R. & Stewart, G.T. (1979) A mathematical model of a heroin epidemic: Implications for control policies. *Journal of Epidemiology and Community Health*, 33, 299-301.
- [59] Maes, P., ed. (1991) Designing autonomous agents: Theory and practise from biology to engineering and back, MIT Press.
- [60] Mandelbrot, B.B. (1982) The fractal nature of geometry. W.H. Freedman and Company.
- [61] Marr, D. (1982) Vision, Freeman.
- [62] Maynard Smith, J. & Haigh, J. (1974) The hitchhiking effect of a favourable gene. *Genetic Research*, 23, 23-35.
- [63] McCulloch, W.S. & Pitts, W. (1943) A logical calculus of the ideas immanent in nervous activity. *Bulletin of Mathematical Biophysics*, 5, 115-133.
- [64] McFarland, D.J. & Sibly, R.M. (1975) The behavioural final common path. *Philosophical transactions of the London Royal Society* 270B, 265-93.
- [65] Miller, G.A. (1956) The magic number seven, plus or minus two: Some limits on our capacity for processing information. *Psychological Review*, 63, 81-97.
- [66] Minsky, M. (1985) The Society of Mind. Simon and Schuster.
- [67] Mitchell, M. (1993) Analogy-making as perception, MIT Press.
- [68] Nersessian, N. (1993) In the theoretician's laboratory: Thought experimenting as mental modeling. In PSA 1992, vol. 2, eds. D. Hull, M. Forbes, & K Okrulik. PSA.
- [69] Palmer, M.W. (1992) The coexistence of species in fractal landscapes. *American Naturalist*, 139, 375-397.
- [70] Pinker, S. (1995) The language instinct. HarperPerrenial.
- [71] Plotkin, H.C. (1988) The role of behavior in evolution. MIT Press.
- [72] Popper, K. R. (1963) Conjectures and refutations, Routledge & Kagen.

- [73] Pribram, K.H. (1974) How is it that sensing so much we can do so little? In The neurosciences third study program, eds. F.O. Schmidt & F.G. Worden, MIT Press.
- [74] Quillian, M.R. (1968) Semantic memory. In Semantic information processing, ed. M. Minsky, MIT Press.
- [75] Ram, A. (1993) Creative conceptual change. Proceedings of the Fifteenth Annual Conference of the Cognitive Science Society, 17-26.
- [76] Ray, T. (1991) An approach to the synthesis of life. In Artificial life II. eds. C.G. Langton, C. Taylor, J.D. Farmer & S. Rasmussen, Addison-Wesley.
- [77] Robert, M. (1990) Observational learning in fish, birds, and mammals: A classified bibliography spanning over 100 years of research. Psych Record 40, 289-311.
- [78] Rodgers, J.L. & Rowe, D.C. (1993) Social contagion and adolescent sexual behavior: A developmental EMOSA model. Psychological Review, 100(93), 479-510.
- [79] Rogers, E.M. (1962) Diffusion of innovations, Macmillan Publishing Company.
- [80] Rumelhart, D.E. & McClelland, J.L. eds. (1986) Parallel distributed processing, Bradford/MIT Press.
- [81] Schank, R.C. (1983) Dynamic memory, Cambridge University Press.
- [82] Schank, R.C. & Leake, D.B. (1989) Creativity and learning in a case-based explainer. Artificial Intelligence, 40(1-3): 353-385.
- [83] Scheuring, I. & Riedi, R.H. (1994) Application of multifractals to the analysis of vegetation pattern. Journal of Vegetation Science, 5, 489-496.
- [84] Schuster, P. & Sigmund, K. (1983) Replicator dynamics. Journal of Theoretical Biology 100, 533-38.
- [85] Shannon, C.E. & Weaver, W. (1963) The mathematical theory of communication. University of Illinois Press.
- [86] Smith, W. J. (1977) The behavior of communicating, Harvard University Press.
- [87] Spector, L. & Luke, S. (1996) Culture enhances the evolvability of cognition. In Proceedings of the 1996 Cognitive Science Society Meeting.
- [88] Squire, L.R. (1992) Memory and the hippocampus: A synthesis from findings with rats, monkeys, and humans. Psychological Review, 99, 195-231.

- [89] Thagard, P. (1980) Against evolutionary epistemology. PSA 1980, ed. P.D. Asquith & R.N. Giere 187-96.
- [90] Tomasello, M. Kruger, A.C., & Ratner, H.H. (1993) Cultural learning. Behavioral and Brain Sciences, 16, 495-552.
- [91] Von Neumann, J. (1966) Theory of self-reproducing automata, University of Illinois Press.
- [92] Vygotsky, L.S. (1978) Mind in society: The development of higher psychological processes, eds. M. Cole, V. John-Steiner, S. Scribner & E. Souberman, Harvard University Press.
- [93] Wallas, G. (1926) The art of thought, Harcourt, Brace & World.
- [94] Walsh, R. & Vaughan, F. (1993) Paths beyond ego, Jeremy P. Tarcher/Perigee.
- [95] Weisberg, R.W. (1986) Creativity: Genius and other myths, Freeman.
- [96] Wright, S. (1969) Evolution and the genetics of populations, University of Chicago Press.

# AUTOCATALYTIC CLOSURE IN A COGNITIVE SYSTEM: A TENTATIVE SCENARIO FOR THE ORIGIN OF CULTURE

## Target Article by Gabora on Origin Culture

By  
**Liane Gabora**  
Center Leo Apostel,  
Brussels Free University,  
Krijgskundestraat 33,  
1160 Brussels, Belgium  
<http://www.vub.ac.be/CLEA/liane/>  
[lgabora@vub.ac.be](mailto:lgabora@vub.ac.be)

## Abstract

This target article presents a speculative model of the cognitive mechanisms underlying the transition from episodic to mimetic (or memetic) culture with the arrival of *Homo Erectus*, which Donald (1991) claims paved the way for the unique features of human culture. The model draws on Kauffman's (1993) theory of how an information-evolving system emerges through the formation of an autocatalytic network. Though originally formulated to explain the origin of life, Kauffman's theory also provides a plausible account of how discrete episodic memories become woven into an internal model of the world, or world-view, that both structures, and is structured by, self-triggered streams of thought. Social interaction plays a role in (and may be critical to) this process. Implications for cognitive development are explored.

Keywords: *abstraction, animal cognition, autocatalysis, cognitive development, cognitive origins, consciousness, cultural evolution, memory, meme, mimetic culture, representational redescription, world-view.*

## I. INTRODUCTION

1. The subject of cultural origins is usually approached from an archeological perspective. For example, by dating artifacts such as tools we learn approximately when humans acquired the ability to make and use those tools. This target article takes a more cognitive approach (see also Barkow et al. 1992; Donald 1991, 1993a, 1993b; Tomasello et al. 1993; Tooby & Cosmides 1989). It outlines a theory of the psychological mechanisms underlying the major cognitive transition that, as Donald (1991) proposes, made possible the characteristic complexity and ingenuity of human culture.

2. The theory proposed here was inspired by an idea originally put forward to explain the origin of life. The origin of life and the origin of culture might appear at first glance to be very different problems. However, at a gross level of analysis they amount to the same thing: the bootstrapping of a system by which information patterns self-replicate, and the selective proliferation of some variants of these self-replicating patterns over others. The theory is thus consistent with the perspective of culture as a form of evolution (Dawkins 1975; Gabora 1997). In keeping with this

evolutionary framework, the term "meme" is used to refer to a unit of cultural information as it is represented in the brain. Thus meme refers to anything from an idea for a recipe to a memory of one's uncle to a concept of size to an attitude of racial prejudice. The rationale for lumping together episodic memories and symbolic abstractions is that they are both food for thought, units of information that can be drawn upon to invent new memes or to clarify existing ones. Memes that have been implemented as actions, vocalizations, or objects are referred to as artifacts.

3. The basic line of reasoning in this paper goes as follows. The bottleneck to cultural evolution appears to be the capacity for a self-sustained stream of thought that both structures and is structured by an internal model of the world, or world-view. It is this capacity that enables us to plan and predict, to generate novelty, and to tailor behavior according to context. The question is: Until discrete memories have been woven into a conceptual web, how can they generate a stream of thought? And conversely, until a mind can generate a stream of thought, how does it weave its memories into a world-view? Kauffman's proposal that life originated with the self-organization of a set of autocatalytic polymers suggests a mechanism for how this comes about. Much as catalysis increases the number of different polymers, which in turn increases the frequency of catalysis, reminding events increase meme density by triggering symbolic abstraction, which in turn increases the frequency of remindings. And just as catalytic polymers undergo a phase transition to a state where there is a catalytic pathway to each polymer present, and together they constitute a self-replicating set, memes undergo a phase transition to a state where each meme is retrievable through a pathway of remindings/associations, and together they constitute a transmittable world-view. In the origin of life scenario, since reactions occur in parallel, autocatalytic closure increases sharply as the ratio of reactions to polymers increases. In the cultural analog, however, the retrieval and invention of memes is funneled through an attention/awareness mechanism, which introduces a bottleneck. Therefore, this transition occurs gradually, as increasingly abstract concepts are perceived and their implications percolate through the memetic network. Social interaction and artifacts facilitate the process, and ensure that the continued evolution of memes does not hinge on the survival of any particular meme host.

## **II. BACKGROUND: THE ORIGIN OF LIFE AND ITS CULTURAL ANALOG**

4. This section will present background material relevant to the central thesis. We begin with a comparison of minds that are and are not able to sustain cultural evolution, and we draw on what is known about human cognition to make some hypotheses concerning the differences between them. We then turn to the paradox of the origin of life, and show how autocatalysis provides a potential solution.

### **II.1 A TRANSITION IN COGNITIVE CAPACITY**

5. In *Origins of the Modern Mind*, Donald (1991) argues convincingly that the capacity for abstract thought is the bottleneck of cultural evolution, and that it came about during the transition from episodic to mimetic culture following the arrival of *Homo Erectus* approximately 1.7 million years ago. Before this time the human memory system was like that of a primate

[FOOTNOTE 1], limited to the storage and cued retrieval of specific episodes. Donald accordingly uses the term episodic to designate a mind in which episodic memory is the only memory system there is. An episodic mind is capable of social attribution, insight and deception, and is sensitive to the significance of events. With much training it can learn arbitrary stimulus-response associations (such as pointing at a token of a certain shape to obtain food). However, it cannot invent symbols or abstractions on its own, or experiment with them. It has great difficulty accessing memories independent of environmental cues, and is unable to improve skills through self-cued rehearsal.

6. In contrast, the mimetic mind has, built upon its episodic foundations, a multimodal modeling system with a self-triggered rehearsal loop. In other words, it can retrieve and recursively operate on memories independent of environmental cues, a process referred to by Karmiloff-Smith (1992, 1994) as representational redescription. Redescribing an episode in terms of what is already known roots it in the network of understandings that comprise the world-view, and the world-view is perpetually restructured as new experiences are assimilated, and new symbols and abstract concepts are invented as needed. A mimetic individual is able to rehearse and refine skills, and therefore exhibits enhanced behavioral flexibility and more precise control over intentional communication. The upshot is cultural novelty. Mime, play, games, tool making, and reproductive memory, says Donald are thus manifestations of the same superordinate mimetic controller. The appearance of sophisticated stone tools, long-distance hunting strategies, and migration out of Africa, as well as the rapid increase in brain size at this time (Bickerton 1990; Corballis 1991; Lieberman 1991), are cited as evidence for the transition from episodic to mimetic culture. Donald claims it is not clear that the mimetic controller must be localized in any single anatomical structure, although it must have functional unity. Mimetic ability seems to encompass a broad panoply of skills associated with several distinct regions of the brain. Since miming accounts for only a small part of what the mimetic mind can do, and since mimetic skill seems to boil down to the capacity to evolve memes, we will use the term "memetic" instead of "mimetic."

7. Donald's proposal is invaluable in that spurs us to consider the cognitive basis of culture, but it leaves us hanging as to what sort of functional reorganization could turn an episodic mind into a memetic one. In particular, it leaves us with a nontrivial problem of origins. In the absence of representational redescription, how are relationships established amongst memes so that they become a world-view? And until a memory incorporates relationships between stored items, how can one meme evoke another which evokes another, etc., in a stream of representational redescription? We know that the brains of an ancestral tribe somehow turned into instruments for the variation, selection, and replication of memes. What happened to get the ball rolling, to enable the process of memetic evolution to take hold? When Grogga, a member of this tribe, had her first experience, there were no previously stored episodes to be reminded of, just external and internal stimuli (such as hunger). As episodes accumulated in her memory, occasionally it happened that an instant of experience was so similar to some stored episode that a retrieval process occurred, and she was reminded of that past episode. Perhaps the retrieval elicited a learned response. For example, the sight of a bumpy, red gourd might have reminded her of a bumpy, yellow gourd that her brother once used to carry water. Embedded in this recollection was the refreshing taste of the water he shared with her. The memory might have inspired her to use the red gourd to carry water. But since her memory consisted only of stored episodes, no

abstractions, this is the only kind of influence it could exert; her awareness was dominated by the stimuli of the present moment. At some point in her life, however, she managed to wilfully direct her attention not to a particular sensory stimulus, nor to the performance of a biological drive satisfying action, but to a chain of symbol manipulation. She kept this stream of thought going long enough to refine a concept or perspective, or invent a novel artifact. But if you need an interconnected world-view to generate a stream of thought, and streams of thought are necessary to connect individual memes into a world-view, how could one have come into existence without the other?

## II.2 COGNITIVE ATTRIBUTES THAT ENABLE ABSTRACT THOUGHT

8. The first step toward an answer is to elucidate as best we can given present knowledge, the cognitive mechanisms that distinguish a memetic mind from an episodic one. This section describes a minimal, biologically plausible cognitive architecture that could qualify as memetic. This cognitive architecture is a best guess, drawn from evidence in cognitive science, neuroscience, and artificial intelligence and the knowledge that the memetic mind grew out of (and therefore the potential for it was implicit in) the architecture of the episodic mind. It has the following attributes. First, it can integrate inputs from the sensorium and the drives with inputs from memory, and it can dispense commands to the motor system. Second, the memory is sparse, content addressable, distributed, modular, and habituates to repeated inputs. This enables it to generate abstractions. Finally, it can manipulate abstractions, consciously and recursively. We discuss each attribute briefly, explain how together they accomplish a memetic task, and then specify what is most likely lacking in the episodic mind.

9. Integration of Memory, Stimuli, and Drives. Vital to both the episodic and memetic mind is a means of integrating sensations, drives, and stored memories to produce a seamless stream of conscious experience and purposeful motor action. The place where this information is coordinated need not correspond to a single anatomical structure (though it is often suggested that the intralaminar nuclei of the thalamus are involved). We will adopt Kanerva's (1988) term, the focus, since it does not imply any commitment regarding centrality or global penetration. We assume that the states of the neurons that comprise the focus determine the content and phenomenal qualities of an instant of awareness. A meme is then a high-dimensional vector of difference relations (or continuous variables) that either is or has been encoded in an individual's focus.

10. Sparse Memory. Our sensory apparatus can register a tremendous amount of information. Where  $n$  is the number of features the senses can distinguish,  $N$ , the number of memes that could potentially be hosted by the focus,  $= 2^{**n}$  for Boolean variables ( $n$  is infinitely large for continuous variables). For example, if  $n = 1,000$ ,  $N = 2^{**1,000}$  memes [FOOTNOTE 2]. Assuming  $n$  is large,  $N$  is enormous, so the memory is sparse in that the number of locations  $L$  where memes can be stored is only a small fraction of the  $N$  perceivable memes. In other words, neural pathways leading out from the focus do not receive inputs from each of its  $n$  slots, but from some fraction of them. The number of different memes actually stored at a given time,  $s$ , is constrained by  $L$ . The set of all possible  $n$ -dimensional memes a mind is capable of storing can be represented as the set of vertices (if features assume only binary values) or points (if features assume continuous values) in an  $n$ -dimensional hypercube, where the  $s$  stored memes occupy

some subset of these points. The distance between two points in this space is a measure of how dissimilar they are, referred to as the Hamming distance. Kanerva (1988) makes some astute observations about this memory space. The number of memes at Hamming distance  $d$  away from any given meme is equal to the binomial coefficient of  $n$  and  $d$ , which is well approximated by a Gaussian distribution. Thus, if meme X is 111...1 and its antipode is 000...0, and we consider meme X and its antipode to be the poles of the hypersphere, then approximately 68% of the other memes lie within one standard deviation ( $\sqrt{n}$ ) of the equator region between these two extremes. As we move through Hamming space away from the equator toward either Meme X or its antipode, the probability of encountering a meme falls off sharply by the proportion  $\sqrt{n}/n$ .

11. In fact the space of possibilities is even larger if we assume that the mind rarely if ever pays attention to all the stimulus dimensions it is capable of detecting. Therefore the number of dimensions the focus pays attention to,  $n$ , is only a subset of the maximum,  $M$ . The strength of the signal on a neural pathway from memory, senses, or drives must surpass some threshold before the dimension of the focus it activates is attended. Since the memory can now store memes of any length up to  $M$ , the number of possible memes is:

$$(1) N = 2^{**}(M+1)-2 = \text{approx: } 2^{**}(M+1)$$

The bottom line is: the memory would probably have to be larger than the number of particles in the universe to store all the permutations of sensory stimuli it is capable of registering. It is therefore sparse.

12. Distributed Representation. In a sparse memory, the probability that a given meme in the focus is identical to one in storage is virtually zero, which means that retrieval should be impossible. In connectionist networks, this problem is solved by distributing the storage of a meme across many locations. Likewise, each location participates in the storage of many memes. The focus is represented as input/output nodes, memory locations as hidden nodes, and their pattern of connectivity as weighted links. An input touches off a pattern of activation which spreads through the network until it relaxes into a stable configuration, or achieves the desired input-output mapping using a learning algorithm. The output vector is determined through linear summation of weighted inputs. Thus a retrieved meme is not activated from a dormant state, but reconstructed. This approach is necessary if we aim to model cognition at a fine-grained level of resolution -- down to the threshold of human discrimination. There is a saying: You never step into the same stream twice; this applies to streams of thought as well as streams of water. Right now I am retrieving a memory of eating cinnamon toast; tomorrow I may retrieve the same memory. But today it is colored by today's mood, today's events; tomorrow it will be experienced slightly differently. It is not the exact same information pattern conjured up time and again. The reconstructive approach enables the memory to abstract a prototype, fill in missing features of a noisy or incomplete pattern, or create a new meme on the fly that is more appropriate to the situation than any meme it has actually experienced (Rumelhart and McClelland 1986). For example, if an autoassociative network has been fed vectors in which feature one is present whenever feature two is present, and vice versa, it will respond to an input that lacks information about feature one, such as \*101, by generating 1101. It may never actually have encountered 1101 before, but given that in its world there exists a correlation between features one and two, this is an appropriate response. In addition to associations between inputs and outputs of features,

the network has learned a higher-level association between two features. In effect, it contains more information than has been fed into it.

13. One problem with distributed representation is that unless stored patterns are perfectly orthogonal, they interfere with one another, a phenomenon known as crosstalk. This is solved by restricting the storage region. For instance, in Kanerva's (1988) Sparse Distributed Memory (SDM) model, a meme is stored in all locations within a hypersphere of addresses surrounding the ideal address. The smaller the Hamming distance between two memes, the more their storage locations overlap, so the higher the probability they are retrieved simultaneously and blended in the focus. A more sophisticated way of implementing this idea, for which there is neurobiological support, is to use a radial basis function (RBF) (Clothiaux et al. 1991; Hancock et al. 1991; Willshaw & Dayan 1990). Once again a hypersphere of locations is activated, but this time activation is maximal at the center of the RBF and tapers off in all directions according to a (usually) Gaussian distribution (see FIGURE 1). Where  $x$  is an  $i$ -dimensional input vector,  $k$  is the center of the RBF, and  $s$  is the width of the Gaussian, hidden nodes are activated as follows:

$$(2) F(x) = e^{-\text{SIGMA}[(x-m)/s]^2}$$

By carving out a hypersphere in memory space, one part of the network can be modified without affecting the capacity of other parts to store other patterns. The further a stored meme is from  $k$ , the less activation it not only receives but in turn contributes to the next evoked meme, and the more likely that its contribution is cancelled out by that of other memes. In neural networks, suitable values for  $k$  and  $s$  are found during a training phase. In the brain,  $k$  values could be modified by changing the pattern of neuronal interconnectivity. Decreasing neuron activation thresholds would increase  $s$ .

14. Organized Modularity. Another way of avoiding cross-talk is to induce a division of labor amongst competing subnetworks; in other words, to make the memory modular (Nowlan 1990; Jacobs et al. 1991). There is abundant evidence of modularity in the brain; its preservation in phylogenetic history suggests that it is not arbitrary. We assume that (1) the world we live in is highly patterned and redundant and that (2) this pattern and redundancy is reflected in the connectivity of the neurons where memes are stored. After birth there is a large-scale pruning of neurons. It seems reasonable that the surviving subset of the  $M$  possible inputs to each neural pathway is determined by biological and cultural selective pressures, instead of at random. These pressures sculpt the pattern of neuronal connectivity such that the  $L$  (out of  $N$  possible) locations can store most of the memes we stand a chance of encountering. This means that in practice the sparseness of the memory does not interfere with its representational capacity. It also means that the probability that a given stimulus activates a retrieval event is not as low as the statistics suggest.

15. Content Addressability. A computer reads from memory by simply looking at the address in the address register and retrieving the item at the location specified by that address. The sparseness of human memory prohibits this kind of one-to-one correspondence. However, content addressability can be feigned, as follows. The feature pattern that constitutes a given meme causes some neurons leading out from the focus to be excited and others to be inhibited. The ensuing chain reaction activates memory neurons where the meme gets stored. The address

of a memory neuron amounts to the pattern of excitatory and inhibitory synapses from focus to storage that make it fire, so there is a systematic relationship between the information content of a meme and the locations it activates. Thus, embedded in the neural environment that supports their informational integrity, memes act as implicit pointers to other memory locations. These pointers prompt the dynamic reconstruction of the next meme to be subjectively experienced, which is statistically similar to the one that prompted it. As a result, the entire memory does not have to be searched in order for a gourd to remind Grogg of a previously encountered gourd. It is worth stressing that there is no search taking place, just information flowing through a system displaced from equilibrium. The current instant of experience activates certain neurons, which in turn activate certain other neurons, which leads to the distributed storage of that experience, which activates whatever else is stored in those locations, which then merges with any salient information from the senses and drives to form the next instant of experience, etc. in an ongoing cycle. What emerges is that the system appears to retrieve memories that are similar, or concepts that are relevant, to the current experience. But that is not magic; it is simply a side effect of the fact that correlated memes get stored in overlapping locations.

16. Habituation. We do not want an ongoing stimulus, such as the sound of rain, to recursively evoke reminders of rain. The nervous system avoids this kind of perseveration as follows. First, neurons have a refractory period during which they cannot fire, or their response is greatly attenuated. Second, they team play; the responsibility for producing a response is shared by a cooperative group of neurons such that when one is refractory another is active. If exactly the same neurons are stimulated repeatedly, they all become refractory, and there is little or no response.

17. Capacity for Symbol Manipulation. The connectionist methods described above are examples of the subsymbolic approach to cognition, which works best for modelling perceptual and low-level cognitive phenomena. These include detecting, representing, and responding flexibly to patterns of correlation, learning fuzzy categories, and solving simple constraint satisfaction problems. Subsymbolic processing makes the world easier to navigate. But the world contains additional structure that our brains are not hard-wired to capture. Therefore, even after memes are stored in memory, they are clustered, rather than being uniformly distributed throughout the space of possible memes; and they contain implicit predicate logic relationships. This is where symbolic processing is useful. Symbolic models of cognition focus on the serial and potentially recursive application of logical operations on symbols, without attempting to represent their internal structure. They are particularly good at modeling the high-level cognitive abilities that are unique to memetic minds, such as planning and deductive reasoning. Arguments for a reconstructive view of retrieval notwithstanding, highly abstract concepts that have been used thousands of times, such as "space" or "equal" or "is," would be unlikely to emerge from memory retaining the associations of any particular usage. Thus it seems reasonable to begin with the working hypothesis that subsymbolic processing predominates for low-level, parallel, automatically generated cognitive phenomena, and that symbolic processing provides a satisfactory approximation for many high-level, serial, consciously directed aspects of cognition. (Creative processes may draw heavily on both.)

18. Let us now examine how a cognitive architecture with these attributes would accomplish a specific task. Consider the situation wherein the sight of a rotting, striped, bumpy, red gourd

reminds Grog of the striped, bumpy, yellow gourd her brother used as to carry water, which generates the desire to have water readily available in the cave. Grog slashes the top off the red gourd and scoops water into it. To her dismay, the water leaks out through a soft decay spot. Just out of sight lies the intestine of a recently killed water buffalo. What sort of cognitive dynamics would prompt Grog to tie one end of the intestine and use it as a waterbag? It is unlikely that the ability to classify gourd and knotted intestine as potentially substitutable instances of the category container is hard-wired. No one in Grog's tribe has previously conceived of an intestine as a container, so social learning is not an option. This task involves a number of difficult skills including abstract reasoning, uncued retrieval, redescription, and manual dexterity. It lies beyond the horizon of what the episodic mind can accomplish.

19. The sight of the red gourd is registered as a vector of features in Grog's focus. This vector determines which synapses leading out from the focus are excited and which are inhibited, which determines how activation flows through her memory network, which in turn determines the hypersphere of locations where "red gourd" is stored. The process of storing to these neurons triggers retrieval from these neurons of whatever has been stored in them. Of course, nothing is retrieved from them if, after red gourd is stored, Grog's attention is directed toward some stimulus or biological drive. But to the extent that memory contributes to the next instant of awareness, storage of red-gourd activates the retrieval of not only red-gourd itself but all other memes stored in the same locations. The next meme to be encoded in the focus is found by evaluating the contributions of all retrieved memes feature-by-feature. Whereas the retrieved copies of red gourd reinforce one another, the other retrieved memes contribute less, and are statistically likely to cancel one another out. They do not cancel out exactly, however, unless the distribution of stored memes within the hypersphere of activated locations is uniformly dense. In this case it is not. The meme yellow gourd container, which got stored when Grog saw her brother carrying water in a yellow gourd, acts as an attractor. The result is that the next meme ends up being red gourd container. Though it is a reconstructed blend, something Grog has never actually experienced, it can still be said to have been retrieved from memory.

20. Grog pours water into the red gourd and, as we know, it leaks out. Her mental model of the world was in error; not all gourds can transport water. Stymied, memory is probed again, with knowledge of relationships between objects and attributes guiding the process. The second probing occurs with intensified activation of the pathway leading from the concave slot of the focus, and inhibition of the permeable slot. Let us now focus on the portion of Grog's memory that deals with four discrete features: bumpy, striped, permeable, and concave (FIGURE 2). These lie on the  $x_1$ ,  $x_2$ ,  $x_3$ , and  $x_4$  axes, respectively, and a black dot represents the center of a distributed hypersphere where a meme is stored. The second probing of memory activates a slightly different set of locations, which evoke the abstract category container, the class of objects that are concave and impermeable, and for which the attributes bumpy and striped are irrelevant. Container was implicit in the meme-space; it covered the two-dimensional yellow region of the original hypercube. More generally, we can view an  $n$ -dimensional meme space as a set of nested hypercubes, such that implicit in the outermost hypercube of memes with all  $n$  dimensions there exist hypercubes of memes with  $n-1$  dimensions,  $n-2$  dimensions, etc. Armed with the category container, Grog dips into memory again to discover what else constitutes a member of this category. The closest thing she can come up with is intestine. Symbol manipulation now kicks in. She realizes that the intestine is impermeable and almost concave.

Knotted at one end, an intestine would constitute another member of the category container. She could therefore carry water in it. She runs off to fetch the intestine.

21. The foregoing discussion may be wrong in the details, but hopefully it captures the gist of memetic cognition. Now we ask: what is the episodic mind lacking? Some are tempted to say that the ability of animals to respond appropriately to salient stimuli, and even learn arbitrary sensorimotor associations, indicates some capacity for symbolic thought. However, animals' learned behavior is stereotyped and brittle: it cannot be adapted to new contexts, which suggests that they use symbols only in an iconic sense. They give no indication of engaging in streams of thought that reorganize memes in ways that make their similarities and differences more explicit. They could not retrieve the memory that an intestine is in the cave, much less realize that it was relevant to the goal of transporting water. Our best-guess model of cognition suggests a number of possible reasons. First, the resolution of the perceptual apparatus might not be high enough to capture enough features of salient stimuli (large M). Second, there might not be enough memory locations to keep these distinctions intact during storage (large L). Third, the density s/N of stored memes might be too low. In other words, there might not be enough different basins of attraction for memes to slide into, or not enough of these attractors might be occupied. Another possibility is that the neuron activation threshold is too high (and thus s too narrow). The end result is the same in all cases: rarely is there a stored meme within retrievable distance of a given meme in the focus. Thus the memory does not encode relationships, so rarely can a stream of interrelated thoughts ensue. In fact, these explanations are connected. M limits L, which in turn limits s. And since if s 0, the memory only retrieves memes identical to the content of the focus and therefore cannot form abstractions, s also limits s.

22. At this point we are in a position to reframe our central question. We want to know how a mind comes to assume a self-sustained stream of thought that progressively shapes and is shaped by, a world-view. Abstract thinking requires each meme that enters the focus to activate one or more memes already stored in memory enough to evoke a retrieval. The memory must be traversed with tunnels that connect related concepts like an apple crisscrossed with wormholes. However, representational redescription is the process that puts related memes within working memory reach of one another; it is what recognizes abstract similarities and restructures the memory to take them into account. How do you get the wormholes without the worms?

### **II.3 THE ORIGIN OF LIFE PARADOX**

23. We will put aside the question of cultural origins for now, and turn to the problem of biological origins. The paradox of the origin of life can be stated simply: if living things come into existence when other living things give birth to them, how did the first living thing arise? That is, how did something complex enough to reproduce itself come to be? In biology, self-replication is orchestrated through an intricate network of interactions between DNA, RNA, and proteins. DNA is the genetic code; it contains instructions for how to construct various proteins. Proteins, in turn, both catalyze reactions that orchestrate the decoding of DNA by RNA, and are used to construct a body to house and protect all this self-replication machinery. Once again, we have a chicken-and-egg problem. If proteins are made by decoding DNA, and DNA requires the catalytic action of proteins to be decoded, which came first? How could a system composed of complex, mutually dependent parts come into existence?

24. The most straightforward explanation is that life originated in a prebiotic soup where, with enough time, the right molecules collided into one another at the same time and reacted in exactly the right ways to create the DNA-RNA-protein amalgam that is the crux of life as we know it. Proponents argue that the improbability of this happening does not invalidate the theory because it only had to happen once; as soon as there was one self-replicating molecule, the rest could be copied from this template. Miller (1955) increased the plausibility of this hypothesis by showing that amino acids, from which proteins are made, form spontaneously when a reducing [FOOTNOTE 3] mixture of oxygen, hydrogen, carbon, nitrogen, water, and ammonia is subjected to high energy. These molecules were all likely to have been present on the primitive earth, and energy could have come in the form of electric discharges from thunderstorms, ultraviolet light, or high temperatures generated by volcanoes. Other experiments have shown that the molecular constituents of DNA and RNA, as well as the fatty acids from which membranes are constructed, can be formed the same way. Unfortunately, the complexity of the DNA-RNA-protein structure is so great, and in the earth's early atmosphere the concentrations of the necessary molecules were so dilute, that the probability of life originating this way is infinitesimally low. Hoyle and Wickramasinghe (1981) likened it to the probability that a tornado sweeping through a junkyard would spontaneously assemble a Boeing 747.

25. The less complex something is, the more feasible its spontaneous generation. The discovery of ribozymes -- RNA molecules that, like proteins, are capable of catalyzing chemical reactions -- brought the hope that the first living molecule had been found. With ribozymes you would not need DNA or proteins to establish a self-replicating lineage; these RNA molecules would do the job of all three. In practice, however, self-replication of RNA is fraught with difficulties. It tends to fold back on itself creating an inert, tangled mess (Joyce, 1987). Furthermore, the probability of a ribozyme assembling spontaneously from its components is remote (Orgel 1987), and even if it managed to come into existence, in the absence of certain error-detecting proteins found in all modern-day organisms, its self-replication capacity would inevitably break down in the face of accumulated error over successive generations (Eigen and Schuster, 1979). Thus it is far from obvious how the chain of self-replicating systems that eventually evolved into you and me got started.

#### **II.4 THE AUTOCATALYSIS THEORY OF THE ORIGIN OF LIFE**

26. Despite the myriad difficulties encountered attempting to get ribozymes to self-replicate, the idea behind it -- that life originated in a simple self-replicating system that over time evolved into the familiar DNA-RNA-protein complex -- was a good one. Once you have some sort of self-replicating structure in place, anything whatsoever that accomplishes this basic feat, natural selection can enter the picture and help things along. Kauffman (1991) suggested that knowing as much as we do about what life is like now may actually get in the way of determining how it began. He accordingly decided to focus on how to get from no life at all to any kind of primitive self-replicating system, and to hand the problem of getting from there to DNA-based life, over to natural selection (as well as self-organizing processes). Given the conditions present on earth at the time life began, how might some sort of self-replicating system have arisen? His answer is that life may have begun not with a single molecule capable of replicating itself, but with a set of collectively self-replicating molecules. That is, none of the molecules could replicate itself, but each molecule could induce the replication of some other molecule in the set, and likewise, its

own replication was induced by some other member of the set. This kind of dual role as both ingredient (or stimulant) and product of different chemical reactions is not uncommon for polymers such as protein and RNA molecules.

27. Polymers induce each other's replication by acting as catalysts. Catalysts speed up chemical reactions that would otherwise occur very slowly. An autocatalytic system is a set of molecules which, as a group, catalyze their own replication. Thus, if A catalyzes the conversion of X to B, and B catalyzes the conversion of Y to A, then A + B comprise an autocatalytic set (FIGURE 3). In an environment rich in X and Y, A + B can self-replicate. A set of polymers wherein each molecule's formation is catalyzed by some other molecule is said to exhibit catalytic closure.

28. It is of course highly unlikely that two polymers A and B that just happened to bump into one another would happen to catalyze each other. However, this is more likely than the existence of a single polymer catalyzing its own replication. And in fact, when polymers interact, their diversity increases, and so does the probability that some subset of the total reaches a critical point where there is a catalytic pathway to every member. To show that this is true we must show that the number of reactions by which they can interconvert increases faster than their total number. Given polymers made up of, say, two different kinds of monomers, of up to a maximum length of M monomers each, then N, the number of polymers, is  $2^{**M+1}$  as per equation (1). Thus as M increases -- which it obviously does, since two of the longest polymers can always join to form a longer one -- the number of polymers increases exponentially. Now we need to show that the number of reactions between them increases even faster. We will be conservative and consider only cleavage (e.g. 110  $\rightarrow$  1 + 10) and ligation (e.g. 1 + 10  $\rightarrow$  110) reactions on oriented polymers (such as protein and RNA fragments). The number of possible reactions R is the product of the number of polymers of a certain length times the number of bonds, summed across all possible lengths:

$$(3) R = (2^{**M})(M-1)+(2^{**M+1})(M-2)+\dots+(2^{**M-(M-2)})(M-(M-1))$$

<ftp://coglit.psy.soton.ac.uk/pub/psycoloquy/1998.volume.9/Pictures/eqn3a.jpg>

$$= \text{SIGMA}[(2^{**n})(n-1)], n=2 \rightarrow M$$

Dividing equation (3) by equation (1), we find that as M increases, the ratio of reactions to polymers increases by a factor of M-2. This means that if each reaction has some probability of getting carried out, the system eventually undergoes a transition to a state where there is a catalytic pathway to each polymer present. The probability of this happening shifts abruptly from highly unlikely to highly likely as R/N increases. This kind of sharp phase transition is a statistical property of random graphs and related systems such as this one. Random graphs consist of dots, or nodes, connected to each other by lines or edges. As the ratio of edges to nodes increases, the probability that any one node is part of a chain of connected nodes increases, and chains of connected nodes become longer. When this ratio reaches approximately 0.5, almost all these short segments become cross-connected to form one giant cluster (FIGURE 4). Plotting the size of the largest cluster versus the ratio of edges to nodes yields a sigmoidal curve. The larger the number of nodes, the steeper the vertical portion of this curve (referred to as the percolation threshold).

29. Of course, even if catalytic closure is theoretically possible, we are still a long way from knowing that it is the correct explanation for the origin of life. How likely is it that an autocatalytic set would have emerged given the particular concentrations of chemicals and atmospheric conditions present at the time life began? In particular, some subset of the R theoretically possible reactions may be physically impossible; how can we be sure that every step in the synthesis of each member of an autocatalytic set actually gets catalyzed? Kauffman's response is: if we can show that autocatalytic sets emerge for a wide range of hypothetical chemistries (i.e., different collections of catalytic molecules), then the particular details of the chemistry that produced life do not matter so long as it falls within this range. We begin by noting that, much as several different keys sometimes open the same door, each reaction can be catalyzed by not a single catalyst, but a hypersphere of catalytic molecules, with varying degrees of efficiency. So we assign each polymer an extremely low a priori random probability P of catalyzing each reaction. The lower the value of P, the greater M must be, and vice versa. Kauffman shows that the values for M and P necessary to achieve catalytic closure with a probability of  $< 0.999$  are highly plausible given the conditions of early earth. Experimental evidence for this theory using real chemistries (Lee et al. 1996, 1997; Severin et al. 1997), and computer simulations (Farmer, et al. 1986) have been unequivocally supportive. Farmer et al. showed that in an artificial soup of information strings capable of cleavage and ligation reactions, autocatalytic sets do indeed arise for a wide range of values of M and P.

30. An interesting question explored in this simulation is: once a set of polymers has achieved autocatalytic closure, does that set remain fixed, or is it able to incorporate new polymer species? Farmer et al. found that some sets were "subcritical" (unable to incorporate new polymers) and others were "supracritical" (incorporated new polymers with each round of replication). Which of these two regimes a particular set fell into depended on P, and the maximum length of the food set polymers.

31. Now the question is: supposing an autocatalytic set did emerge, how would it evolve? The answer is fairly straightforward. It is commonly believed that the primitive self-replicating system was enclosed in a small volume such as a coacervate or liposome to permit the necessary concentration of reactions (Oparin 1971; Morowitz 1992; Cemin & Smolin, in press). Since each molecule is getting duplicated somewhere in the set, eventually multiple copies of all molecules exist. The abundance of new molecules exerts pressure on the vesicle walls. This often causes such vesicles to engage in a process called budding, where it pinches off and divides into two twins. So long as each twin contains at least one copy of each kind of molecule, the set can continue to self-replicate indefinitely. Replication is far from perfect, so an offspring is unlikely to be identical to its parent. Different chance encounters of molecules, or differences in their relative concentrations, or the arrival of new food molecules, could all result in different catalysts catalyzing a given reaction, which in turn alters the set of reactions to be catalyzed. So there is plenty of room for heritable variation. Error catastrophe is unlikely because, as mentioned earlier, initially each reaction can be catalyzed not by a single catalyst but by a hypersphere of potential catalysts, so an error in one reaction does not have much effect on the set at large [FOOTNOTE 4]. Selective pressure is provided by the affordances and limitations of the environment. For example, say an autocatalytic set of RNA-like polymers arose. Some of its offspring might have a tendency to attach small molecules such as amino acids (the building blocks from which proteins are made) to their surfaces. Some of these attachments inhibit

replication, and are selected against, while others favor it, and are selected for. We now have the beginnings of the kind of genotype-phenotype distinction seen in present-day life. That is, we have our first indication of a division of labor between the part of the organism concerned with replication (in this case the RNA) and the part that interacts with the environment (the proteins)

32. The autocatalytic origin of life theory circumvents the chicken-and-egg problem by positing that the same collective entity is both code and decoder. This entity does not look like a code in the traditional sense because it is a code not by design but by default. The code is embodied in the physical structures of the molecules; their shapes and charges endow them with propensities to react with or mutually decode one another such that they manifest external structure, in this case a copy of its collective self. Since autocatalytic sets appear to be a predictable, emergent outcome in any sufficiently complex set of polymers, the theory suggests that life is an expected outcome rather than a lucky long-shot.

### **III. THE EMERGENCE OF AUTOCATALYSIS IN A COGNITIVE SYSTEM**

33. We have taken a look at two paradoxes -- the origin of culture and the origin of life -- which from hereon will be referred to as OOC and OOL respectively. The parallels between them are intriguing. In each case we have a self-replicating system composed of complex, mutually interdependent parts, and since it is not obvious how either part could have arisen without the other, it is an enigma how the system came to exist. In both cases, one of the two components is a storehouse of encoded information about a self in the context of an environment. In the OOL, DNA encodes instructions for the construction of a body that is likely to survive in an environment like the one in which its ancestors survived. In the OOC, an internal model of world encodes information about the self, the environment, and the relationships between them. In both cases, decoding a segment of this information storehouse generates another class of information unit that coordinates how the storehouse itself gets decoded. Decoding DNA generates proteins that, in turn, orchestrate the decoding of DNA. Retrieving a memory or concept from the world-view and bringing it into awareness generates an instant of experience, a meme, which in turn determines which are the relevant portion(s) of the world-view to use in constructing the next instant of experience. In both cases it is useful to think of the relevant class of information units as states in an information space, each of which can act on a hypersphere of other states. In the memory model it was the hypersphere of related memes, and in the OOL model it was the hypersphere of potential catalysts.

34. We have argued that the most likely bottleneck in the OOC is the establishment of a network of inter-related memes, a world-view, that progressively shapes and is shaped by a stream of self-triggered thought. We want to determine how such a complex entity might come to be. Donald claims that the transition from episodic to memetic culture would have required a fundamental change in the way the brain operates. Drawing from the OOL scenario presented above, we will explore the hypothesis that meme evolution begins with the emergence of a collective autocatalytic entity that acts as both code and decoder. This idea was mentioned briefly in Gabora (1996a; 1996b; 1997); here it is fleshed out it in greater detail.

### III.1 ESTABLISHING THE CAPACITY FOR ABSTRACTION

35. In the OOL case we asked: what was lying around on the primitive earth with the potential to form some sort of self-replicating system? The most promising candidate was catalytic polymers, the molecular constituents of either protein or RNA. Here we ask an analogous question: what sort of information unit does the episodic mind have at its disposal? It has memes, specifically memories of episodes. Episodic memes then constitute the food set of our system.

36. Next we ask: what happens to the food set to turn it into a self-replicating system? In the OOL case, food-set molecules catalyzed reactions on each other that increased their joint complexity, eventually transforming some subset of themselves into a collective web for which there existed a catalytic pathway to the formation of each member molecule. An analogous process might conceivably transform an episodic mind into a memetic one. Food-set memes activate redescriptions of each other that increase their joint complexity, eventually transforming some subset of themselves into a collective web for which there exists a retrieval pathway to the formation of each member meme. Much as polymer A brings polymer B into existence by catalyzing its formation, meme A brings meme B into conscious awareness by evoking it from memory. As in Section II.2, a retrieval can be a reminding, a redescription of something in light of new contextual information, or a creative blend or reconstruction of many stored memes.

37. How might Grog's mind have differed from that of her ancestors such that she was able to initiate this kind of transformation? In the OOL case, it was crucial that the polymers be catalytic. We simply gave each polymer a small, random probability  $P$  of catalyzing each reaction. In the OOC case, we assume that each of Grog's  $L$  memory locations where the  $s$  memes are stored has a RBF with a Gaussian distribution of width  $s$  centered on it. Thus the probability that one meme evokes another is determined by  $s$  rather than a random probability  $P$ . Let us consider what would happen if, due to some genetic mutation, Grog's activation threshold were significantly lower than average for her tribe. Thus  $s$  is wider, which means that a greater diversity of memes are activated in response to a given experience, and a larger portion of the contents of memory merge and surface to awareness in the next instant. Since the memory is content-addressable, when meme  $X$  goes fishing in memory for meme  $X$ , sooner or later this large hypersphere is bound to catch a stored meme that is quite unlike  $X$ . For example, let us say that Grog sees rabbits every day, so there are lots of rabbit memories stored in her brain. For simplicity, let us say they consist of a sequence of ten 0's followed by a five bit long variable sequence. She happens to look off in the distance and see a grazing water buffalo, which gets represented in her focus as 000000011101010. The buffalo meme will be referred to as meme  $X$ . Because the hypersphere is wide, all the rabbit memories lie close enough to meme  $X$  to get evoked in the construction of  $X$  (as is  $X$  itself). Since all the components from which  $X$  is made begin with a string of seven zeros, there is no question that  $X$  also begins with a string of seven zeros. These positions might code for features such as has eyes, eats, etc. The following set of three 1's in the rabbit memes are cancelled out by the 0s in the buffalo memes, so in  $X$  they are represented as \*. These positions might code for features such as floppy ears. The last five bits constituting the variable region are also statistically likely to cancel one another out. These code for other aspects of the experience, such as, say, the color of the sky that day. So  $X$  turns out to be the meme 0000000\*\*\*\*\*, the generic category animal; which then gets stored in memory in the next iteration. This evocation of animal by the buffalo episode is not much of a stream of thought, and it does not bring her much closer to an interconnected conceptual web, but it is an

important milestone. It is the first time she ever derived a new meme from other memes, her first creative act.

### III.2 ESTABLISHING A STREAM OF THOUGHT

38. Although lowering the neuron activation threshold was what enabled Groga to create an abstraction, the penalty for having too low a threshold would be very high, because successive thoughts would not necessarily be meaningfully related to one another. Abstract thought, unlike episodic thought, cannot rely on the continuity of the external world (i.e., if a desk is in front of you now it is likely to still be in front of you now) to lend coherence to conscious experience. Too low a threshold might be expected to result in a cognitive rendition of superconductivity, where lowering resistance increases correlation distance and thus a perturbation to any one pattern percolates through the system and affects even distantly related patterns. (The free-association of a schizophrenic seems to correspond to what one might expect of a cognitive system with this property (see Weisberg 1986).) However, if the threshold is extremely high, such that distributions do not overlap, the attended meme must be identical to one stored in memory to evoke a retrieval. To produce a steady stream of meaningfully related yet potentially creative remindings, the threshold must fall within an intermediate range. This is consistent with Langton's (1992) finding that the information-carrying capacity of a system is maximized when its interconnectedness falls within a narrow regime between order and chaos.

39. Thus thoughts do not leap from one unexplored territory of meme-space to another, but meander from one meme to a similar one in a region that has proven fruitful and is therefore exceptionally clustered with memes. This not only increases the frequency of remindings and abstractions, it provides a thread of continuity linking one meme to the next. Organized modularity also enhances continuity by precluding the activation of irrelevant memes. Since statistical similarity is preserved across sequentially evoked memes in a train of thought, thinking can be viewed as an internal form of meme self-replication. It could be argued that the correlation between consecutive memes is so low that this hardly deserves to be called a form of self-replication. One would not want consecutive memes to be identical. Surely Eigen and Schuster's error catastrophe argument applies here; that is, the copying fidelity of this process is so low that errors would quickly accumulate and in no time the lineage would die. But this argument does not apply. The only reason it is a pitfall for biological evolution is that copying error tends to impair the capacity to self-replicate. So long as offspring are as good as their parents at reproducing themselves, and live long enough to do so, it does not matter how much error is introduced from one generation to the next. It is only when a generation dies without having reproduced that there is a problem. In the biological world, once something is dead it cannot bring forth life [FOOTNOTE 5]. But in memetic evolution this is not necessarily the case. To show why this is so, say that half-way through the train of consecutive memes in Einstein's brain that culminated in the theory of relativity, a tiger burst in through the window. The correlation between the relativity meme of one instant and the tiger-perception-meme of the next instant would be almost zero. This momentous memetic lineage would come to a screeching halt. But would it be lost forever? No. Sooner or later, once the tiger situation was taken care of, the relativity stream of thought would resume itself. Memory (and external artifacts) function as a memetic sperm bank, allowing a defunct ancestral line to be brought back to life and to resume

self-replication. The upshot is that in culture you can get away with a much higher error rate than in biology.

40. Cultural evolution has not only an internal form of replication, but also an internal means of generating variation. In a stream of thought, consecutive memes are not exact replicas; each meme is a variation of its predecessor. It also has an internal form of selection. Selection comes in the form of drives, needs, attention-focusing mechanisms, and the associative organization of memory, which constrain how one meme evokes another. Thus all the components of an evolutionary process take place in the mind of an isolated individual. The memory-driven generation of a stream of correlated memes can be viewed as a coevolutionary relationship between replication, variation, and selection, and the process of representational redescription can itself be redescribed as the selective generation of variant replicants. Embedded in the outer, inter-individual sheath of memetic evolution we find a second intra-individual sheath, where the processes of replication, variation, and selection are not spatiotemporally separated but intimately intertwined. Together they weave a stream of thought, one meme fluidly transmuting into the next.

41. Thus the semantic continuity of a stream of thought makes memory navigable despite its sparseness. Once "animal" has been evoked and stored in memory, the locations involved habituate and become refractory (so, for instance, animal does not recursively evoke animal). However locations storing memes that have some animal features, but that were not involved in the storage of animal, are still active. Thus animal might activate "tiger" which might evoke "hyena" etc., strengthening associations between the abstract category and its instances. Other abstractions, such as container, form in analogous fashion. As Grogga accumulates both episodic memes and abstractions, the probability that any given attended meme is similar enough to some previously stored meme to activate it increases. Therefore reminding acts increase in frequency, and eventually become streams of remindings, which get progressively longer. Grogga is now capable of a train of thought. Her focus is no longer just a spot for coordinating stimuli with action; it is now a forum for abstractive operations that emerge through the dynamics of iterative retrieval.

### **III.3 AUTOCATALYTIC CLOSURE IN A WEB OF SPARSE, DISTRIBUTED MEMES**

42. We have seen how our best-guess model of human cognition achieves a stream of thought. How do we know that streams of thought will induce a phase transition to a critical state where for some subset of memes there exists a retrieval pathway to each meme in the subset? In the OOL case, we had to show that  $R$ , the number of reactions, increases faster than  $N$ , the number of polymers. We found that  $R/N$  increased by a factor of  $M^2$ , where  $M$  was the maximum number of monomers per polymer. Because of the highly parallel nature of this system, it was reasonable to equate potential reactions with actual reactions, and therefore to assume that the new polymers resulting from these reactions actually exist (and can themselves partake in reactions). Similarly, we now want to show that some subset of the memes stored in an individual's mind inevitably reach a critical point where there is a path by which each meme in that subset can get evoked. But here, it is not reasonable to assume that all  $N$  perceptible memes actually exist (and can therefore partake in retrieval operations). Their number is severely curtailed by the number of memory locations, the variety of perceptual experience, and the fact

that meme retrieval, though distributed at the storage end, is serial at the awareness end. The rate at which streams of thought reorganize the memetic network is limited by the fact that everything is funneled through the focus; we can only figure one thing out at a time. This presents a bottleneck that was not present in the OOL scenario. As a result, whereas OOL polymers underwent a sharp transition to a state of autocatalytic closure, any analogous transition in inter-meme relatedness is expected to take place gradually. So we need to show that  $R$ , the diversity of ways one meme can evoke another, increases faster than not  $N$  but  $s$ , the number of memes that have made it through this bottleneck. That is, as the memory assimilates memes, it comes to have more ways of generating memes than the number of memes that have explicitly been stored in it.

43. This brings us to another complication, which further prolongs cognitive development. Since short, simple molecules are more abundant and readily formed than long, complex ones, in the OOL case it made sense to expect that the food set molecules were the shortest and simplest members of the autocatalytic set that eventually formed. Accordingly, in simulations of this process the direction of novelty generation is outward, joining less complex molecules to form more complex ones through AND operations. In contrast, the memetic food set molecules are complex, consisting of all attended features of an episode. In order for them to form an interconnected web, their interactions tend to move in the opposite direction, starting with relatively complex memes and forming simpler, more abstract ones through OR operations. The net effect of the two is the same: a network emerges, and joint complexity increases. But what this means for the OOC is that there are numerous levels of autocatalytic closure, which convey varying degrees of world-view interconnectedness and consistency on their meme host. These levels correspond to increased penetration of the  $(n-1, n-2)$ -dimensional nested hypercubes implicit in the memory space. Since it is difficult to visualize a set of nested, multidimensional hypercubes, we will represent this structure as a set of concentric circles, such that the outer skin of this onion-like structure represents the hypercube with all  $n$  dimensions, and deeper circles represent lower-dimensional hypercubes. Obviously, not all the nested hypercubes can be shown. The points of our original hypercube are represented as points along the perimeter of these circles, and  $k$  values (centermost location where a meme is stored) are shown as large, black dots. The outermost shell encodes memes in whatever form they are in the first time they are consciously encountered. This is all the episodic mind has to work with. In order for one meme in this shell to evoke another, they have to be extremely similar at a superficial level. In a memetic mind, however, related concepts are within reach of one another because they are stored in overlapping hyperspheres.

44. Under what conditions does that  $R$  increase faster than  $s$ ? As it turns out, abstraction plays a crucial role. To determine how abstraction affects  $R$ , let us assume for the moment that memory is fully connected. Clearly this is not the case, but this simplification illustrates some trends which also apply to a memory wherein sparseness is compensated for by restricted distributed activation. We will be conservative and limit the sort of retrieval event under consideration to abstraction, and the redescription of a meme as an instance of an abstraction (including analogical thought). Abstractions have  $n$  dimensions, where  $n$  ranges from a minimum of  $m$  to a maximum of  $M$ .  $R_{\text{sub}A}$ ), the number of ways a retrieval can occur through abstraction, equals the number of retrieval paths allowed by an  $n$ -dimensional abstraction, multiplied by the number of  $n$ -dimensional abstractions, summed over all values of  $n$  from  $m$  to  $M-1$ . The number of

retrieval paths equals the number of memes that are instances of an n-dimensional abstraction =  $2^{**}(M-n)$ . The number of n-dimensional abstractions is equal to the binomial coefficient of M and n. The result is multiplied by two, since an abstraction can evoke an instance, and likewise, an instance can evoke an abstraction.

$$(4) R_{\text{sub}A} = 2(2^{**}(M-m)[M::m]+2^{**}(M-(m+1))[M::m+1]+\dots 2[M::M-1]) \\ = 2(\Sigma(2^{**}M-n[M::n])) \text{ for } (n-m) \rightarrow M-1$$

The key thing to note is that lower-dimensional memes allow exponentially more retrieval paths. Abstraction increases s by creating a new meme, but it increases R more, because the more abstract the concept, the greater the number of memes a short Hamming distance away (since  $|x_i - k_i| = 0$  for the irrelevant dimensions). A second thing to note is that the number of abstractions at a given value of n increases up to  $M/2$ . Taken together, these points mean: the more deeply a mind delves into lower-dimensional abstractions, the more the distribution rises and becomes skewed to the left. The effect is magnified by the fact that the more active a region of meme space, the more likely that an abstraction will be positioned there, and thus abstractions beget abstractions recursively through positive feedback loops. Reminding incidents also contribute to R. Hence the more likely it is that some meme will get activated and participate in a given retrieval. So, whereas R increases as abstraction makes relationships increasingly explicit, s levels off as new experiences have to be increasingly unusual in order to count as new and get stored in a new constellation of locations. Furthermore, when the carrying capacity of the memory is reached, s plateaus, but R does not. Thus as long as the neuron activation threshold is large enough to permit abstraction and small enough to permit temporal continuity, the average value of n decreases. Sooner or later the system is expected to reach a critical percolation threshold such that R increases exponentially faster than s.

45. So long as R does indeed eventually increase faster than s, Grog's memory becomes so densely packed that any meme that comes to occupy the focus is bound to be close enough in Hamming distance to some previously stored meme(s) to evoke it. The memory (or some portion of it) is holographic in the sense that there is a pathway of associations from any one meme to any other. Together they form an autocatalytic set. What was once just a collection of isolated memories is now a structured network of concepts, instances, and relationships -- a world-view. This most primitive level of autocatalytic closure is achieved when stored episodes are interconnected by way of abstractions just a few onionskin layers deep, and streams of thought zigzag amongst these superficial layers. A second level occurs when relationships amongst these abstractions are identified by higher-order abstractions at even deeper onionskin layers, etc. Once Grog's memory has defined an abstraction, identified its instances, and chunked them together in memory, she can manipulate the abstraction much as she would a concrete episode. Reflecting on an idea amounts to reflecting it back and forth off onionskin layers of varying depths, refining it in the context of its various interpretations. The conscious realization of the logical operators and, or and not, are expected to significantly transform Grog's world-view by enabling conscious symbol manipulation. Other particularly useful abstractions such as mine, depth, or time, as well as frames (Barsalou, in press), scripts (Schank & Abelson 1977), and schemas (Minsky 1985), are also expected to induce reorganization. Just as in a sand pile perched at the proverbial edge of chaos a collision between two grains occasionally triggers a chain reaction that generates a large avalanche, one thought occasionally triggers a chain reaction of others that

dramatically reconfigure the conceptual network. Rosch's (1978) work on basic level categories suggests that the way we organize information is not arbitrary but emerges in such a way as to maximize explanatory power. It would not be surprising to find that the number of categories and their degree of abstraction exhibit the same kind of power law relationship as one finds in other emergent systems (Bak, Tang, & Weisenfeld 1988).

46. How does an interconnected world-view help Grogña manifest the skills that differentiate a memetic mind from an episodic one? The capacity to maintain a stream of self-triggered memes enables her to plan a course of action, and to refine behavior by incorporating kinesthetic feedback into a meme sequence. The ability to generate abstractions opens up a vast number of new possibilities for Grogña. It allows her to incorporate more of the structure of the world into her mental model of it. This increases behavioral flexibility by enabling her to define elements of the world in terms of their substitutable and complementary relationships. (For example, if she usually makes bows out of wood X, but she cannot find any wood X, and if wood Y is as strong and flexible as wood X, then wood Y might substitute for wood X.) The power of abstraction also enables her to express herself artistically by extricating memes from the constraints of their original domain and filtering the resulting pattern through the constraints of other domains. For example, she can translate the scene before her into a sequence of motor commands that render it as a cave painting or stone carving, or transform the pattern of information that encodes the sorrow she experienced at her child's death into a song. Finally, abstraction enables Grogña to communicate with others through spoken or nonverbal forms of language. This brings us to the issue: how does the world-view replicate?

### **III.4 SOCIAL INTERACTION FACILITATES WORLD-VIEW EMERGENCE**

47. Now that we have an autocatalytic network of memes, how does it self-replicate? In the OOL scenario, polymer molecules accumulate one by one until there are at least two copies of each, and their shell divides through budding to create a second replicant. In the OOC scenario, Grogña shares concepts, ideas, stories, and experiences with her children and tribe members, spreading her world-view meme by meme. Categories she had to invent on her own are presented to and experienced by others much as any other episode. They are handed a shortcut to the category; they do not have to engage in abstraction to obtain it.

48. Recall how the probability of autocatalysis in Kauffman's simulation could be increased by raising either the probability of catalysis or the number of polymers (since it varied exponentially with M). Something similar happens here. Eventually, once enough of Grogña's abstractions have been assimilated, her tribe members' memories become so densely packed that even if their neuron activation thresholds are higher than Grogña's, a version of Grogña's world-view snaps into place in their minds. Each version resides in a different body and encounters different experiences. These different selective pressures sculpt each copy of Grogña's original world-view into a unique internal model of the world. Small differences are amplified through positive feedback, transforming the space of viable world-view niches. Individuals whose activation threshold is too small to achieve world-view closure are at a reproductive disadvantage, and over time eliminated from the population. Eventually the proclivity for an ongoing stream of thought becomes so firmly entrenched that it takes devoted yogis years of meditation to even briefly arrest it. There is selective pressure for parents who monitor their child's progress in abstraction

and interact with the child in ways that promote the formation of new abstractions the next level up. Recall the discussion in section II.4 concerning the incorporation of new polymer species by supracritical autocatalytic sets. This kind of parental guidance is analogous to handcrafting new polymers to be readily integrated into a particular autocatalytic set; in effect it keeps the child's mind perpetually poised at a supracritical state. Language provides a means for individuals to mutually enrich one another's world-views, and to test their world-views against each other, and in so doing prompt one another to penetrate deeper and deeper into the onion.

49. Clearly social processes are an integral component of cultural evolution. In fact the origin of culture is often unquestioningly equated with onset of the capacity for social transmission. However, as many authors have pointed out (e.g. Darwin 1871; Plotkin 1988), although transmission is wide-spread in the animal kingdom, no other species has anything remotely approaching the complexity of human culture. Moreover, although in practice transmission plays an important role, is it crucial? If, for example, you were the only one human left on the planet, but you were able to live forever, would meme evolution grind to a halt? If you were to come up with some unique dance, would you not be exploring the space of possible dance memes even though no one was watching? If you found an ingenious way to fix a broken toaster, would you not still have invented a novel meme?

50. In biological evolution, transmission and replication go hand in hand; genetic information gets replicated and is transmitted to offspring. But that is not necessarily the only way of getting the job done. In memetic evolution, the most obvious means of meme replication is through social processes such as teaching or imitation, but there is a second form of replication that takes place within an individual. We noted earlier that in the mind of someone engaged in a stream of thought, each meme is a statistically similar variant of the one that preceded and prompted it. It is in this sense that they self-replicate without necessarily being transmitted to another host. Thus there need not necessarily be more than one individual for a meme to evolve. Nevertheless, although intra-individual meme replication is sufficient for evolving memes, the culture of a single individual would be extremely impoverished compared to that of a society of interacting individuals, because the number of memes increases exponentially as a function of the number of interacting memetic-level individuals. As a simple example, a single memetic individual who invents ten memes is stuck with just those ten memes. A society of ten interacting individuals, only one of whom has reached the memetic stage and can invent ten memes, is no better off; there are still just ten memes. In a society of ten noninteracting individuals, each of whom invents ten memes but does not share them, each individual still has only ten memes. But in a society where each of the ten interacting individuals invents ten memes and shares them, each individual ends up with one hundred memes. The bottom line is: culture as we know it, with its explosive array of meaningful gestures, languages, and artifacts, depends on both intra-individual and inter-individual meme replication.

51. In fact it is possible that cognitive closure as described above first occurred at the level of the group, within a collection of interacting individuals, and cognitive closure at the level of the individual came into existence some time later. (The two need not be mutually exclusive; it is possible that group-level closure could persist after the arrival of individual-level closure.)

## **IV. IMPLICATIONS**

52. In this section we explore some implications of the autocatalytic cognition hypothesis. This is the most speculative section of what is admittedly a speculative paper.

### **IV.1 WHY DON'T ANIMALS EVOLVE CULTURE?**

53. As noted in Section III.2, the penalty for having too low a neuron activation threshold is very high. Each meme has little relevance to the one that preceded it, and thinking is so garbled that survival tasks are not accomplished. On the other hand, too high a neuron activation threshold is not life-threatening. The focus is virtually always affected by external stimuli or internal drives, and memory is reserved for recalling how some goal was accomplished in the past. This may be the situation present in most brains on this planet, and though not harmful, it has its own drawbacks. A stream of thought dies out long before it produces something creative. However, this may not be of practical consequence to other species. The advantages of a stream of thought would largely be lost on nonhuman animals because they have neither the vocal apparatus nor the manual dexterity and freedom of upper limbs to implement creative ideas. (Language, for example, drastically increases the degrees of freedom of what can be expressed.) No matter how brilliant their thoughts were, it would be difficult to do something useful with them. Moreover, in an evolutionary line there is individual variation, so the lower the average activation threshold, the higher the fraction of individuals for which it is so low that they do not survive. It seems reasonable to suggest that animals are not prohibited from evolving complex cognition *a priori*, but that there is insufficient evolutionary pressure to tinker with the threshold until it achieves the requisite delicate balance to sustain a stream of thought, or to establish and refine the necessary feedback mechanisms to dynamically tune it to match to the degree of conceptual fluidity needed at any given instant. It may be that humans are the only species for which the benefits of this tinkering process have outweighed the risks.

### **IV.2 PSYCHOLOGICAL CONSIDERATIONS**

54. Initially a child is expected to be unselective about meme acquisition, since (1) it does not know much about the world yet, so it has no basis for choosing, and (2) its parents have lived long enough to reproduce, so they must be doing something right. However, just as importing foreign plants can bring ecological disaster, the assimilation of a foreign meme can disrupt the established network of relationships amongst existing memes. Therefore the child develops mental censors that ward off internalization of potentially disruptive memes. Censors might also be erected when a meme is embarrassing or disturbing or threatening to the self-image (Minsky 1985). This could be accomplished by temporarily increasing the activation threshold so as to prematurely terminate the meme's assimilation into the world-view. Much as erecting a fence increases the probability that people will stay on either one side or the other, censorship warps the probability that a meme will partake in any particular stream of thought, such that the individual either avoids the censored meme or dwells on it excessively. (This seems to be consistent with our bipolar attitude toward highly censored subjects such as aggression and sexuality.) Thus censorship precludes incorporation of a meme into the autocatalytic portion of the memory, and thereby interferes with its holographic nature.

55. Categorization creates new lower-dimension memes, which makes the space denser, and increases susceptibility to the autocatalytic state. On the other hand, creating new memes by combining stored memes could interfere with the establishment of a sustained stream of thought by decreasing the modularity of the space, and thereby decreasing density. If cross-category blending indeed disrupts conceptual networking, one might expect it to be less evident in younger children than in older ones, and this expectation is born out experimentally (Karmiloff-Smith 1990). There is evidence of an analogous shift in human history from an emphasis on ritual and memorization toward an emphasis on innovation (Donald, 1991). As world-views become more complex, the artifacts we put into the world become more complex, which necessitates even more complex world-views, etc.; thus a positive feedback cycle sets in.

56. We mentioned that animals are hard-wired to respond appropriately to certain stimuli, as are humans. However, the ability of humans to develop world-views with which they can make decisions about what action to take may obviate the need for some of this hard-wiring. Genetic mutations that interfere with certain regions of hard-wiring may not be selected against, and may actually be selected for, because in the long run they promote the formation of concepts that generate the same responses but can be used in a more context-sensitive manner. However, this increases the amount of computation necessary to achieve a workable world-view.

#### **IV.3 THE ROLE OF AUTOCATALYSIS IN EVOLUTIONARY PROCESSES**

57. Returning briefly to the origin-of-life puzzle, recall that traditional attempts to explain how something as complex as a self-replicating entity could arise spontaneously entail the synchronization of a large number of vastly improbable events. Proponents of such explanations argue that the improbability of the mechanisms they propose does not invalidate them, because it only had to happen once; as soon as there was one self-replicating molecule, the rest could be copied from this template. However, Kauffman's theory that life arose through the self-organization of a set of autocatalytic polymers suggests that life might not be a fortunate chain of accidents but rather an expected event.

58. Although there is much evidence for this hypothesis, definitive proof that it is the correct explanation of how life originated will be hard to come by. However, if we are interested in the more general question of how information evolves, we now have another data point, another evolutionary process to figure into the picture. Culture, like biological life, is a system that evolves information through variation, selection and replication. In fact, it has two layers of replication, one embedded in the other, and to actualize the inner layer of replication, all members of the culture must establish their own personal world-view, which generates their own unique autonomous stream of sequentially activated self-similar patterns. Consistent with Kauffman's assertion that the bootstrapping of an evolutionary process is not an inherently improbable event, the "it only had to happen once" argument does not hold water here because the cultural analog of the origin of life takes place in the brain of every young child. Autocatalysis may well be the key to the origin of not only biological evolution, but any information-evolving process.

## V. CONCLUSIONS

59. Cultural evolution presents a puzzle analogous to the origin of life: the origin of an internal model of the world that both generates and is generated by streams of self-sustained, internally driven thought. In this target article we have explored a plausible scenario for how cultural evolution, like biological evolution, could have originated in a phase transition to a self-organized web of catalytic relations between patterns. TABLE 1 presents a summary of how the components of the proposed theory of cultural autocatalysis map onto their biological counterparts.

<b>EVOLUTIONARY SYSTEM</b>	<b>BIOLOGY</b>	<b>CULTURE</b>
INFORMATION UNIT	Polymer Molecule	Meme
INTERACTION	Catalysis	Reminding, retrieval, Reconstruction
AUTOCATALYTIC SET	Catalytically closed set of polymer molecules; primitive organism	Network of inter-related memes; World-view
REPLICATION	Duplication of each molecule, segregation via budding	Correlation between consecutive memes; Social learning; teaching, Associations, drives; social pressures, affordances and limitations of environment
SELECTION	Physical constraints on molecules, affordances and limitations of environment	affordances and limitations of environment
VARIATION	Novel food molecules, nonspecificity of catalysis, replication error	Sensory novelty, blending; expressive constraints, misunderstanding, etc.

**TABLE 1: Components of an autocatalytic theory of biological evolution and their cultural counterparts**

60. The scenario outlined here is nascent. Putting the pieces together would require the cooperation of neuroscientists, developmental psychologists, cognitive scientists, sociologists, anthropologists, archeologists, and perhaps others. Nevertheless, I know of no other serious attempt to provide a functional account of how memetic evolution got started. Whether or not the scenario outlined here turns out to be precisely correct, my hope is that it draws attention to the problem of cultural origins, suggests what a solution might look like, and provides a concrete example of how we gain a new perspective on cognition by viewing it as an architecture that has been sculpted to support a second evolutionary process, that of culture.

## **ACKNOWLEDGEMENTS**

I would like to thank David Chalmers, Merlin Donald, Bruce Edmonds, Harold Edwards, Stuart Kauffman, Francis Heylighen, Norman Johnson, Wolfgang Klimesch, William Macready, and Mario Vaneechoutte for helpful discussion and comments on the manuscript.

## **FOOTNOTES**

1. Although this may not be completely accurate; see Donald (1993) and accompanying commentary.
2. This number is perhaps better appreciated when we realize that its magnitude is 10,300
3. In a reducing atmosphere there is no free oxygen present. The presence of ferrous (FeO) rather ferric (Fe<sub>2</sub>O<sub>3</sub>) iron in primitive rock leads us to believe that the earth's atmosphere was reducing when life began. (It is no longer so today.)
4. See Kauffman (1993) for an interesting discussion of why error catastrophe becomes a serious problem as the parts of the system becomes more co-adapted.
5. The niche it filled still exists, so there is still selective pressure for it to evolve all over again. But the information has to re-evolve (as opposed to being retrieved from storage).

## **REFERENCES**

- Bak, P., Tang, C., & Weisenfeld, K. (1988) Self-organized criticality Phys. Rev. A 38: 364.
- Barkow, J. H., Cosmides, L. & Tooby, J. (1992) The adapted mind: Evolutionary psychology and the generation of culture. Oxford University Press.
- Barsalou, L. W. (In press) Perceptual symbol systems. Submitted to Behavioral and Brain Sciences.
- Bickerton, D. (1990) Language and species, University of Chicago Press.
- Cemin, S. C. & Smolin, L. (in press) Coevolution of membranes and channels: A possible step in the origin of life. Submitted to Journal of Theoretical Biology (October, 1997).
- Corballis, M. C. (1991) The lopsided ape: Evolution of the generative mind, Cambridge University Press.
- Darwin, C. (1871) The descent of man, John Murray Publications.
- Donald, M. (1991) Origins of the modern mind, Harvard University Press.

- Donald, M. (1993a) Precis of Origins of the modern mind: Three stages in the evolution of culture and cognition. *Behavioral and Brain Sciences* 16: 737-791.
- Donald, M. (1993b) Human cognitive evolution: What we were, what we are becoming. *Social Research* 60 (1): 143-170.
- Eigen, M. & Schuster, P. (1979) The hypercycle: A principle of natural self-organization. Springer.
- Farmer, J. D., Kauffman, S. A., & Packard, N. H. (1987) Autocatalytic replication of polymers. *Physica D*, 22 (50).
- Hoyle, F. & Wickramasinghe, N. C. (1981) Evolution from space, Dent.
- Joyce, G. F. (1987) Non-enzymatic, template-directed synthesis of informational macromolecules. In: Cold Spring Harbor Symposia on Quantitative Biology 52. Cold Spring Harbor Laboratory, New York.
- Gabora, L. (1996a) Culture, evolution and computation. In Proceedings of the Second Online Workshop on Evolutionary Computation. Society of Fuzzy Theory and Systems.
- Gabora, L. (1996b) A day in the life of a meme. *Philosophica*, 57, 901-938. Invited manuscript for special issue on concepts, representations, and dynamical systems.
- Gabora, L. (1997) The origin and evolution of culture and creativity. *Journal of Memetics: Evolutionary Models of Information Transmission* Vol. 1, Issue 1.
- Hinton, G. E. & J. A. Anderson. (1981) Parallel models of associative memory. Lawrence Erlbaum Associates.
- Kanerva, P. (1988) Sparse distributed memory, MIT Press
- Karmiloff-Smith, A. (1990) Constraints on representational change: Evidence from childrens drawing. *Cognition* 34: 57-83.
- Karmiloff-Smith, A. (1992) Beyond modularity: A developmental perspective on cognitive science, MIT Press.
- Karmiloff-Smith, A. (1994). Precis of Beyond modularity: A developmental perspective on cognitive science. *Behavioral and Brain Sciences* 17 (4): 693-745.
- Kauffman, S. A. (1993) Origins of order, Oxford University Press.
- Langton, C. G. (1992) Life at the edge of chaos. In: Artificial life II, eds. C. G. Langton, C. Taylor, J. D. Farmer & S. Rasmussen, Addison-Wesley.

Lee, D. H., Granja, J. R., Martinez J. A., Severin, K. & Ghadiri, M. R. (1996) A Self-Replicating Peptide. *Nature*. 382, 8 August, 525-528.

Lee, D. H., Severin, K., Yokobayashi & Ghadiri, M. R. (1997) Emergence of Symbiosis in Peptide Self-Replication through a Hypercyclic Network. *Nature*, 390, 11 December: 591-594.

Lieberman, P. (1991) *Uniquely human: The evolution of speech, thought, and selfless behavior*, Harvard University Press.

Miller, S. L. (1955) Production of some organic compounds under possible primitive earth conditions. *Journal of the American Chemistry Society* 77:23-51.

Minsky, M. (1985) *The society of mind*, Simon and Schuster.

Morowitz, H. J. (1992) *The beginnings of cellular life*, Yale University Press.

Olton, D. S. (1977) Spatial memory. *Scientific American* 236, 82-98

Oparin, A. I. (1971) Routes for the origin of the first forms of life. *Sub. Cell. Biochem.* 1:75.

Orgel, L. E. (1987) Evolution of the genetic apparatus: A review. In *Cold Spring Harbor Symposia on Quantitative Biology Vol. 52*. Cold Spring Harbor Laboratory, New York.

Plotkin, H. C. (1988) *The role of behavior in evolution*, MIT Press.

Rumelhart, D. E. & J. L. McClelland. (1986) *Parallel distributed processing: Explorations in the microstructure of cognition*. MIT Press.

Rosch, E. (1987) Principles of categorization. In E. Rosch & B. B. Lloyd (Eds.) *Cognition and categorization*. Erlbaum.

Schank, R., & Abelson, R. P. (1977) *Scripts, plans, goals, and understanding: An inquiry into human knowledge structures*. Lawrence Erlbaum Associates.

Severin, K., Lee, D. H., Kennan, A. J. & Ghadiri, M. R. (1997) A Synthetic Peptide Ligase. *Nature*, 389, 16 October, 706-709.

Tomasello, M., Kruger, A. C. & Ratner, H. H. (1993) Cultural learning Behavioral and Brain Sciences, 16, 495-552.

Tooby, J. & Cosmides, L. (1989) Evolutionary psychology and the generation of culture, Part I. *Ethology and Sociobiology*, 10, 29-49.

Weisberg, R.W. (1986) *Creativity: Genius and other myths*, Freeman.

## REVIEW

By

Liane Gabora

Center Leo Apostel, Brussels Free University, Krijgskundestraat 33, B-1160 Brussels, Belgium

Of The

**The Meme Machine**

Susan Blackmore

Oxford: Oxford University Press, 1999, Cloth: ISBN 0-19-850365-2

*The Meme Machine* follows through on Dawkins' (1976) fascinating suggestion that culture, like biology, evolves through the processes of variation, selection, and replication. *TMM* does a nice job of laying out the basic idea - that, much as organic life evolves into more complex forms through progressive adaptation to environmental constraints, ideas and artifacts build on what came before in response to the necessities of human survival. Like Richard Brodie's (1996) *Virus of the Mind*, and Aaron Lynch's (1996) *Thought Contagion*, *TMM* explores how this kind of evolutionary perspective on culture can shed light on various aspects of the human experience, such as why we talk so much, believe in alien abduction and fantasize about sex. (Though the chapter titled 'An orgasm saved my life' never gets around to explaining how an orgasm saved someone's life.)

### Is Imitation What Distinguishes Humans From Animals?

*TMM* is about imitation rather than innovation. Blackmore lays this out explicitly: "The thesis of this book is that what makes us different is our ability to imitate." (p. 3). The story goes: humans can imitate, and animals cannot, and that is why we alone have evolved culture. The idea that conformity, not creativity, is the hallmark of the human condition, will inevitably strike some readers as so unappealing and counter-intuitive, that they won't get past Chapter One. And in the end, the case she makes may not be convincing. Nevertheless, the possibility is at least worth considering, and it is interesting to see how it fares.

Blackmore begins by arguing that imitation is natural and ubiquitous in humans, but almost non-existent in animals. Though the first part of this argument is clearly true, the second is highly controversial. As it turns out, while the book was being written, the consensus was swinging in favor of the view that animals *do* imitate. (An article by Byrne and Russon 1998 and the accompanying commentary, provide an insightful review of this topic.) Despite the fact that animal culture is impoverished at best, they do appear to have some imitative capacity. So from the outset, there are doubts as to whether Blackmore will succeed in pulling her 'imitation drives culture' thesis off, but the reader doesn't know for sure.

The next few chapters spell out how memes function as replicators, and discusses how the world looks from the 'meme's eye view' (Dennett 1995). The 'imitation drives culture' hypothesis leads Blackmore to circumscribe what counts as a meme and can be culturally transmitted. For

example, she limits the transmission process to imitation of one human by another. If a child learns to peel a banana by watching her mother, a meme has replicated, but if the child learns this skill from a cartoon character on television, no replication has taken place. By the end of the book (particularly in the chapter on the internet) she relaxes this claim somewhat. Human-made artifacts now seem to play a role in her vision, though elements of the natural world still don't. Thus if a child figures out how to peel a banana by watching the petals of a flower unfold, her flower-inspired 'how to peel a banana' meme is *not* transmittable. In the blink of an eye, Blackmore discards the possibility that *any* experience can be food for thought and thus food for culture, on the grounds that it is "extremely confusing" (p. 45). The worldview implied by the Shroödinger equation is extremely confusing too, but its batting average as a predictor of experimental outcomes is unsurpassed. 'Confusing' is not synonymous with 'wrong'.

Blackmore also claims that "perceptions and emotions are not memes because they are ours alone and we may never pass them on." (p. 15). It follows that the feeling evoked by a painting depicting a stormy night at sea has no relationship to what the artist was feeling at the time ... that a teacher's attitude of compassion has no impact on the cultural dynamics of the classroom. Thus it isn't certain how Blackmore's narrow definition of meme clears up the confusion.

## The Origin of Culture

Chapter Six focuses on the dramatic increase in brain size that began two and a half million years ago during the transition from *Australopithecus* to *Homo*. It is here that hopes of pulling off the 'imitation drives culture' thesis take a nosedive. Blackmore correctly notes that the archaeological record reveals a sudden *increase* in tool variety at this time. She does not seem to be aware that this contradicts the thesis of her book. That is, if imitation were the bottleneck to culture, then prior to the origin of culture there would have been variation everywhere, and the onset of imitation would have funneled this variation in a few of the most useful directions. To bolster the thesis that imitation is the cultural bottleneck with archaeological evidence, Blackmore would have had to find a period of time where tool variety sharply *decreased*.

The evidence she cites is, in fact, consistent with the thesis that creativity, rather than imitation, was the bottleneck to culture. The lack of cultural complexity in animals - despite the evidence that they can imitate when put to the test - is also consistent with this proposal. Imitative capacity remains latent or hidden until there is variation for it to work on. In 'Meme and Variations', my computer model of cultural evolution, when I set the agents' ability to imitate to 1 and their ability to invent to 0, what happened is ... nothing (Gabora 1995). There has to be something worth imitating before the ability to imitate will manifest itself. Novelty can then breed more novelty. Or as one choreographer (whose name I forget) put it: "If we don't do what our predecessors did, we're doing what our predecessors did."

The material that follows, dealing with the relationship of memetics to language and various social issues such as beauty and birth control, is well done. It provides a wealth of intriguing alternatives to explanations offered by sociobiology and evolutionary psychology over the last few decades. Sharpening the contrast between these approaches - for example, contrasting the memetic explanation for gossip with that proposed by Barkow (1992) - would have made this part of the book even more valuable.

## **Memetics and Profound Stuff**

Like Dawkins, Dennett, Brodie and Lynch, Blackmore cannot resist going to great lengths to convince us that religion and the New Age movement are nothing but hotbeds of manipulative tricks. "Religions build theories about the world and then prevent them being tested. Religions provide nice, appealing and comforting ideas and cloak them in a mask of 'truth, beauty, and goodness'. The theories can then thrive despite being untrue, ugly, or cruel ... I do defend the idea that science, at its best, is more truthful than religion." (p. 202-3). One can't help but wonder here 'What about religion *at its best*? Didn't religion give birth to science?' I for one am getting bored with this kind of thing. Those who devote their lives to religion may similarly deride science and gloat over how concepts such as 'fallen from grace' and 'purification of the soul' are vastly more meaningful than concepts like 'Turing test', and 'corpus callosum'. Instead of pitting one against the other, it would have been more satisfying to hear about how different experiences open up different memetic niches, which in turn generate different experiences, i.e. the 'embryology' of the process by which a mind incorporates a given memeplex. I'd like to hear what is going on when a meme (whether it be 'evolution' or 'human redemption') percolates far into a conceptual network touching every meme it encounters, resulting in a more consistent or satisfying worldview. In fact there is some disparity between the 'science rules' attitude and the relative paucity of theory or data. If the title leads you to expect material on computer models, cognitive science, complexity, information theory, etc. you might be disappointed. There isn't a lot on the workings of the memetic machinery.

From there the book moves onto topics like consciousness, the meaning of life and the concept of self. Much of this is intriguing; for example, the cross-cultural comparison of near death experiences. However to define the 'self' as "a bunch of memes" (p. 231) is misleading (particularly if emotions and attitudes don't count as memes!) It's like saying a chair is just a bunch of sticks. Much as 'chairness' resides in the way sticks are organized, the self arises from the way memes are structured and interact with one another, something about which Blackmore could have said more.

## **Is Memetics a Second Form of Evolution or Just an Analogy?**

Blackmore does not feel obliged to dig her spurs too deep into evolutionary theory, on the grounds that borrowing concepts from biology can lead cultural theorists astray: "There need be no memetic equivalent of the phenotype or the vehicle, any more than there are equivalents for strictly genetic concepts like alleles, loci, mitosis, and meiosis." (p. 66). Repeatedly she insists "We must remember they are only analogies." (p. 42) However, this stance also discourages potentially fruitful directions. Consider what happens when someone familiar with the concept of snow skiing first hears of water skiing. He might continue to include 'snow' as vital to the concept of 'skiing', and view water skiing as *analogous* to skiing. Or he might generalize the definition of 'skiing' to include both snow skiing and water skiing. Neither is objectively more *correct*, but the generalized concept of skiing more readily invites the application of knowledge gained through the study of snow skiing to water skiing. This is not equivalent to saying that water skiing *is* snow skiing; it's just a way of organizing world knowledge that is less likely to waste time by reinventing the wheel. Let us suppose that, a few years before water skis hit the market, you had been a venture capitalist, and were approached by two inventors who had both

come up with the idea of water skis. One inventor had spent the last few years working for a company that makes snow skis, and talked at length about how the length and width of a ski affect glide and balance, why skis curl up at the tip, etc. The other knew little about snow skis and argued this didn't matter because snow and water are not the same thing, it's "just an analogy". Which inventor would you have been more likely to fund?

Similarly, we can either say that DNA, chromosomes, etc. are vital components of evolution, and memetic change is *analogous* to evolution, or we can redefine evolution as any process wherein the iterated variation and selection of information induces adaptation to environmental constraint, and view memetic change as a second form of evolution. In so doing, one finds that concepts such as fitness, epistasis, drift, mutation, morphology, niches and attractors provide an extremely useful scaffold upon which to investigate how ideas unfold as they are assimilated by one individual after another and this scaffold gives the ideas their own unique slant. (See Radnitzky and Bartley 1987 and Hull 1988a, 1988b.)

The example Blackmore provides to illustrate how biological concepts can mislead is the distinction between genotype and phenotype. Most memeticists maintain that something along these lines is useful; i.e. there is a need to differentiate between the mental representation of a meme in the mind, and the implementation of it as behavior, vocalization or artifact. Blackmore argues that this distinction does not account for the difference between copying the result (as when someone watches you make soup and then makes it themselves) and copying the instruction (as when you give someone a recipe for soup). In fact, this raises no problem at all for the genotype/phenotype distinction. Both soup and recipe are artifacts, phenotypic expressions of different but related mental representations. As one's understanding of biological concepts increases, the danger of misapplying them decreases.

The exception to Blackmore's avoidance of biological concepts are the two chapters on memetic altruism. The idea that not only genes, but memes, impel us to behave altruistically toward others who bear copies of them, is a promising answer to a problem that has plagued sociobiology for some time. Blackmore does a good job of explaining the concept, though it would have been appropriate to mention that it has been around for some time (Heylighen 1992, 1992, Gabora 1996, 1997, 1998a, Evers 1998).

## Wrapping Up

I was not convinced of the pivotal role given to imitation, but you may well be. I am perhaps not the most objective of readers. First, because I harbour some possibly overly-idealistic memes about the human condition. And second because my own research supports the alternative proposition that creativity was the bottleneck to culture. A story can be told. Whereas biological creativity is random and 'breadth-first' - i.e. generate as many variants as possible and hopefully one of them will do the trick - cultural creativity is non-random and 'depth-first' - i.e. generate novelty *strategically*, using knowledge of relationships (Gabora, 1996, 1997). For a mind to be capable of strategic creativity, its initially discrete memories and stimulus-response associations must transform into an interconnected conceptual web, or worldview, wherein related concepts are connected by way of abstractions. In Gabora (1998b), I demonstrate how an interconnected

worldview could emerge through an autocatalytic process instigated by a genetic mutation which lead to decreased neuron activation threshold.

*TMM* is worth taking a look at. The issues it addresses are big and important, and memetics offers some compelling answers. Readers should be aware, though, that many of the ideas Blackmore discusses are better developed in other scholarly work in memetics, which often goes uncited here. Blackmore adds her own wrinkles, but she is not quite as alone out there on the wild, dangerous frontiers of human inquiry as the book might suggest.

If we accept the premise that memes evolve, let's use everything we've got and see if memetics has the potential to do for the cognitive and social sciences what the theory of natural selection has done for biology. *The Meme Machine* is a start. It is not inconceivable that the next century will bring forth more books on cultural evolution than this century has on biological evolution.

## References

- BARKOW J. H. 1992. Beneath New Culture is Old Psychology: Gossip and Social Stratification. In J. H. Barkow, L. Cosmides and J. Tooby, editors, *The Adapted Mind: Evolutionary Psychology and the Generation of Culture* Oxford University Press, New York, NY.
- BRODIE R. 1996. *Virus of the Mind: The New Science of the Meme*, Integral Press, Seattle, WA.
- BYRNE R. W. and A. Russon 1998. Learning by Imitation: A Hierarchical Approach, *Behavioral and Brain Sciences*, 21:667-721.
- DAWKINS R. 1976. *The Selfish Gene*, Oxford University Press, Oxford.
- DENNETT D. C. 1995. *Darwin's Dangerous Idea: Evolution and the Meanings of Life*, Allen Lane, London.
- EVERS J. R. 1998. A Justification of Societal Altruism According to the Memetic Application of Hamilton's Rule, Abstract of a paper presented at the First Workshop on Memetics.
- GABORA L. 1995. Meme and Variations: A Computer Model of Cultural Evolution. In L. Nadel and D. L. Stein, editors, *1993 Lectures in Complex Systems*, SFI Studies in the Sciences of Complexity, Lectures Volume VI, Addison-Wesley, Redwood City, CA.
- GABORA L. 1996. A Day in the Life of a Meme, *Philosophica*, 57:901-938.
- GABORA L. 1997. The Origin and Evolution of Culture and Creativity, *Journal of Memetics - Evolutionary Models of Information Transmission*, 1.
- GABORA L. 1998a. Memetics. In P. Bouissac, editor, *The Encyclopedia of Semiotics*, Oxford University Press, Oxford.

GABORA L. 1998b. Autocatalytic Closure in a Cognitive System: A Tentative Scenario for the Evolution of Culture, *Psycoloquy*, 9:67.

HEYLIGHEN F. 1992a. Evolution, Selfishness and Cooperation, *Journal of Ideas*, 2:70-76.  
[Download]

HEYLIGHEN F. 1992b. Selfish Memes and the Evolution of Cooperation, *Journal of Ideas*, 2:77-84. [Download]

HULL D. L. 1988a. *Science as a Process: An Evolutionary Account of the Social and Conceptual Development of Science*, University of Chicago Press, Chicago, IL.

HULL D. L. 1988b. Interactors Versus Vehicles. In H. C. Plotkin, editor, *The Role of Behavior in Evolution*, The M.I.T. Press, Cambridge, MA.

LYNCH A. 1996. *Thought Contagion: How Belief Spreads through Society: The New Science of Memes*, Basic Books, New York, NY. [JASSS review]

RADNITZKY G. and W. W. Bartley III 1987. *Evolutionary Epistemology, Rationality, and the Sociology of Knowledge*, Open Court, La Salle, IL.

# **MACROMEMETICS: TOWARDS A FRAMEWORK FOR THE RE-UNIFICATION OF PHILOSOPHY**

**By**

**Derek Gatherer**

School of Biomolecular Science,  
Liverpool John Moores University,  
Byrom St., Liverpool, L3 3AF. UK.

## **Abstract**

The review of the philosophical and scientific antecedents of memetics by Elan Moritz [27] is extended to cover 20<sup>th</sup> century philosophy. It is proposed that '*macromemetics*', i.e., the study of the evolution of entire meme pools, is in many respects a similar enterprise to that attempted by the cultural evolutionist school of anthropology which flourished in the late 19<sup>th</sup> and early 20<sup>th</sup> centuries. Memetics may provide a new, and more rigorously Darwinian, slant to this rather neglected branch of anthropology. Furthermore, it is proposed that much of the terminology and conceptual apparatus of memetics may be reconciled with that used by recent Continental philosophers in their studies of culture, thus suggesting a possible Darwinian framework for the re-unification of Western philosophy.

**Key words:** meme, culture, evolution, Structuralism, philosophy, reductionism, anthropology, civilization.

## **1 Introduction and Summary**

This article reviews the divided state of Western Philosophy, and points out certain areas in which the Continental and Anglo-American schools may be reconciled by the use of a memetic approach. Firstly, the similarities between Popper's hierarchical dissection of his *World 3* concept and the hierarchical nature of the gene/meme analogy are discussed. Special attention is paid to the lowest level, that of fundamental propositions or memetic '*nucleotides*', and the highest level, that of the meme pool. The study of meme pool evolution is compared with the cultural evolutionary school of social anthropology. Additionally, incipient memetical ideas are detected and examined in the work of Peirce, Saussure, Wittgenstein, Toynbee, Foucault and Derrida. The principal conclusion is that wherever philosophy is concerned with informational entities, a memetical approach may be applied. Both the cultural subject matter of Continental philosophy and the logical and linguistic concerns of the Anglo-American tradition can be re-expressed memetically. Memetics thus provides some possible common ground for the reunification of the two traditions.

## **2 Western Philosophy Divided**

Since the time of Immanuel Kant (1724-1804), Western philosophy has been split into two main traditions, namely Continental philosophy which took its lead from J.G. Fichte (1762-1814) and developed via G.W.F. Hegel (1770-1831) and Karl Marx (1818-1883) into 20<sup>th</sup> century

Existentialism and Structuralism, and secondly Anglo-American philosophy which continued the radical empiricist tradition of David Hume (1711-1776) as developed by Bertrand Russell (1872-1970). The labels *Continental* and *Anglo-American* are unfortunate in that they imply some kind of geographical separation of the two traditions. While this is true to a certain degree, many of the principal figures in the Anglo-American tradition have been native German-speakers, such as Ludwig Wittgenstein (1889-1951), Karl Popper (1902-1994) and the various members of the Vienna circle. Conversely, Continental philosophy is now increasingly popular in Britain and the USA, although usually in academic departments of literature and cultural studies rather than those of philosophy (concise reviews of the two traditions are provided by Scruton [40] and Pears & Kenny [29]).

Anglo-American philosophy primarily concerns itself with logic, philosophy of science and linguistic analysis of meaning. Continental philosophy, on the other hand, is interested in politics, aesthetics, culture and '*mind*', all topics generally regarded as too diffuse and intangible for the more logic- and science-oriented Anglo-Americans. In recent years a certain cross-fertilization has taken place, with Wittgenstein's later work being of influence in both main branches [51]. Additionally, with the development of what is known as Cognitive Science, the subject of '*mind*' and consciousness is becoming respectable again in the Anglo-American philosophical world (reviewed by Gardner [17]).

The principal thesis of this essay is that memetics, being applicable to all transmitted information whether scientific or more generally cultural, provides a means for reuniting Western philosophy. Many of the concerns of Continental philosophers, such as the state of society and social change, are approachable using memetics, and much of their terminology and conceptual framework can be re-expressed in memetic terms. This is not so unlikely as one might initially suspect, since the roots of Structuralism are in 19<sup>th</sup> century cultural evolutionism. Before dealing with this proposal in more detail, some consideration will be given to the meme-gene analogy and its relevance to the subject of '*culture*'.

### **3 The Hierarchical Structure of the Meme Pool and Popper's World 3**

Memes are not transmitted independently. A religious education, for instance, imparts an enormous bundle of memes to an individual which are generally delivered all together or not at all. They are, to borrow from genetic terminology, '*linked*'. Indeed many of these memes may be dependent on each other (in what Speel [41] has termed a '*memeplex*'). The process of growing up and living in a certain culture at a certain time means that an individual is very likely to share a vast quantity of memes with other individuals in the same circumstances. This is what may be termed the meme pool of that group or society. Where a society is highly pluralistic, several meme pools may coexist and partially overlap. Unlike higher eukaryotic gene pools [52], meme pools are not closed systems between which absolutely no interaction can take place. The Roman Empire, for instance, played host to a vast plethora of cultures and religions. Much meme flow between diverse meme pools took place; for instance the interaction between Greek and Jewish thought which produced the work of Philo of Alexandria (c. 30 BCE - 40 CE) and had a seminal influence on early Christianity. However, despite this extensive interaction the Jewish and Hellenistic meme pools remained sufficiently different to be clearly recognizable as distinct cultures. Although this cultural diversity was reduced in the Middle Ages, Christianity was

regularly riven with heresies and schisms even before the Reformation. Memetic mutation and recombination, like their genetic analogues, are omnipresent features of meme pools. Nevertheless, the tendency of memes to be transmitted en bloc rather than as an independent assortment, suggests that we are justified in regarding the meme pool as a useful concept. A more difficult question is how to subdivide the meme pool.

Dawkins [9] cites the work of Karl Popper as one of his inspirations for the meme concept. Popper [34, 36] introduced the term '*World 3*' to refer to the objective contents of thought. His discussion of these contents is almost identical to Dawkins' presentation of examples of memes.

Compare Dawkins' original definition:

*'Examples of memes are tunes, ideas, catch-phrases, clothes fashions, ways of making pots or of building arches'*[9],

with Popper's definition of World 3:

*'theoretical systems.... problems and problem situations.... critical arguments..... the contents of journals, books and libraries'* ([34], p. 107).

It is clear from this that Popper and Dawkins have similar concepts in mind. Although Popper's first three specific examples could be grouped under the heading of Dawkins' third example i.e. '*ideas*', his inclusion of printed artifacts such as the contents of libraries demonstrates that Popper, like Dawkins, also wishes to include the physical manifestations of transmitted information in his definition. In other words, both authors fail to make a clear distinction between what Speel [42] terms the memotype and the phemotype. Popper also intends that World 3 should include all objective thought contents both past and present. Concepts evolve in World 3 just as genes evolve in World 1, but World 3 does not itself evolve. Meme pools are collections of memes available to human populations at points in space and time, and they therefore evolve, diversify, go extinct etc. World 3 may thus be regarded as the set of all meme pools, possible and actual. However, the point in this instance is that the contents of meme pools correspond to the contents of World 3, and that both may be hierarchically structured (Popper [35]). This process strengthens the meme/gene analogy and eliminates the possible criticism that memes are difficult to define precisely, since genes are also non-discrete elements arranged hierarchically in the genome.

The lowest and most fundamental level is that of simple propositions. These will be shared by almost all the higher level structures. For instance, some propositions of logic may fall into this category. Higher level memetic structures such as religions or science or political beliefs may be widely different, or even incompatible, but are often based on the same simple propositions. Despite the uneasy relationship between Christianity and science over the last 400 years, a case can be made that both are derivatives of Platonism. They share what Peter van Inwagen [46] has called the Common Western Metaphysic. A further injection of ancient Greek thought into Christianity was provided by Thomas Aquinas (1224-1274) who reintroduced Aristotelian logic. Indian philosophy, by contrast, has its own system of logic known as Nyaya (Phillips [32]), and a metaphysical structure that is quite different to the Western Metaphysic outlined by

van Inwagen [46]. A little more will be said later about the Logical Atomist school of Anglo-American philosophy for whom the simple propositions of Western logic were the main topic of study. However, at present it will suffice, following Popper, to regard fundamental propositions as the primary basis of any higher level memetic structures. In the gene-meme analogy they are the memetic '*nucleotides*'. Most strands of Western thought, whatever their differences at more complex levels, have a common basis at this lowest '*nucleotide*' level.

Popper is less specific about the intermediate levels of his hierarchy but envisages a progressive building upwards to '*complex ideas*'. This would presumably include entities like scientific theories, specific religious doctrines or points of political dogma. These complex memes have many component parts, and may stand alone, but are at their most effective when combined with other complex memes. For example, Einstein's Special Theory of Relativity is a robust meme which may be transmitted alone, but will generally be replicated more accurately (and possibly also more efficiently) if when delivered as part of a general education in physics including other complex memes such as Maxwell's Equations and the Copenhagen Interpretation.

Above the level of the complex meme - which may perhaps be analogous to the gene as a functional, possibly '*selfish*', but not strictly independent unit - we have co-transmitted aggregates of memes. Examples of these might include religions, scientific disciplines or artistic schools considered as wholes. These are what might be termed the '*linkage groups*' of memetics or '*memeplexes*' (Speel [41]).

Finally we arrive at the total meme pool, or World 3. As emphasized above, Popper's choice of the word '*World*' is appropriate in that he wishes to include every single content of objective thought. There is only one World 3. On the other hand, meme pools exist in the plural, being more analogous to the gene pools of biology in that they may be permanently separate, or at most in intermittent communication, and located at varying points in time and space. Since in a biological context, gene pools correspond roughly to species, or at least to sub-species or incipient species, it is here proposed that meme pools are the appropriate species-level entity in a memetic context. This is contrary to Benzon [3] who sees '*paradigms*' as the species-level memetic entity. If the meme pool/species equivalence is maintained, then paradigms are closer to gene-level entities, like Popper's '*complex ideas*' discussed above. Although meme pools are not completely isolated in the way that higher eukaryotic gene pools are, the analogy is still acceptable if one thinks of prokaryotic gene pools where much horizontal transfer of genetic information takes place (Speel [42]).

### **3.1 Meme pools and the total cultural apparatus of societies**

A unanimously acceptable definition of '*culture*' has always been difficult to obtain (see the review by White [47] for a history of the controversy). Two broad strands of opinion may be identified; those that believe that culture is what people think, do and produce, and most importantly see culture as a learned and retransmitted entity (Leach [24] p. 9), and in contrast those who see it as primarily an advanced biological adaptation to the environment. The latter school has recently been strengthened by support from sociobiologists (Wilson [49]). Although this is a major controversy in anthropology, from the point of view of memetics it is perhaps not necessary to reach a final decision in favour of one view over the other. Sociobiologists see

culture as the product of selection on behaviour, but that the primary consequences of that selection are genetic, and only secondarily memetic. For sociobiologists, culture is only likely to be efficiently replicated when the appropriate genetic basis is already there. For instance, a sociobiologist might point out that belief in the undesirability of incest may stem ultimately from the undesirable medical consequences of inbreeding. Those who are genetically predisposed to avoid incest (e.g. through genetically influenced feelings of disgust) may do so highly efficiently, and thus raise less inbred and consequently healthier children to whom they will also pass their incest-avoidance genes. Incest avoidance is an almost (but not quite, e.g. ancient Egypt - Hopkins [19]) universal feature of human societies, which tends to support the opinion that it is an adaptation with an ultimately genetic origin. However, the resulting behaviour of incest avoidance may still be treated as a meme, i.e., one may learn an incest taboo from a person who genetically avoids incest, even though one may not be oneself genetically predisposed to avoid incest. The incest avoidance meme is also subject to selection, and will spread if incest avoidance produces healthier progeny who then learn incest avoidance themselves. The question of learning versus sociobiology thus seems to lose some of its difficulty when viewed from a memetical standpoint.

Conversely, ideas, like genes, that are maladaptive to individuals may still spread within populations. Incest avoidance memes, like incest avoidance genes, increase the fertility of those who carry them, but suicide cults are obviously highly counter-adaptive to the individuals who sacrifice their lives under their influence. However, a suicide cult meme may procure an advantage to the culture in which it occurs, where the incidence of such suicide is tightly culturally controlled. For example, kamikaze fighter pilots were a formidable military weapon for early-modern Japanese culture. On the other hand, the James Jones mass suicide in Guyana was a unique event which eliminated the entire culture in which it occurred. A genetic predisposition to suicide may exist (possibly associated with some genetic form of depressive illness), but suicidal behaviour may also be learned by indoctrination. Thus, both adaptive and maladaptive behaviours may be considered as selectable memes, whether or not there is also a genetic basis. Genetics has also traditionally focused, until very recently, on the individual as the unit of selection. Part of the appeal of memetics is its compatibility with gene-centred models of selection (Dawkins [9], Williams [48]). Focusing on the spread of memes permits explanation of behaviours which are maladaptive to the individual.

In summary, the traditional divide between culture-as-transmitted-information and culture-as-advanced-adaptation may not be so serious, if transmitted information is also subject to selection. Nature versus nurture debates of the past generally relied on the assumption that only nature was subject to Darwinian selection. Nurture, i.e. transmitted information, was generally regarded as being independent of external forces, a pure product of human superstition and/or rationality, or alternatively subject to dialectics of a Hegelian or Marxist variety. Memetics, by positing a strictly Darwinian process for nurture, thus questions the division between environment and biology, since both are evolving under the same dynamic.

Meme pools may exist in isolation for thousands of years, as was the case for many of the indigenous peoples of the globe before the European expansion beginning in the late 15<sup>th</sup> century. As Benzon [3] has illustrated, the isolation of early migratory groups of humans would have resulted in both genetic and cultural divergence, by both genetic/memetic drift and selection

for adaptation to local conditions. As the globe became more populated, isolated meme pools regained contact. When Christopher Columbus stepped ashore at Watling Island in 1492, contact was established between two meme pools whose nearest common ancestor was at least some 12,000 years in the past. New meme pools may result from the contact of previously isolated ones, or one meme pool may virtually swamp and displace another, as was the case in the Americas in the aftermath of Columbus. Much anthropological investigation has been carried out in 'acculturation', i.e. the consequences of culture contact, usually between an indigenous society and a Western colonising power (Beattie [2] p. 241). If memetics is to be of any value for the study of society as a whole, it is important to judge to what extent memetic phenomena may be seen at work at the level of the meme pool. This is not likely to be an easy task: macroevolution is still the most challenging problem of biological evolutionary theory, and is exacerbated by the fact that macroevolutionists are limited to observation rather than experimentation (Stanley [44]). Fortunately for would-be '*macromemeticists*', much study has already been devoted to evolution of cultures as wholes. Unfortunately, much of that work is by now rather old and somewhat out of favour. Nevertheless, a reassessment from a memetic point of view is desirable.

## 4 The Cultural Evolutionary School of Social Anthropology

In 1786, Sir William Jones reported the results of his comparative studies in the Sanskrit, Latin and ancient Greek languages. His conclusion was that all three were descended from a common progenitor (Gardner [17] p. 196). This primitive phylogeny of the Indo-European languages was one of the first strands of evidence for any kind of evolutionary process, and was a factor which profoundly influenced Darwin:

'*We find in distinct languages striking homologies due to community of descent, and analogies due to a similar process of formation*' (Darwin 1871, quoted by Dennett [11]).

The study of cultural evolution was thus technically a predecessor of that of biological evolution, but fell behind in the following century and a half.

Darwin's contemporary, Herbert Spencer (1820-1903), was the first to take an evolutionary view of society as a whole, seeing it as progressing through various stages with Victorian English civilisation rather suspiciously at the pinnacle [43]. Unfortunately, in the hands of his lesser successors, Spencer's ideas deteriorated into a pseudo-scientific justification for imperialist exploitation of '*inferior*' societies (rev. by Dobzhansky [15]). Spencer, like Darwin, had a broad view of evolution as primarily a result of natural selection on naturally occurring variation producing gradual and progressive change. Those anthropologists who adopted the evolutionary theory of culture also took the gradualist line. One of the first of these was Lewis Morgan (1818-1881), who set out an evolutionary theory of society with three stages, namely savagery, barbarism and civilisation, which he posited as corresponding to the evolution of one species from another (Beattie [2] p. 6). Just as Darwin had concerned himself with the origin of species, so were the evolutionary anthropologists concerned with the origin of civilisation.

A remarkable pioneer was Edward Burnett Tylor (1832-1917) who was the first to use the comparative method. Preparing a database of customs, practices and beliefs and subjecting them to statistical analyses, Tylor's '*social arithmetic*' sounds like a premonition of the reductionist

approach of memetics ([25] p. 23). Tylor (1871, quoted by Leach ([24] pp. 38-39)) defines his field of study as follows:

*'....that complex whole which includes knowledge, belief, art, morals, law, custom, and any other capabilities acquired by man as a member of society.....the first step in the study of civilisation is to dissect into details, and classify these into their proper groups.'*

Tylor's use of the word 'acquired' demonstrates that he was interested in all that was transmitted and copied. The 'proper groups' into which he sought to break down this transmitted information include:

*'weapons, textile arts, myths, rites and ceremonies.....laws of marriage and property.....special moral and religious doctrines'* (Leach [24] p. 39).

Tylor is in the same subject area as memetics and approaches it in the same reductionist spirit. He even sketches an idea of memetic linkage in his concept of '*adhesion*', ie. identification of those cultural traits which tends to be co-transmitted, even in different cultures (Gardner [17] p. 228). Tylor's only deficiency is the absence of a neo-Darwinian technique with which to study the evolution of culture, which of course was not available at that time.

By the early 20<sup>th</sup> century the cultural evolutionists had split into two schools, those which saw the development of civilisation as an inevitable progression from a less civilised state which had occurred many times at different locations, and the '*diffusionists*' who put forward the idea of civilisation having originated once in ancient Egypt and radiating out from there (Kuper [23] p. 3). In more neo-Darwinian terms they differed on the question of whether civilisation was polyphyletic or monophyletic. There seems no a priori reason from a memetic point of view to prefer either opinion, but archaeological evidence has resolved the debate against the diffusionists.

The major anthropological theorist at the turn of the century was James Frazer (1854-1941) whose work, *The Golden Bough*, published in several volumes from 1890 onwards, was a grand synthesis searching for regularities or general laws in cultural evolution. Despite the immense arsenal of examples that Frazer brought to bear on the question, the eventual general consensus was that he had failed, and ever since then one influential strand of social anthropology has sought to deny that anything approximating to general laws may be found in the subject (Beattie [2] p. 44). Memetics holds out some fresh hope that a general body of laws capable of explaining social change may be found, and that that body is already with us, namely Neo-Darwinian theory applied to informational transmission.

Cultural evolutionism did not quite die in the aftermath of Frazer, but persisted to a certain extent in the work of Bronislaw Malinowski (1884-1942), Alfred Radcliffe-Brown (1881-1955) and Emile Durkheim (1858-1917). These thinkers tended to see societies as organic wholes, a view often referred to (not always congenially) as '*functionalism*'. Whereas the memetic approach sees societies as meme pools, the functionalist metaphor was more that of a physiological system. Cultural components of societies were seen as the organs of a body rather than as memes under selection (Mair [25] p. 36). One of the adverse consequences of this was the tendency to see

societies as balanced homeostatic entities which played down the process of change ([2] pp. 56-59).

It should not be thought that modern anthropology is totally adverse to the concept of memetics. Anthropologists have devoted a great deal of time to understanding the belief basis of culture, which is ideal material for memetic analysis. For example, much study has been carried out into the potlatch, a custom of the Kwakiutl people of Canada's western seaboard. The potlatch involves the conspicuous consumption, and occasionally destruction, of wealth and the products of labour (Beattie [2] p. 198). Even more extreme examples have been identified by anthropologists. For instance, in 1856 the Xhosa nation of southern Africa, under great pressure from European encroachment, came under the influence of a prophet who ordered them to destroy food stocks and cattle. Many starved to death as a consequence (Beattie [2] p. 263). The anthropological literature also provides interesting examples of other phenomena of memetic relevance. One interesting instance is the rapid spread of the Ghost Dance religion among the Plains Indians in the late 1890s. The Plains Indian culture was itself a fairly recent development which resulted from the westward retreat of the indigenous peoples of the eastern seaboard of the USA from the process of European colonisation. New meme pools were created as previously isolated tribes banded together in new political and social units. Interestingly, the Navajo people who were indigenous to the mid-West and, unlike the Plains tribes, were not a new and transient cultural entity, were completely resistant to the Ghost Dance religion and apparently much amused by it (Beattie [2] p. 261). The Ghost Dance illustrates how societies of diverse memetic origin and subject to intense pressures may evolve memetically in unpredictable and sudden ways. The Navajo, by contrast, were more memetically homogeneous and under less pressure from the European colonists. Another example which was consequent on the contact of previously isolated meme pools with colonial powers was the so-called '*cargo cult*' of Melanesia which spread rapidly throughout the region in the early 20th century.

These examples illustrate how complex memetic analysis may become. Claude Levi-Strauss refers to '*la pensee sauvage*' or '*mytho-logics*', implying that the memetic constitution of a culture may have virtually nothing in common with our own Western meme pool. Exactly what constitutes the basis of our meme pool can be found in the work of Peter van Inwagen [46], who has deduced a '*Common Western Metaphysic*', a set of assumptions agreed upon by all Western thought, whether religious or scientific. The phrase '*mytho-logics*' is used to imply that even the rules of logic, the memetic nucleotides alluded to earlier in this article, may be different in non-Western societies. The logic is mythical. Other anthropologists do not take such an extreme view but nevertheless are prone to make much reference to Wittgenstein's arguments about '*language games*' and '*family resemblances*' (Leach [24], Wittgenstein [51]). Wittgenstein will be dealt with further in the section on philosophy. Levi-Strauss's conclusions, although tremendously influential, are by no means undisputed. Lucien Levy-Bruhl was a forerunner of Levi-Strauss in his idea that primitive societies had fundamentally different systems of reasoning to their Western counterparts, but had abandoned this view by the end of his life (Gardner [17] pp. 223-259).

The fact that modern anthropologists have occasionally thought along lines that have brought them very close to what we would now term memetics is illustrated by White [47]. Writing about a definition of culture in a purely anthropological context, White postulates the unit of culture to

be a '*symbolate*', which is defined as '*a thing or event dependent on symboling*'. Symbolates produce both '*somatic*' culture (ie. beliefs, behaviour, rituals, customs etc.), and '*exosomatic*' culture (i.e. artifacts, buildings, clothing, machines etc.). The symbolate thus corresponds quite well to Dawkins' definition of the meme as including '*ways of making pots or building arches*', as well as beliefs, religion etc. The reason why White's system, like that of his proto-memetical predecessor Tylor mentioned above, does not quite correspond to memetics is that no explicit reference to selection or evolution of symbolates is made. However, in his reference to somatic and exosomatic symbolates, White provides categories which correspond to the phenotype and the memotype (Speel [42]).

#### 4.1 Evolutionary Analysis of Civilisations

Anthropology tends to deal with indigenous peoples who have no written history. Where written records are available, the historian takes over. Some historians have seen patterns in the rise and fall of civilisations that are interpretable in memetic evolutionary terms. The most prominent exponents of this approach have been Oswald Spengler (1880-1936) and Arnold Toynbee (1889-1976). Toynbee's vocabulary is explicitly biological, with words like '*species*', '*genus*' and '*mutation*' regularly utilised, albeit in a somewhat loose manner. If a civilisation can be seen as analogous to a large, sophisticated meme pool, then, like the gene pools of biological species, civilisations may speciate, evolve directionally in response to selection, or go extinct.

Toynbee [45] identified 34 civilisations of which he regarded 15 as being still in existence. It is notable that Toynbee gives a great deal of prominence to ideology as a defining characteristic, thus suggesting strong parallels between his definition of separate civilisations and the concept of the meme pool. The lifetime of civilisations may be short - for instance the Minoan civilisation of Crete only lasted about 6 centuries from 2000 BCE to 1400 BCE - or very long - such as the Mayan civilisation of the Yucatan which may have lasted as long as 43 centuries from about 2500 BCE to its final destruction by the Spanish in 1680 CE. Just as no species is immortal, so it appears that civilisations also have a finite lifetime, after which memetic resources are exhausted and, if no '*speciation*' has taken place, extinction is the result. A more sudden extinction may be the consequence of a genocidal obliteration of meme pools by rival civilisations.

As far as '*speciation*' is concerned, Toynbee regarded some civilisations as derivatives of others. For instance, he posited that the Sumeric civilisation of the Euphrates-Tigris Delta gave rise to the Babylonian civilisation between 1700 and 1500 BCE. The Babylonian civilisation was abruptly terminated in 538 BCE at the destruction of Babylon by the Medes and Persians. However, the meme pool of Babylonian civilisation left a potent memetic residue in the form of astrology, which survives to a certain extent even today (much to the dismay of scientists). An earlier offshoot of the Sumeric civilisation was the Indic civilisation of northern India which survived until the end of the Gupta Empire around 475 CE. Toynbee regarded later Indian civilisation as being sufficiently different from the Indic to constitute a different civilisation which he termed Hindu, and which survives today. There is therefore a limited memetic continuity between the modern Indian meme pool and Sumeric ancestors. The modern Western meme pool developed in Europe in the Dark Ages, but is an offshoot of the Hellenic civilisation which included the Roman Empire and began in Greece and the Aegean around 1300 BCE. The predecessor of the

Hellenic civilisation was the Minoan which originated in Crete around 2000 BCE. The Minoan meme pool is thus the ultimate ancestor of our own.

Speciation and extinction, since they are often definite historical events, are easier to identify than directional evolution under selection. Perhaps the high rate of technological development in Western society since the mid-18th century can be seen as such a directional evolution, where the selective pressure is the economic and political power associated with new technology (Hull [20]). Another example may be the progressive elaboration of an immensely complicated calendar by the Mayan civilisation where the selective pressure was the social power and status associated with accurate ability to predict eclipses and other astronomical events. Notice that the selective pressure posited in these cases is associated with power of one sort or another accruing in the hands of elites (priests, scientists) who are capable of generating novel memes or combinations of memes. This is perhaps the easiest answer but not necessarily the only one. Alternatives are suggested by Benzon [3], and in the work of the 19th century American philosopher, Charles Sanders Peirce, who will be dealt with at greater length below.

Toynbee's system suggests that there may be reasonable grounds for placing meme pools at the species level in the gene-meme analogy. Of course one may have some doubts about his interpretation of various details, and many would dispute his diffusionist tendencies (see above), but the general picture he paints, of lines of descent, extinctions and speciations, is highly compatible with a view that treats meme pools as equivalent to species. However, as briefly mentioned previously, meme pools are rarely totally isolated. Toynbee may see the Minoan meme pool as the ultimate ancestor of the Western, but it is arguable that Middle Eastern memes constitute a strong component of the Western meme pool through the influence of Judaism on Christianity. There could be no biological equivalent of this extensive meme flow from one species-level entity to another (except in a bacterial system with extensive horizontal transmission (Speel [42]).

Pursuing this line of analogy just a little further, we might also speculate that, just as small gene pools are more likely to disappear than large ones, small societies are often the most vulnerable to memetic extinction. What is the necessary size of the meme pool to maintain cultural viability? The Tikopia, for instance, subjects of a classic anthropological study by Raymond Firth in the 1920s (Leach [24] p. 27), were a population of about 1300 individuals on a Pacific island of some three square miles in size. This population had been there for some thousands of years, a meme pool in total isolation but nevertheless viable. Only contact with European missionaries finally destroyed it. However, much smaller meme pools like those of the Pitcairn mutineers or the Californian Yahi isolate (Diamond [13], [14]), numbering only some dozens of founding individuals, are frequently subject to a form of cultural decay rather analogous to genetic load (Muller [28]), which might be termed 'memetic load'.

Benzon [3] presents an interesting scheme for the evolution of civilisations through four '*cognitive ranks*' of paradigmatic sophistication, namely 1 - preliterate, 2 - literate but non-numerate, 3 - numerate/algorithmical/scientific, and finally 4 - computational/data manipulative. Each of these is postulated to increase its memetic complexity in a sigmoidal manner, before a transition to the next rank. This analysis suggests many interesting lines of research, for instance

in how cognitive rank transition may correspond to paradigm shift in science(Kuhn [22]) or to sociological phenomena.

In summary, memetics suggests many new possibilities for the analysis of civilisations. The state of modern Western civilisation, and how it has changed over time, are primary concerns for modern Continental philosophers. Therefore a discussion of how memetics relates to Continental philosophy is relevant at this point.

## 5 Memetics and 20<sup>th</sup> Century Philosophy

Popper's hierarchical concept of World 3 leads to the postulation above of the fundamental propositions of logic as memetic nucleotides. It is important to note that we are not here concerned with what may or may not be true about ancient or modern systems of logic. Logic is learned and therefore subject to the same selective pressures as all other memes.

The basic propositions of logic were very much the concern of Anglo-American philosophers at the turn of the century. Bertrand Russell (1872-1970) and Alfred North Whitehead (1861-1947), having produced their magnum opus on the logical foundations of mathematics, Principia Mathematica, were concerned to extend this solid foundation to all language. This endeavour, Logical Atomism, was designed to provide an ideal language which would be the foundation of all sciences. Their failure to do so resulted in Whitehead's rapid movement into mysticism, while Russell turned to other activities such as history and political campaigning.

Despite the despondency of the Cambridge philosophers, the torch of Logical Atomism was taken up by a young Austrian, Ludwig Wittgenstein. In his main early work, Tractatus Logico-Philosophicus [50], Wittgenstein describes a logical world built from '*atomic facts*', the constituents of simple propositions, which are either true or false. '*The World*' is the totality of true propositions. There are also pseudo-propositions which arise owing to the illogical structure of human language. For memeticists, the '*truth-value*' of a proposition is of less importance than its ability to replicate; both genuine propositions and pseudo-propositions can be considered as memes.

At this point, it should be pointed out that caution must be exercised in co-opting Wittgenstein and the Logical Atomists as forerunners of memetics. Wittgenstein's world of propositions is static and he has little if anything to say about change, let alone evolution, whereas memes are to be regarded as the component parts of an evolving meme pool, much of which, it must be remembered, may be pseudo-propositional. Wittgenstein ([50], Section 6) regards logic, the a priori part of science, mathematics, ethics and philosophy as pseudo-propositional. Interestingly, an early draft of the Tractatus includes the phrase 'a theme in music is a proposition' (Prototracatus 3.16021 quoted by Kenny [21]), but the final draft omits this phrase, confining itself to consideration of what is or is not the case, and excluding the vagaries of the symbolic, ethical or artistic as things which, as Wittgenstein insisted, cannot be said, only shown. Nevertheless, a memeticist may retort that what can be shown can be transmitted, and therefore can evolve. One philosopher who was very much concerned with evolution of the symbolic was the founder of what became the discipline of Semiotics, Charles Sanders Peirce (1839-1914).

## 5.1 Memes and Pragmatism

Peirce [30] was one of the first to reject Hegelian dialectic and assert that consciousness evolves in a Darwinian manner, regarding mental variations as random and negative selection as removing those that are non-functional (tychastic evolution, in his terminology). Peirce's view of thought as consisting of a complex array of atomistic signs and symbols, comes close to the meme concept. His following description of the evolution of symbols bears many similarities to Dawkins' idea of mutating and recombining memes:

*'Symbols grow. They come into being by development out of other signs.....we think only in signs...it is only out of symbols that a new symbol can grow... a symbol, once in being, spreads among the symbols. Such words as force, wealth, law, marriage, bear for us different meanings from those they bore for our barbarous ancestors' (Peirce [31]).*

Peirce's symbols do not have to be linguistic - his theory leaves room for music and other cultural manifestations. The modern discipline of semiology is a development of Peirce's thought. Aside from Peircian semiology's historical interest as a precursor of the meme concept, his theory of the nature of truth, known as Pragmatism, offers a model for the way that memes can evolve. His world of signs has more flexibility, if less order, than that of early Wittgenstein's Logical Atomism. The basic description of a meme, or sign, does not contain any truth function, that is a meme may be an undeniable fact or a highly absurd piece of nonsense or what Wittgenstein [50] would have termed a pseudo-proposition. There is nothing to indicate how 'true' a meme may be until it is tested in the world. The Pragmatic approach to 'truth' by trial and error offers a potential cultural equivalent of Darwinian evolution in genetics.

## 5.2 Popper and Evolutionary Epistemology

Before leaving Anglo-American philosophy, Popper must be mentioned again. His concept of World 3 and its similarity to the meme pool have already been discussed. This is one of the strands of thought that leads to the philosophical school known as Evolutionary Epistemology which developed in parallel to the decline of Logical Atomism. In the inter-war era, even Wittgenstein abandoned philosophy for some years and his immediate followers in the Vienna Circle were broken up by the Nazi invasion of Austria. Wittgenstein's later work emphasised the context dependence of language in the form of 'language games'. In Philosophical Investigations (Wittgenstein [51]), he ruminates on how an analysis of language treated independently of the language game or 'form of life' in which it is used is essentially meaningless. This led to the post-war school known as Linguistic Philosophy led initially by J.L. Austin (1911-1960) [1]. Opposition to this tendency was provided by Popper, although Moritz Schlick (1882-1936) of the Vienna Circle seems to be prefiguring Popper's concept of World 3 in his Allgemeine Erkenntnislehre [31].

The term '*Evolutionary Epistemology*' was coined by D.T. Campbell [5] in a commentary on Popper's work, and has been reviewed recently by Henry Plotkin [33]. A crucial paper is by Leda Cosmides [8], the title of which, 'Has natural selection shaped the way humans reason?' neatly sums up the whole approach, echoing Peirce but in a neo-Darwinian format. Evolutionary Epistemology has become the major trend in epistemology within the Anglo-American world,

thus placing the Darwinian analysis of thought at the very centre of that tradition. In order to begin to reconcile Anglo-American epistemology with Continental post-structuralist thought, it is necessary to examine the Darwinian roots of Structuralism.

### 5.3 Saussure and Signifiers

The founder of 20<sup>th</sup> century Continental philosophy was a Swiss linguist named Ferdinand de Saussure (1857-1913). Virtually unknown in his lifetime, his posthumous influence was immense, due to the publication in 1916 of notes taken at his lectures by two pupils. Many of his concerns were the same as those of the Anglo-American philosophers of his day: the nature of language, meaning and its correspondence with reality etc., but his results were very different from those of his contemporaries Whitehead and Russell. Saussure was originally a comparative linguist working on the evolution of the Indo-European languages, but his later theories tended to emphasise the '*synchronic*' i.e., the existing structure of language at a given point in time, over the '*diachronic*', i.e. the evolution of language over the course of time. Nevertheless, much of his terminology is still compatible with the memetic approach.

Saussure's theory revolves around the notions of the '*signifier*' and the '*signified*'. To use an example provided by Sarup ([37] p. 3), in the case of an apple, the signifier is the sound image made by the word '*apple*', but it is the concept of an apple which is the signified (not, as one might imagine, the apple itself). The '*sign*' in Saussure's terminology is the relationship between the signifier and the signified, and it is arbitrary, depending on convention. A case has already been made for equating the propositions and pseudo-propositions of Logical Atomism with memes, but in this case the correspondence is not so easy to tease out. Is the signified the meme? Or the signifier? Or the sign?

A further difficulty is provided by the fact that Saussure's followers, the Structuralists, like the Logical Atomists, were not particularly interested in change. Structuralism emphasises the study of structural relations existing at one moment in time, ie. the '*synchronic*', over the way that these relations change through time, ie. the '*diachronic*', and thus relegates evolution to a position of lesser importance. As Structuralism has turned into Post-Structuralism, there has been a tendency to concentrate on the signifier rather than the signified, which has been interpreted as an attempt to remove the one-to-one correspondence between propositions and reality. This presents a considerable philosophical challenge (especially for Anglo-Americans), but in effect it brings Structuralism closer to memetics. The potential ambiguity present in Saussure's complex triadic system of signifier, signified and sign is removed. For the Post-Structuralists, the signifier is now the dominant unit and can be considered as analogous to the meme. We thus have 'the play of the signifiers' (*le jeu des signifiants*) much beloved of the school of Post-Structuralism known as the Deconstructionists. The process of breaking a text down into its component signifiers is a similarly reductionist process to memetics. Memeticists analysing a complex belief system are concerned with identifying, dissecting and describing the memes that are present in it, in terms of their replicative powers, adaptiveness, selfishness etc.

The leading Deconstructionist Jacques Derrida has presented the notion that we are made out of language. This seems a strange idea to many scientists and Anglo-American philosophers. However, Daniel Dennett [10] has used the meme concept to say something very similar about

consciousness. Dennett sees memes as a kind of software for the '*virtual machine*' of consciousness which runs on the '*hardware of the brain*'. To say that we (or our consciousnesses) are '*made of*' language, following Derrida, is not too far from Dennett's view that our consciousnesses are '*made*' from the complex interaction of memes.

As mentioned above, the deconstructive process which Derrida and his followers apply to texts is the sort of process that memeticists have to apply to culture in general, in order to reduce it to its basic components, the memes or memetic nucleotides. One example of the lengths to which Deconstruction will go to dissect a text is provided by Derrida [12] who devotes a 139 page book to an examination of the word '*Geist*' in the work of Heidegger. Whereas with the Deconstructionists the process is designed to '*shatter*' meaning (Sarup [37]), in memetics the process is intended to dissect the way that culture has evolved. Deconstruction sees itself as a process of reading texts in a radical new way, different to the traditional way of seeing things. So too is memetics a new approach to culture and science and indeed all human learned behaviour and belief - not merely just a revival of cultural evolutionism but a more truly Darwinian reductionist approach to culture.

#### **5.4 Foucault and the Episteme**

Just as Deconstructionism and detailed memetic analysis display parallels at the lower levels of the memetic hierarchy, similarities can also be seen between the '*macromemetic*' approach of cultural evolutionism and the theory of succeeding '*epistemes*', or total systems of thought, described in the famous structuralist work of the 1960s, *Les Mots et les Choses* [16], by Michel Foucault (1926-1984). Foucault advances an argument for a saltatory view of cultural change, with epistemes replacing each other in periods of revolutionary cultural change. The early cultural evolutionists considered above were gradualists in the spirit of early Darwinism. Now that it is apparent that punctuated equilibria may be a feature of biological evolution (Stanley [44]), macromemetics may also admit revolutionary changes. As Madan Sarup points out [37], there are many similarities between Foucault's notion of the '*normative society*' and Thomas Kuhn's independently developed concept of normal science interrupted by periods of scientific revolution [22]. Kuhn's '*paradigms*' are similar to Foucault's epistemes, except that they apply more narrowly to science rather than to culture as a whole.

One of Foucault's principal concerns was to reject any possibility of a single unifying theory. He thus rejects Marxism. It might be objected that he would have done the same to memetics. However, in defence of our nascent discipline, it must be emphasised that memetics is not really a totalising metanarrative. It is not really a '*grand recit*', like Marxism, where the world is evolving towards a certain inevitable endpoint, for instance in Marxism the endpoint is the achievement of world Communism. Darwinian evolution has no '*guiding hand*' behind it. It may have direction, but that is only in response to consistent selection pressure (Gould [18]). If the selection pressure changes, the direction trend of the response may be diverted or go into reverse. The occasional structuralist accusation that evolution is a '*grand recit*'/metanarrative is therefore unjustified.

## 6. Conclusion: The Role of Memetics

Dawkins [9] acknowledges his indebtedness to Popper, and to Cloak [7], for paving the way to the meme concept. Memetics thus has both philosophical and anthropological precedents. On the philosophical side, Popper's World 3 is populated with many elements which might be described as memes. Cloak comes from the anthropological tradition, and much 19<sup>th</sup> and early 20<sup>th</sup> century anthropology had a strong evolutionary component. Dawkins' real contribution can be seen as the fusion of cultural evolutionism with a reductionist approach derived from the gene selectionist school of evolutionary biology. Cavalli-Sforza and Feldman's '*culturgens*' [6] and perhaps even White's '*symbolants*' [47], were partial anticipations of this approach. Saussure's work on '*signifiers*' stemmed from his early research in the evolution of languages. 20<sup>th</sup> century structuralist philosophy thus has its roots in evolutionary theory. The emphasis that Structuralists place on discontinuous change and an opposition to metanarratives, has led some of the proponents of Structuralism to insist that it is incompatible with Darwinian evolution. However, now that it is appreciated that discontinuous changes e.g. saltations, may occasionally appear in biological evolution, the incompatibility no longer applies. One of Anglo-American philosophy's most flourishing branches is Evolutionary Epistemology, which also derives from Popper's Darwinian approach to the themes of consciousness and the history of ideas. Memetics thus represents a possible framework for the reconciliation of the two main branches of Western philosophy.

## References:

- [1] Austin, J.L. (1979) eds. Urmson, J.O. and Warnock, G.J. *Philosophical Papers* 3rd ed. Oxford University Press, Oxford.
- [2] Beattie, J. (1966) *Other Cultures: Aims, Methods and Achievements in Social Anthropology*. Routledge and Kegan Paul Ltd., London.
- [3] Benzon, W. (1997) Culture as an evolutionary arena. *Journal of Social and Evolutionary Systems*, in press. Also at <http://www.newsavanna.com/wlb/CE/Arena00.shtml>
- [4] Buchler, J. (1955) ed. *Philosophical Writings of Peirce*. Dover Publications Inc., New York.
- [5] Campbell, D.T. (1974) Evolutionary Epistemology. In: Schlippe, A. (ed) *The Philosophy of Karl Popper*. Open Court, La Salle. 413-463.
- [6] Cavalli-Sforza, L.L. (1986) Cultural evolution. *American Zoologist*, 26, 845-855.
- [7] Cloak, F.R. (1975) Is a cultural ethology possible? *Human Ecology*, 3, 161-182.
- [8] Cosmides, L. (1989) The logic of social exchange: Has natural selection shaped how humans reason? Studies with the Wason selection task. *Cognition*, 31, 187-276.
- [9] Dawkins, R. (1976) *The Selfish Gene*. Oxford University Press, Oxford.

- [10] Dennett, D.C. (1991) *Consciousness Explained*. Little Brown, Boston. 138.
- [11] Dennett, D.C. (1996) *Darwin's Dangerous Idea: Evolution and the Meanings of Life*. Penguin Books, London.
- [12] Derrida, J. (1989) *Of Spirit: Heidegger and the Question*. Translated by Bennington, G. and Bowlby, R. The University of Chicago Press, Chicago.
- [13] Diamond, J.M. (1987) Survival in extreme isolation. *Nature*, vol. 325 p. 394.
- [14] Diamond, J.M. (1994) The last people alive. *Nature*, vol. 370 pp. 331-332.
- [15] Dobzhansky, T. (1961) *Mankind Evolving*. Bantam Books Inc., New York.
- [16] Foucault, M. (1973) *The Order of Things: An Archaeology of the Human Sciences*. Vintage Books, New York.
- [17] Gardner, H. (1987) *The Mind's New Science: A History of the Cognitive Revolution*. Basic Books Inc., New York.
- [18] Gould, S.J. (1989) *Wonderful Life: The Burgess Shale and the Nature of History*. Norton, New York.
- [19] Hopkins, K. (1980) Brother-sister marriage in Roman Egypt. *Comparative Studies in Society and History*, 22, 303-354.
- [20] Hull, D.L. (1988) *Science as a Process*. Chicago University Press, Chicago.
- [21] Kenny, A. (1973) *Wittgenstein*. The Penguin Press, London.
- [22] Kuhn, T.S. (1970) *The Structure of Scientific Revolutions*. 2nd ed. The University of Chicago Press, Chicago.
- [23] Kuper, A. (1996) *Anthropology and Anthropologists: The Modern British School*. 3rd ed. Routledge, London.
- [24] Leach, E. (1982) *Social Anthropology*. Fontana Press, Glasgow.
- [25] Mair, L. (1972) *An Introduction to Social Anthropology*. 2nd ed. Oxford University Press, Oxford.
- [26] Montagu, M.A.F. (1962) Culture and the Evolution of Man. Oxford University Press Inc., New York.
- [27] Moritz, E. (1990) Memetic Science: 1 - General Introduction. *Journal of Ideas*, 1, 3-23.  
Also at <http://www.sepa.tudelft.nl/webstaf/hanss/morihp0.htm>

- [28] Muller, H.J. (1950) Our load of mutations. *American Journal of Human Genetics*, 2, 111-176.
- [29] Pears, D. and Kenny, A. (1994) Mill to Wittgenstein. In: Kenny, A. (ed) *The Oxford Illustrated History of Western Philosophy*. Oxford University Press, Oxford.
- [30] Peirce, C.S. (1893a) Evolutionary love. Originally published in *The Monist*, reprinted in Buchler [4] Chapter 27.
- [31] Peirce, C.S. (1893b) Logic as semiotic: the theory of signs. *Collected Papers of Charles Sanders Peirce*, 2 274-302, reprinted in Buchler [4] Chapter 7.
- [32] Phillips, S.H. (1995) South Asian Philosophy. In: Solomon, R.C. and Higgins, K.M. (eds.) *World Philosophy: a Text with Readings*. McGraw-Hill Inc., New York. 65-119.
- [33] Plotkin, H. (1994) *The Nature of Knowledge: Concerning Adaptations, Instinct and the Evolution of Intelligence*. Allen Lane The Penguin Press, London.
- [34] Popper, K.R. (1972) *Objective Knowledge: An Evolutionary Approach*. Clarendon Press, Oxford.
- [35] Popper, K. (1974a) Autobiography of Karl Popper. In: Schlippe [39], 14.
- [36] Popper, K. (1974b) Evolution and World 3. In: Schlippe [39], 1048-1080.
- [37] Sarup, M. (1988) *An Introductory Guide to Post-Structuralism and Postmodernism*. Harvester Wheatsheaf, Hemel Hempstead.
- [38] Schlick, M. (1925) Allgemeine Erkenntnislehre. Springer Verlag, Leipzig. Translated by Brown, G. Section 32.
- [39] Schlippe, A. (1974) ed. *The Philosophy of Karl Popper*. Open Court, La Salle, Illinois.
- [40] Scruton, R. (1994) Continental philosophy from Fichte to Sartre. In: Kenny A (ed) *The Oxford Illustrated History of Western Philosophy*. Oxford University Press, Oxford.
- [41] Speel, H-C. (in press) Memetics: On a conceptual framework for cultural evolution. In: Heylighen, F. and Aerts, D. (eds.) *The Evolution of Complexity*, Kluwer, Dordrecht. Also at
- [42] Speel, H-C. (1997) A short comment from a biologist on William Benzon's essay 'Culture as an evolutionary arena'. *Journal of Social and Evolutionary Systems*, in press. Also at
- [43] Spencer, H. (1868) *Social Statics*. Williams and Norgate, London.
- [44] Stanley, S.M. (1981) *The New Evolutionary Timetable: Fossils, Genes and the Origin of Species*. Basic Books Inc., New York.

- [45] Toynbee, A. (1972) *A Study of History*. Revised and abridged by the author and Caplan, J. Oxford University Press in association with Thames and Hudson, London.
- [46] van Inwagen, (1993) *Metaphysics*. Oxford University Press, Oxford.
- [47] White, L.A. (1959) The concept of culture. *American Anthropologist*, 61, 227-251.  
Reprinted in Montagu (o cit.)
- [48] Williams, G.C. (1966) *Adaptation and Natural Selection*. Princeton University Press, Princeton.
- [49] Wilson, E.O. (1978) *On Human Nature*. Harvard University Press, Cambridge, Mass.
- [50] Wittgenstein, L. (1922) *Tractatus Logico-Philosophicus*. Translated by Ogden and Richards. Routledge and Kegan Paul, London.
- [51] Wittgenstein, L. (1958) *Philosophical Investigations*. Translated by Anscombe, G.E.M. Basil Blackwell, Oxford.
- [52] Wright, S. (1931) Evolution in mendelian populations. *Genetics*, 16, 97-159.

# **WHY THE “THOUGHT CONTAGION” METAPHOR IS RETARDING PROGRESS OF MEMETICS**

**By**

**Derek Gatherer**

School of Biomolecular Sciences  
Liverpool John Moores University  
Byrom St., Liverpool L3 3AF  
[D.Gatherer@livjm.ac.uk](mailto:D.Gatherer@livjm.ac.uk)

## **Abstract**

The most generally accepted definition of the meme, as a ‘unit of information residing in a brain’ (Dawkins 1982), implies a meme-host duality which is the basis of many current developments in memetics, in particular the notion that the passage of such memes (or homoderivative mnemonics, following Lynch 1998) from mind to mind constitutes a process that may be considered as ‘thought contagion’. A critique of religious belief and other non-rational systems of thought, as ‘mind viruses’ (Dawkins 1993), has been built upon such a meme-host duality. This paper provides two objections to the ‘thought contagion’/‘mind virus’ theory: a) that the concept of a transmitted belief, as opposed to transmitted information, is highly problematic, and b) that in any case the concept of a meme-host duality is equally suspect. It is suggested that the least philosophically problematic constitution for a science of memetics would be to adopt a behaviourist stance towards memes, to restrict the use of the term to those replicating cultural phenomena which can be directly observed or measured (Benzon 1996). This would release us from the difficulties of the indefinable meme-host relationship, and also have the merit of making memetics more directly comparable to animal behavioural ecology, to the existing branch of social psychology known as social contagion theory, and to the sociological field of empirical diffusion studies.

**Key words:** meme, belief, religion, linguistic behaviourism, mind virus, thought contagion, social contagion, diffusion of innovations.

## **1 Introduction and Summary**

This article is designed to complement the on-going debate concerning the definition of the term ‘meme’, by investigating some of the limitations consequent on certain usages of the word. Broadly, my intention is to show that what is probably the most popular, even one might say the ‘orthodox’, definition of the meme as a ‘unit of information residing in a brain’ (which I shall term Dawkins B, see Dawkins 1982), presents us with serious philosophical difficulties that may hamper (and indeed are already hampering) the development of memetics as a science. I shall argue that an earlier and broader picture of the meme as a ‘unit of cultural transmission, or a unit of imitation’ (which I shall call Dawkins A, see Dawkins 1976), is in many respects a better working definition, despite the fact that it is now generally regarded as obsolete. The conclusion of this argument is that the best current definition is that of Benzon (1996) although, as with any definition, there is scope for refinement.

I shall focus on two main theses, which develop as a consequence of discussing the difficulties that lurk within Dawkins B:

1. There can be no population memetics.
2. Memetics cannot be used to study why beliefs spread.

In brief, thesis 1 is consequent on the impossibility of quantifying meme frequencies in populations, since this requires a clear meme-host relationship and that is precisely what we cannot obtain in most cases, as will be shown. Thesis 2 is consequent on the impossibility of accurately identifying when belief has been replicated or transmitted. It is possible to accept thesis 2 while still rejecting thesis 1. However, if thesis 1 is accepted, thesis 2 is a necessary consequence.

Some may be horrified at thesis 1. Part of the initial appeal of memetics was that it seemed to suggest that the powerful theoretical tools of population genetics might be applied to revitalise the venerable but rather moribund field of cultural evolution. The present author has also subscribed to that view (Gatherer 1997a, b, c, 1998) so this article is partly a self-criticism. If we can have no *population* memetics, some might feel that there is no point in having *any* memetics. However, memetics may still be scientifically applied in other ways, and may still make a contribution to the humanities. In particular, I shall attempt to argue that memetics may be best constituted as a science by adopting a behaviourist perspective. By behaviourist, I do not mean the rigid 'neo-behaviourist' tradition of Skinner (eg. Skinner 1972), but in the broader sense of Watson (see Cohen 1979). Under the terms of such a 'soft' behaviourism, only observable behaviours and artefacts would be considered, and memetics would cease to concern itself with unobservable mental entities such as beliefs or thoughts. Benzon (1996) has previously come to this conclusion, but by a different route. This critique is particularly directed against those who see memetics as a science of 'thought contagion' (eg. Lynch 1996a), presumably extrapolating from the well-established branch of social psychology known as social contagion theory. This latter dates back to LeBon's 'Psychologie des foules' in 1895, and Baldwin's study, in 1897, of what later became known as the 'Werther Effect' (apparently contagious suicidal behaviour, see Phillips 1974; reviews of social contagion by Hamilton and Hamilton 1981, Levy and Nail 1993). Importantly though, social psychology is largely (although admittedly not exclusively) constituted as a behaviourist and empirical science, and thus avoids many of the philosophical difficulties that 'thought contagion' theory finds itself in.

The first step is to examine why: a) we cannot have a meme-host duality, and why therefore: b) we cannot quantify meme frequencies per capita, and why we therefore: c) cannot have any population memetics.

## **2 Thesis 1: Memes - To Have and Have Not**

### **2.1 Dawkins A and Dawkins B**

I shall refer to Dawkins' original formulation of the meme as Dawkins A (Dawkins 1976). Several years later, Dawkins presented a revised version of the meme which I shall refer to as

Dawkins B (Dawkins 1982). According to Dawkins (1982), the revised version was intended to clarify the relationship between memes and cultural products. Here it will be argued that the Dawkins B definition did not clarify the situation but trapped memetics in a conceptual impasse from which it has not yet emerged.

For reference, the original formulations are as follows:

**Dawkins A:** '*..a unit of cultural transmission, or a unit of imitation.*' (Dawkins 1976, p.206); '*Examples of memes are tunes, ideas, catch-phrases, clothes fashions, ways of making pots or of building arches*' (Dawkins 1976, p.206); '*Popular songs and stiletto heels are examples. Others, such as the Jewish religious laws...*' (Dawkins 1976, p.209); '*Perhaps we could regard an organised church, with its architecture, rituals, laws, music, art and written tradition, as a co-adapted stable set of mutually-assisting memes.*' (Dawkins 1976, p.212); '*Memes for blind faith have their own ruthless ways of propagating themselves.*' (Dawkins 1976, p.213).

**Dawkins B:** (*referring to the original Dawkins A definition, above*) '*..I was insufficiently clear about the distinction between the meme itself, as replicator, on the one hand, and its 'phenotypic effects' or 'meme products' on the other. A meme should be regarded as a unit of information residing in a brain..... It has a definite structure, realized in whatever physical medium the brain uses for storing information....I would want to regard it as physically residing in the brain.*' (Dawkins 1982, p.109); '*The phenotypic effects of a meme may be in the form of words, music, visual images, styles of clothes, facial or hand gestures.....*' (Dawkins 1982, p.109).

The central flaw in the 'thought contagion' and mind virus' hypotheses may be summed up in a single phrase: these theories require individuals to *have* memes. In Dawkins B, the state of having a meme is taken as synonymous with the state of having a certain informational pattern in the brain. But, as will be demonstrated, this is a poor basis for building any theory of cultural evolution. Others have criticised memetics by claiming variously that the meme is a useless abstraction, or that memetics has nothing to contribute to the theory of cultural evolution, or that culture does not evolve (eg. Rose et al. 1984, Barbrook 1996, Gould 1996, Harms 1996, Pinker 1998). I do not support such criticisms, which call the validity of memetics into question. My criticisms, by contrast, are intended to strengthen the theoretical basis of the memetics movement. If we are to have a cultural evolutionary theory, then we need units of replication/selection, and the meme concept would seem to be as good a way as any of approaching this problem. The mistake lies in the frequent assumption that individuals *have* memes. But because, as will be shown, we cannot assign memes *to* individuals with sufficient reliability or regularity, we cannot produce meme frequencies, defined as the proportion of individuals in a population that possess or have a meme. Gene frequencies *are* absolutely necessary to population genetics, and if there is to be a population memetics we correspondingly need to have unambiguous meme frequencies. There are some circumstances in which we can derive a statistic of this sort but, as will be demonstrated, this kind of 'meme frequency' statistic, if we are to have it at all, must be defined in a purely behaviourist manner.

By contrast, individuals do *have* genes. They may or may not pass them on to the next generation of individuals. However, whatever the reproductive success or otherwise of an individual, that individual carries those genes around in its body for the duration of its life. A population

geneticist may derive allele frequencies as the proportion of individuals who carry one gene sequence or another. Gene frequencies are of course abstractions, pure quantitative values, but they are abstractions which relate directly to a physical reality, ie. the reality of gene sequences within the bodies of individuals. Thus, providing the required technical methods are available, one may derive an unambiguous estimation of gene frequencies which are comparable between one population and another.

Memes, on the other hand, are more difficult to pin down. Dawkins A included concepts, cultural artefacts and subjective states such as beliefs, whereas Dawkins B restricted memes to units of neural information giving rise to behaviours or the production of artefacts (Dawkins 1982, p.109). Dawkins is careful to stipulate that when a meme is transmitted from one brain to another, it is not necessary that exact neural configurations are reproduced. Two individuals who are exhibiting the same mental state, entertaining the same idea, performing the same behaviour, are taken to have the same meme, for all intents and purposes - even if their neural configurations are not identical. Thus at this most basic physical level, Dawkins B memes are less tangible than genes. They are merely informational specifications within brains.

## 2.2 Memes in Action

Let us take an example to see how such a Dawkins B framework would apply to a simple system. There is a method of tying a necktie which is called the Windsor knot (named after King Edward VIII). Windsor knots provide a better knot volume and more elegant shape than the standard method of tying a necktie. There is even a mutant form, the half-Windsor, which is slightly simpler and which many people tie in preference to the full Windsor. If we first apply Dawkins A, the knot itself is a cultural replicator, and therefore a meme, as is *also* the idea of tying the knot. *Both* the Windsor knots themselves and the neural activity involved in tying them are taken to be capable of spreading through populations (for instance, the Windsor knot has made a distinct comeback now that ties are a little wider again). By contrast, Dawkins B would restrict the meme to the neural correlates of knot-tying activity, with the knot itself simply becoming the cultural artefact, or 'meme product' (Dawkins 1982, p.109) consequent on the activity of the meme in the mind/brain of the tie-wearer.

To reiterate, in Dawkins B, two people who know how to tie a Windsor knot have the same meme, even if their neural configurations producing this activity are widely different. Likewise, when a Windsor knot is taught, the pupil need not have the same neural configuration as the teacher. The important thing is that both individuals can think of the Windsor knot, and can tie it correctly. Under these circumstances, it is said that meme transmission or replication has occurred.

## 2.3 Two Difficulties

### 2.3.1 Difficulty no.1 - How do we observe a (Dawkins B) meme?

Two difficulties arise as a consequence of this Dawkins B definition. The first of these is the less serious and is acknowledged by Dawkins (1982), although he gives little indication that he thinks it *is* a difficulty. This difficulty is that the flexibility that he allows in neural configuration

corresponding to a Dawkins B meme, the fact that it is *just* a unit of information, presents us with no means of identifying memes separate from their manifestation in the outside world, i.e., separate from their 'meme products'.

Even if one were able to use some method of scanning an individual's brain to discover the internal neural configuration, we would not be able to identify the Dawkins B memes of that individual. As far as neurobiology is currently concerned, Dawkins B memes must be abstractions. This is not a problem as far as mathematical modelling is concerned, nor is it a problem for systems theory-oriented memeticists (eg. de Winter 1984). However, for those with a more empirical bent, those who would like to study memes 'in the wild', it is evident that memes, *even under the rules of Dawkins B*, are only defined by *their manifestations*. Those of us who wish to do empirical memetics are thus left in the position of looking at 'meme products', which *are* admissible as actual memes under the more flexible rules of Dawkins A. Little therefore seems to be achieved by the 'refinement' of Dawkins A to Dawkins B, other than to require that we continually remind ourselves that the cultural artefacts we are studying are not the 'real' memes but that these 'real memes' are unobservable informational entities in some brains somewhere.

This first difficulty, although inconvenient, is not fatal. It produces a certain philosophical awkwardness, but the general schema of Dawkins B would still be functional if we were able to assign the abstract memes unambiguously to individuals, and thus derive meme frequencies per capita of human population. But this is something we *cannot* do, for the following reasons.

### **2.3.2 Difficulty no.2 - How do we quantify a (Dawkins B) meme frequency in an observed population?**

This second difficulty is considerably more serious, since it disqualifies both the population memetics and 'thought contagion' models of cultural change. Both of these *depend* on a meme/carrier-of-meme or meme/host relationship, a necessary consequence of Dawkins B. In the population memetics approach, the meme is regarded as coding for a meme product which may or may not contribute to the fitness of the individual that carries it. In the thought contagion model, the relationship is considered as more than that of virus/host. This is more than just a nicety of semantics, since the decision to adopt one or the other basic perspective can lead to different abstract mathematisations of memetics. For instance, compare Cavalli-Sforza and Feldman (1981) for a population memetics approach where memes (or 'cultural traits' as they term them) are primarily transmitted 'vertically' within families in a manner precisely analogous to the vertical transmission of genes and an iterative generational model is used with an emphasis on fitness and selective advantage; versus Laland (1992), Lynch (1998) and Takahasi (1998) for contagion approaches which emphasise 'horizontal transmission' and the reproductive, rather than selective, efficacy of the meme. Cavalli-Sforza and Feldman (1981) also present models of the contagion variety, where transmission of the 'cultural trait' is horizontal between siblings and unrelated individuals, and the mathematics uses equations from epidemiology rather than population genetics. As Wilkins (1998) points out, both the epidemiological and population genetics methods are derived from the same basic mathematical models of differential reproduction, but the emphases in each case are different.

However, whatever the relative merits of the mathematical systems, in neither case is it possible *in practice* to accurately derive any quantification of meme-host relationships. The example of the Windsor knot will again serve as an illustration. Let us imagine a room of 100 individuals, only one of whom can tie the Windsor knot, and that individual is the only one wearing the Windsor-knotted tie. If we consider the frequency of the Windsor knot meme to be the frequency of the actual occurrence of the knot (taking a Dawkins A-type interpretation), then we have a frequency of  $p=0.01$ . If, on the other hand, we take a more orthodox Dawkins B interpretation, and consider the frequency of the Windsor knot meme to be the frequency of the neurally encoded ability to tie the knot, i.e., the mental informational unit corresponding to the knot, then again we have a frequency of  $p=0.01$ . The two approaches would seem initially to be in agreement.

This situation becomes rapidly more ambiguous when our tie-wearing individual proceeds to tie the knot around the necks of all the other individuals in the room. The frequency of our meme as defined by the occurrence of the knot (Dawkins A) becomes  $p=1$ . However, the frequency of the neurally encoded ability (Dawkins B) remains  $p=0.01$  (unless we allow that some of the individuals learned how to tie the knot while the process was being performed on them, in which case it will increase).

This thought experiment demonstrates that Dawkins A has the advantage of telling us about the change in frequency of a cultural artefact. However, it tells us nothing about any neural events underlying that. We have a change in Dawkins A meme frequency from  $p=0.01$  to  $p=1$ , but that does not reveal the fact that our knot-tying individual is the only one with the knowledge of how to tie the knot. Conversely, under a Dawkins B interpretation, we have no change in meme frequency, but a radical change in the cultural situation of our population. Thus our choice of meme definition places us in two very different situations regarding analysis of the above thought experiment. Furthermore, under the Dawkins B schema we cannot *in any case* identify memes objectively, but only *infer* them from their 'meme products'. If we do not see the single individual tying all the knots, we might incorrectly infer that the other 99 individuals in the room have acquired the Dawkins B meme for knot-tying, and mistakenly place our Dawkins B meme frequency value at  $p=1$ . Again, if one of the other individuals in the room secretly knows how to tie the Windsor knot but is not disposed to divulge this fact, then we actually have a Dawkins B meme frequency of  $p=0.02$ , but we cannot ever know this empirically. Only a selective event (eg. a million dollar prize for Windsor knot tying ability) might suddenly reveal the real Dawkins B meme frequency.

## 2.4 Dawkins A is Better than Dawkins B

To recap, Dawkins A provides us with information about the *cultural state* of our group of 100 individuals, but tells us nothing about what is going on inside their heads. Dawkins B tells us nothing about the change in cultural state of our group, but rather focuses on *unobservable and merely inferred events* going on inside their heads. Since memetics is a theory of cultural evolution, Dawkins A is preferable as *it allows us to look at culture*. Dawkins B diverts us into conjectures about which memes reside in whose heads, regardless of the objective cultural state of our population. This is the principal reason why there has been little or no empirical memetics in the last 22 years. Dawkins B is not a sound basis for any empirical science, as it relegates the

observable, objective manifestations of culture to a subsidiary status. It takes all the data we can gather and dismisses it as mere 'meme product', and leaves us tangled in endless speculation about unobservable mental units of information.

Dawkins A, by treating the objective bits and pieces of culture as memes, frees us from the requirement to chase unobservable entities. It also frees us from the meme-host relationship, since Dawkins A memes, as artefacts and behaviours, need not have hosts. Indeed artefacts definitely do not have hosts. Of course Dawkins A is a broad definition and *also* classifies unobservable mental entities as memes. But Dawkins B turns the unobservable mental entities into the *main objects of study*.

## 2.5 Further Difficulties with the Meme-Host Relationship

Recall the thought experiment involving the room of 100 individuals and the Windsor knot. We saw that in order to identify accurately the true number of individuals who can tie the Windsor knot, ie. the frequency of the hosts of the Windsor-knot-tying meme under the rules of Dawkins B, we might have to produce selective conditions which reveal the full extent of ability to tie the knot. However, even if we can do this, we still cannot say that the individuals capable of tying the knot are *permanent* hosts for any Windsor knot Dawkins B meme, in the way that they may be hosts for any particular gene. All that we *see* is knot tying activity, and a resulting change in the frequency of knots. But knot-tying activity alone (and what other evidence do we have that the individual 'has' the Dawkins B meme?) does not give grounds for postulating any permanent meme-host relationship.

Hewlett and Cavalli-Sforza (1986), in their extensive study of cultural traits in the Aka pygmies, adopt just such a technique of asking who can perform a particular skill (50 skills are considered ranging from hunting techniques, fruit gathering, and sexual behaviour, to songs and dancing). They are therefore able to build up a picture of who is able to do what, and from whom they learned it. The result, however, tells us more about social structure and relations than it does about selection pressures or any changes in the cultural state of the society under consideration (except for the observation that the Aka have recently abandoned hunting with bow-and-arrows for hunting with the more lethal crossbow - but note that in this single instance it is no longer ability but an artefact that is under consideration). Notably, many of the skills are possessed by high percentages of the population. For instance 99% of adult males know all net-hunting techniques, 98% know all food gathering techniques, 100% know all sexual techniques. Among adult females, percentages are equally high. The only distinct partition among the adult population is that females are less knowledgeable concerning hunting (only 59% know all hunting techniques) while males are less knowledgeable about food preparation (only 47% of adult males are any good in the kitchen as compared to 100% of adult females). The homogeneity that arises from Hewlett and Cavalli-Sforza's analysis is precisely a consequence of their decision to examine ability to carry out skilled activities, as opposed to a more behaviourist approach of observing and quantifying performance of skills over time. Referring back to the thought experiment above, they are in the same position as one who attempts to quantify knot-tying *ability*.

A later paper from Cavalli-Sforza's school (Guglielmino et al. 1995) adopts a wider perspective and compares the total cultural pools of different populations. This approach does not require the difficult task of assigning Dawkins B memes (or cultural traits) *to* individuals, since a population is deemed to possess a cultural trait if that trait, usually defined as an artefact or behaviour, exists within it somewhere.

As we shall see later, to postulate that anyone who *can* perform a certain cultural activity, *has the meme for* that activity, leads us into even deeper problems where conceptual memes are considered.

Under Dawkins B, there is no easy way to circumvent this problem. Among the unsatisfactory alternatives, one might postulate *another* meme for 'wearing the Windsor knot', independent of the one for tying it. This might enable us to say that the frequency of the Dawkins B 'knot-tying' meme in our room is  $p=0.01$ , and remains so, but the frequency of the Dawkins B 'knot-wearing' meme goes from  $p=0.01$  to  $p=1$ . But what is the neural process that underlies merely passively wearing the knot? Whereas one might easily concede that performance of a skill involves a specific neural activity, which may be reproduced every time the skill is performed, passive states are less easily coupled to neural processes. A similarly difficult example is the cultural trait for 'acceptance of the Fascist salute' (Cavalli-Sforza and Feldman 1981, p.64). *Tying* the knot, or *giving* the salute, are behaviours which presumably must be underlain by neural activities of some sort, unobservable though they certainly are. But what neural activity underlies simply wearing a certain tie, or accepting a certain salute, or travelling to work by train? One may of course easily quantify numbers of individuals travelling by train, or the number of individuals wearing ties, but these statistics would *not* represent Dawkins B memes. Thus a 'tie-wearing' Dawkins B meme is not a solution to the problem.

## 2.6 How do we Quantify the Frequency of a Conceptual Meme?

With conceptual memes, the problems become even more difficult. The memetics of concepts, specifically religious concepts, is considered by Dawkins (1993), but the major venture into this territory is by Lynch (1998). Lynch uses the term 'mnemon' to indicate an abstract unit of memory.

*'Thus, the principle abstractions manipulated with memetics theory are memory abstractions, or mnemons'* (Lynch 1998, Section 4).

This is not dissimilar to the Dawkins B meme, the 'unit of information residing in a brain', although of course one might insist that Lynch's mnemons refer only to the memory aspect of neural architecture. What is clear is that Lynch considers *awareness of* simple statements, or *belief in* simple statements, to be mnemons, and that the transmission or replication of such mnemons is the basis of the process of 'Thought Contagion' (Lynch 1996a).

Thus Lynch makes statements such as the following Lynch 1998, Section 12):

*'Take the mnemons expressed by the following 3 statements, for instance:*

*Mnemon A: "There is only one true God".*

*Mnemon B: "Christ is Lord".*

*Mnemon C: "Unbelievers are damned".*

*Mnemon D: "Earthly life is better among believers".'*

[There are of course actually four of them, but I quote verbatim. As I have attempted to demonstrate throughout this article, 'Thought Contagion' theory pays scant regard to the accurate quantification of memes].

These statements, or more exactly *awareness of* these statements, are taken by Lynch to be encoded or somehow instantiated in memory. The reader is referred to his article (Lynch 1998) for further examples involving birth control, church attendance etc. Individual concepts are thus taken as mnemonics which exist in the heads of their hosts.

Let us accept this system at face value for the moment, and consider such a concept-mnemon: 'Napoleon died in 1821'. This has been physically replicated in print many times throughout history books. An individual who is aware of this might be said to 'have' this meme/mnemon, under Dawkins B or Lynch. However, this individual is also capable of transmitting a vast variety of other related memes/mnemonics, such as 'Napoleon died in 1820', 'Napoleon died in 1819' and so on. Now does this individual *have* these mnemonics as well? The fact that the first option is factually correct and the other two are false merely perhaps makes it more likely that the first option will be transmitted (perhaps in the selective context of a history exam). However, any individual who has heard of Napoleon *at all* will be *capable* of constructing and transmitting an infinite variety of mnemonics of the form 'Napoleon died in  $x$ '. If we accept that any individual who is *capable of tying* a Windsor knot *has* the Windsor knot meme (and the Dawkins B interpretation *would* seem to imply this, and so would Hewlett and Cavalli-Sforza 1986), then we must accept that anyone who is *capable of transmitting the concept* 'Napoleon died in 1819', *has* that mnemonic: 'awareness of the concept that Napoleon died in 1819'. If we can transmit something, we must be aware of it, unless we are transmitting in an unconscious state.

Similarly, to return to Lynch's example, anyone who has the mnemonic A, above: 'There is only one true God', must also have the mnemonics: 'There are only two true gods', 'There are no true gods' etc, since that individual would be capable of transmitting these mnemonics. Even if that individual does not actually transmit any of the mutant derivatives of the first mnemonic, that person would still be *aware of* the alternatives and thus, according to Lynch's rules, must *have* the mnemonics for those alternatives. Since the alternatives are in principle infinite (ie. 'there are only  $x$  true gods', for  $x = 0$  to infinity), Lynch's system of conceptual mnemonics, taken as *awareness of* any particular statement, inevitably leads us into a picture of the human mind as host to an *infinite* array of potentially transmitted concepts. This is clearly unsatisfactory.

The problem of dealing with conceptual mnemonics/memes as stored statements, or stored awareness of statements, is a further strong stimulus to abandon the strained host-meme correspondence. We have to consider memes as cultural entities which do not *belong to*, or parasitise, any individual. For this reason, Dawkins A is again a preferable definition. Individuals may be involved in the *production* of memes, but those memes/mnemonics are *not* resting latent in

the individual as a gene does. This position will be developed further once we have considered thesis 2.

In summary, since memes (defined by Dawkins B) are not directly observable, we can only study their meme products. Furthermore we cannot assign Dawkins B memes to individuals with any reliability since the relationship of a meme product, whether behaviour, artefact, or concept, to the hypothesised unobservable Dawkins B meme that produces it, is always a difficult one. Consequently, we cannot derive meme frequencies per unit of human population with any reliability. Without such reliability, there can be no population memetics.

### 3 Thesis 2: Lies, Damned Lies and Memetic Transmission

#### 3.1 Beliefs Concerning a Simple Statement

Thesis 2 is that memetics cannot be used to study why beliefs spread. Note that the phrase 'How Belief Spreads Through Society' is the subtitle of the book on memetics by Lynch (1996a). One possible difference that might be posited between the mnemon 'Napoleon died in 1821' (N1821) and other mnemons of the type 'Napoleon died in x', is that N1821 *is believed* and the others are not. Perhaps I *have* the meme N1821 *because I believe it*. The study of belief states has fascinated, perhaps obsessed, memeticists ever since Dawkins, and indeed memetics is best known to the general public for the aggressive critique it has provided of religious belief as a 'mind virus' (Dawkins 1993).

However, there are also serious problems with the attribution of memetic status to belief states. Lynch (1998) hold beliefs in statements to be mnemons *in their own right*. We therefore have 'awareness of the statement that abortion is a mortal sin' as *one* mnemon and 'belief that abortion is a mortal sin' as a *different* mnemon. Similarly (Lynch 1998, Section 9):

'Mnemon P is the belief that "bee pollen invigorates".'

Mnemon Q is awareness of the "bee pollen invigorates" proposition.'

Bearing this in mind, let us further analyse the example of the statement 'Napoleon died in 1820' (N1820). If a teacher deliberately transmits this false statement to a pupil and the pupil believes it, then we cannot say that *belief* has been replicated or transmitted, since the teacher does not believe in this statement. According to Lynch's system, the mnemon 'awareness of N1820' *has* been replicated, but the mnemon 'belief in N1820' *has not*. The same applies to misunderstandings, where the teacher attempts to transmit the correct statement (N1821) but the student ends up believing an incorrect version. Transmission of belief is clearly far more problematic than transmission of information. Where transfer of information alone is concerned and belief is disregarded, the misunderstanding is merely a mutational event, an error of information flow. Likewise, deliberate lying would still be an accurate transmission of information, regardless of the credence given to that information. Despite these difficulties in the transmission of belief, memeticists frequently give the impression that transmission of belief is a straightforward matter of the replication of one person's belief in the mind of another (eg. Dawkins 1993, Lynch 1996a).

No matter how often a teacher repeats a statement, there can be no guarantee that the pupil will believe the statement. Many factors may influence the outcome, but that outcome is not a matter of simple replication of the teacher's beliefs in the mind of the pupil. Belief is not transmissible, but is something that happens *after* the transmission of information. While we *can* often observe the communication of information, we *can never* directly observe transmission of belief.

*Information and belief are not the same kind of thing.* Lynch's system, that has 'awareness of x' and 'belief in x' as two variant forms of *the same kind of thing*, ie. as mnemonics, is based on a highly doubtful ontology.

### 3.2 Belief at the Dawn of Memetics

Lynch's conflation of information with belief was not the first incidence of this error in memetics. One of the examples that Dawkins (1976) gives of a meme is 'belief in life after death', although he actually quotes this from a personal communication by Humphrey (Dawkins 1976, p.206-207). When a meme 'parasitises a brain' (quoting Humphrey again - and here we have one of the earliest uses of the meme-host perspective), it is the transmission of belief that is taken to matter. Religious leaders acquire disciples, then larger sets of followers and eventually a whole church by transmitting the necessary set of memes for their particular religion, secured by other memes such as 'faith' and 'fear of divine retribution' which are co-adapted to ensure that the other memes are not easily cleansed from their hosts. This is the standard memetic view of religion, initiated by Dawkins (1976) and developed further by Dawkins 1993 and Lynch 1996a. What is under consideration in these analyses is the spread of belief.

Or is it? Dawkins' original discussion soon veers away from belief and into more abstract territory, referring to the 'idea of God' (Dawkins 1976, p.207). This is not the same thing as belief, in fact is rather akin to the idea of a flat earth, but Dawkins fails to point out the distinction. Like Lynch, Dawkins appears to see awareness of the concept of God, and belief in the actual existence of God, as similar kinds of thing, just two memes - different memes admittedly, but nevertheless the fact that one is a concept and the other a belief is not seen as a serious problem. Later, in discussing scientific theories, he slides from memes-as-beliefs to meme-as-ideas with a disconcerting rapidity. For instance, we soon have: '....all biologists [who]...believe in Darwin's theory', equated with: '..every individual who understands the theory.....' (Dawkins 1976, p.210). The absurdity of this is illustrated by the fact that I understand everything about fundamentalist Christian theology; raptures, pre-millenarianism, dispensationalism, total depravity, predestination and irresistible grace (yes, I am Scottish, in case you were wondering). However I don't believe a single word of it. By contrast, when it comes to Darwinism, I both understand and believe.

To recap on thesis 2, that memetics cannot be used to study how beliefs spread: an individual may have a set of beliefs, but these cannot be memes, since a) they cannot be transmitted. All that can be transmitted is information. Belief is not itself information, but an attitude towards information. By having statements and belief in statements as mnemonics, Lynch is saying that the attitude towards a mnemonic is itself a mnemonic. Additionally, b) the only empirical evidence we have of belief is through behaviour, whether linguistic behaviour (eg. uttering 'I believe' - a performative utterance in the jargon of linguistic philosophy), or any other cultural manifestation of belief. One may copy the behaviour, but whether this actually leads to belief is another matter

entirely. Therefore beliefs themselves do not qualify as memes/mnemons in any meaningful sense.

## 4 Belief Systems as Complexes of Behaviours

Having stated my two theses: 1, that there can be no population memetics using Dawkins B or the mnemonic system (Lynch 1998), since there can be no reliable calculation of meme or mnemonic frequencies per unit of population, and 2, that memetics cannot be used to study how beliefs spread, since belief is not transmissible but occurs, or fails to occur, independently in the recipient after transmission of information, I shall attempt to put forward an alternative. This is a behaviourist scheme, which treats memes as cultural events, behaviours or artefacts which may be transmitted or copied. Outside the occurrence of the event, the practice of the behaviour, or the lifetime of the artefact, the meme has no existence. The meme does not 'go anywhere' when it is not manifested. It is *not* stored in some neural data bank, some internal meme repository. Lynch (1998, Section 4) would have us believe that the unexpressed meme: 'resides very redundantly in someone's brain', but again this conjures up images of vast memory banks of mnemonics encoding 'awareness of statement x', not a view of the architecture of the brain that is very consonant with current thinking in either neurobiology or linguistics. By contrast, the behaviourist meme *is* its own manifestation. This definition is basically Dawkins A, but with the mentalistic component of that meme definition stripped out. It is a similar definition to that of Benzon (1996).

Let us take the example of a laboratory scientist, an evolutionary ecologist, carrying out a complex set of daily behaviours: fruit flies may have to be fed, progeny counted, genomic DNA prepared, agarose gels loaded, data processed on a computer etc. The belief system behind this may be Darwinism, and the particular set of behaviours may be designed to quantify genetic differences in fly genomes from different habitats. None of the behaviours in the laboratory make much sense without the presupposition of belief in Darwinian theories of evolutionary change. The scientists also say certain typical things to each other; they count allele frequencies, they discuss selective pressures, they debate the significance of results for the theory. These are transmitted behaviours; anyone hoping to become a successful evolutionary fly ecologist will have to copy and master them all. They are memes in their own right. Likewise, a priest of a religion may perform a standard set of ritual activities, say standard things in a religious context, live in a certain way and encourage her flock to do likewise (or perhaps to have lots of children). Novice priests have a lot of memes to copy too. These things presuppose the belief system. Religions, political doctrines and scientific theories may be referred to in terms of the behaviour they produce. This is not quite the same thing as saying that they are defined in terms of this behaviour, but where memetic transmission is concerned, much of what is transmitted is behavioural. In both cases, however, there is still room for individuals who perform the behaviours without believing the underlying theory. There are skeptics and jokers, subversives and cynics, in every profession.

Nevertheless, behaviour, and the artefacts produced by behaviour, such as micropipettes, laboratories and cathedrals, *is* more easily quantified than conceptual abstractions. It is possible to measure the increase in the number of Christian churches per capita in the Roman Empire, to measure the number of scientific articles on the subject of Darwinism per total articles published,

to measure the percentage of Aka hunters carrying crossbows, to measure the percentage of votes cast for political parties defining themselves as socialist etc. *These* are our meme frequency statistics, or as near as we shall ever get to such a thing. *These* memes are behaviours, or artefacts that are the products of behaviour, and *not* abstract informational instantiations in individual brains.

And, crucially, individuals do not *have* any of these memes. They build them, say them, do them, make them, assent to them or deny them, but the memes are entirely outside the human beings that generate them. These meme frequency statistics are *not* per capita of human populations, and therefore *do not constitute a body of data which is formally analogous to that of population genetics*. For this reason, there can be no population memetics.

## 5 Social Contagion Theory and its Relevance to Memetics

In this section I ought to be clear that I am discussing the social contagion school of social psychology, and not the thought contagion wing of the memetics movement.

Social contagion psychologists refer to a wide variety of socially transmitted phenomena. These can include everything from major social problems such as aggression (Wheeler and Caggiula 1966), mass hysteria phenomena (Stahl and Lebedun 1974) and antisocial behaviour (Sigelman and Sigelman 1976), right down to such things as contagious laughter (Freedman and Perlick 1979), jaywalking (Lefkowitz et al. 1955), hyperventilation (Moss and McEvedy 1966), binge eating (Crandall 1988) and coughing (Pennebaker 1980). The reader is referred to Levy and Nail (1993) for further details and a panoramic review of the field.

One theme of contagion research is the modification of epidemiological modelling to the particular phenomenon under consideration (Hamilton and Hamilton 1981). In most cases however, the approach is purely empirical and no mathematization is presented. Additionally, social psychologists tend to have little to say about evolution. Social contagion phenomena are either generally seen as unimportant to the evolution of society as a whole or, perhaps more frequently, the social psychologist is raised in a tradition which rejects the entire notion of social evolution. Thus in 'traditional' sociology and social psychological contagion theory, there is little if any reference to selective pressures, competition etc.

Social contagion theorists have tended to split their field into three areas: disinhibitory, echo and hysterical (Levy and Nail 1993). In disinhibitory contagion, what seems to be involved is not the transmission of any meme so much as contagious release of inhibitions which then serves to permit an aspect of the unconscious to manifest itself. The connections with Freudian psychology are obvious (Freud 1959).

Social contagion theory tends to direct itself towards the immediate, spontaneous and transient, such as crazes, manias, and fads. Much of the discourse of pop memetics is concerned with such things. However, if memetics is to explain cultural evolution (and that is surely its aim) then it needs a theory that can deal with more permanent changes in behaviour. Blumer in the late 1930s (cited by Levy and Nail 1993) attempted to enlarge the domain of social contagion theory to include more meaningful social changes such as financial panics, patriotic hysteria, the dancing mania of the Middle Ages and various witch hunts of more recent times. These 'social

epidemics', as Blumer termed them, can result in more serious consequences, since patriotic hysteria can cause genocide or war, financial panics can wreck economies etc.

It should also be remembered that, for the social psychologist, contagion theory is just one of a variety of ways of explaining the behavioural influence of one individual on another in society. To quote Levy and Nail (1993), social psychologists also have to consider:

*'...conformity, obedience, persuasion, compliance, deindividuation, social norms, contagion, interdependence, leadership, reactance, social facilitation, social inhibition, social loafing, and vicarious learning...,'*

as well as the more obvious factors such as education and indoctrination. Memeticists tend to lump all these factors together into transmission or replication (there may be a subtle difference between these last two but in memetics they are often used interchangeably). However, contagion only represents a small proportion of all memetic events. To claim, as Lynch (1996a, b, 1998) does, that an epidemiology-derived modelling system can be applied to the spread of beliefs through society, is very likely to fall foul of social psychologists, who recognise a more subtle taxonomy of human discourse.

Where social contagion theory enters the evolutionary camp, in the work of Cavalli-Sforza and Feldman (1981), the point is made that innovations tend to spread horizontally whereas long-standing elements of culture tend to be transmitted vertically from elder members of families (and are thus analysed using a population memetics approach). Innovation, especially technological innovation, can be an important fact producing major changes in society (eg. the introduction of motor cars, antibiotics, computers etc.), but once established as part of a culture, the horizontal, social contagion model of transmission is frequently superseded by the vertical, population memetics model.

Crucially, however, Cavalli-Sforza and colleagues almost always tend to study the spread of the empirically observable (Cavalli-Sforza et al. 1982, Cavalli-Sforza 1986, Guglielmino et al. 1995, and to a lesser extent Hewlett and Cavalli-Sforza 1986). The distinction between traditional social contagion theory (whether evolutionary or sociological) and 'thought contagion' theory is that in the latter, mental events are under consideration.

Before leaving the subject of social contagion, consider the phenomenon of mass hysteria, a recent subject of much discussion in the Memetics Mailing List. Participants at last year's most publicised funeral were seen on worldwide television to exhibit waves of weeping, shouting and, much commented on, spontaneous and what one might call 'culturally inappropriate' applause. Are we to suppose that the wave of applause represented the passage of mnemons from head to head in a rapid wave of mnemonic propagation? And that these mnemons were then immediately translated into the meme product of actual applause? What is happening in a person's head when he hears others applauding and impulsively joins in? What is a 'clapping' mnemonic anyway? I know what clapping is, I have done it often enough. I do not belong to a culture in which there is no clapping (*then* it might be conceivable to suggest that I have no clapping mnemonic). Since I must 'have' a clapping mnemonic, what mnemonic do thought contagionists propose is spreading

through a crowd who all already know what clapping is? Is it the 'clap now' mnemonic perhaps? How does the 'clap now' mnemonic differ from the simple 'clapping' mnemonic?

I cannot answer these questions, but if thought contagionists claim that their theory can explain phenomena such as hysterical contagious applause, then they must be able to provide cogent answers to them.

In any case, a more parsimonious approach would surely be simply to adopt a behaviourist stance and just record who is clapping, where, and for how long; no mnemonics, no abstractions of memory instantiation, no homogenic instantiation events, *just behaviour*.

## 6 Transmission of Innovations

Another related area which has recently been the subject of much discussion in memetic circles, is that of the theory of diffusion of innovations, principally the work of Rogers and his school. Rogers and Shoemaker (1971) trace its origins back to the work of Tarde in 1903, almost contemporaneous with the origins of social contagion theory.

This field however, is mostly about the production of artefacts and/or the adoption of cultural practices, and therefore cannot be easily reconciled with Dawkins B, or the mnemonic approach of Lynch (1998). Typical questions for cultural diffusion studies would be: what is the rate of growth of in the proportion of houses with satellite television and their geographical distribution?, or, what is the growth in the use of catalytic converters and unleaded petrol in motor vehicles?, and so on. Diffusion theory makes no speculation concerning mnemonics or other internal elements in human minds.

For example, in the bibliography of 337 classic diffusion theory-related publications, produced by Brown (undated, probably 1964), only four are reconcilable with a Dawkins B-type memetic approach. Only one of these concerns itself directly with religion: Hawley's study of the uneven acceptance of missionary Catholicism in the southwestern USA (Hawley 1946), although two theoretical papers by Rashevsky deal with beliefs, prejudices (Rashevsky 1951) and ideologies (Rashevsky 1952). The fourth paper is a study of the spread of rumours (Rapaport and Rebhun 1952). The last of these subjects is probably equally compatible with a behaviourist approach, since one may simply record who transmits the rumour (verbal behaviour) and who has already been the recipient of an attempted transmission, without the requirement for any mnemonics of 'having rumour', 'believing rumour' etc. Among the more typical subjects for diffusionists are vaccine use, agricultural practices, medical practices, adoption of fluoridation, spread of cowries as a currency, arithmetical techniques and horse-riding (Brown, undated, probably 1964).

It is significant that neither religion nor belief occur in the index of Rogers and Shoemaker's compendious review of the diffusion research field (Rogers and Shoemaker 1971). These authors list over 1500 articles on diffusion research, but are only able to add one more paper on the diffusion of beliefs, that of Erasmus (1952). Like the social contagion psychologists, diffusionists are overwhelmingly empirical investigators. Their material is the observable and the quantifiable, not the abstract and internal.

## 7 Who Uses Which Definition?

Dawkins B has almost become the orthodox definition of the meme. The principal point of this paper is that Dawkins B represents a retrograde step from Dawkins A, and that it has been responsible for the spawning of all manner of 'thought contagion' speculation which is an active hindrance to empirical work in the subject, and a source of ridicule in the scientific community at large (as those of us who are academics can verify).

Those cultural evolutionists who do not describe themselves as memeticists nevertheless often use concepts that are similar to those of Dawkins. For instance, Lumsden and Wilson's definition of the 'culturgen' as: 'the node of semantic memory' (Lumsden and Wilson 1985, p.348) seems to be very close to Lynch's (1998) mnemon, although importantly, unlike Lynch, they do not construct semantic memory as a stack of 'awareness of x' modules. Richerson and Boyd (1978) are less specific concerning their 'culture-types', which is a unit of selection (not necessarily of replication) in a cultural system:

*'...natural selection should act on culture-type, increasing the frequency of those items of culturally coded information in a population which.....produce phenotypes that are more successful in passing the culture-type to the next generation' (Richerson and Boyd 1978, p.132)*

If, by 'culturally coded information', Richerson and Boyd (1978) imply neural informational units, then this is clearly like Dawkins B. However it is not entirely clear if this is what is implied, since: '..the precise mechanism of inheritance of the code is likely to be largely irrelevant' (Richerson and Boyd 1978, p.132). Whether or not a Dawkins B-type model is implied, Richerson and Boyd do imply a meme-host duality in that they refer to: 'cultural codes as they are transmitted from one individual to another' (Richerson and Boyd 1978, p.132), and their mathematical treatment follows from this.

Cavalli-Sforza and Feldman (1981), as discussed above, tend generally towards a Dawkins A definition in theoretical discussion of their 'cultural traits' (also see Cavalli-Sforza and Feldman 1983, Feldman and Cavalli-Sforza 1984, Feldman et al. 1985), and in practice tend to be more behaviourist, if somewhat irregularly (Cavalli-Sforza et al. 1982, Hewlett and Cavalli-Sforza 1986, Cavalli-Sforza 1986, Guglielmino et al. 1995).

Those who use more orthodox Dawkins B interpretations or close relatives, include de Winter (1984), Dennett (1990, 1991) and Speel (1997), and it is the source for the computer modelling of a simple society, carried out by Doran (1997). Ball (1984) begins with a resolutely Dawkins B interpretation, but then takes it further in claiming that *all* mental contents are memes, including those aspects of information which are products of Skinnerian conditioning. Memes thus constituted are no longer units of imitation but units of mental content. A similar view is expressed by Gabora 1997, who sees memes as the units of mental experience, down to and including sensory qualia such as 'a vivid impression of red'. Preti and Miotto (1997) also define memes as simply 'mental representations'.

Benzon (1996) places memes in the external environment as cultural replicators. Briefly, his central point is that it is artefacts that are copied, and therefore it is artefacts that replicate (albeit

somewhat haphazardly). Like Benzon, I do not deny that psychological traits exist - neither of us is a behaviourist in the sense of Skinner or Ryle - but that they are too intangible to be copied in any meaningful sense of the word. The appeal for a 'soft' behaviourist memetics here, and in Benzon (1996), constitutes a restriction of the Dawkins A definition to the artefactual, and the elimination of those aspects of Dawkins A which are mentalistic.

Thus reconstituted, human memetics might have more in common with the thriving animal behaviourist wing of the subject. For instance Ficken and Popp (1995), when measuring the evolution of the 'gargle vocalisation' of the black-capped chickadee, produce statistics of how many songs of certain types have been produced in a certain period. There is no attempt to assign song memes to individual birds, since individuals have a large repertoire of songs. Similarly, Guglielmino et al. (1995) seek to define which memes (or cultural traits) belong to which cultures, and to work through the evolutionary implications of meme distribution at the societal level. Here we have the meme as artefact/behaviour studied in its own right, without the confusing mentalistic stance that currently plagues human memetics.

## **8 Criticisms Anticipated**

Some may reply to the above by saying:

### **8.1 That Memes are Only Abstractions**

One might argue that, since memes are abstractions, we need not worry overly about their ontological status (de Winter 1984). All one needs are tokens for mathematical analysis. This is unsatisfactory, as mathematical modelling is only of value if the model is a reasonable representation of reality. Outside of quantum physics, scientists are generally constrained by the requirement that their models should mirror nature. Even in quantum theory, mathematicians were only driven to the counter-intuitive by the overwhelming weight of experimental evidence. Memetics should be firmly grounded in reality.

### **8.2 That Memetics is Still in its 'Mendelian' Stage, and We Will One Day be Able to Identify Memes Directly**

Another similar criticism might be that before the discovery of DNA, geneticists had to be content with a highly abstract definition of the gene. Prior to the molecular era, genes were only identifiable in terms of phenotypes. Thus a fly geneticist could only identify the white (*w*) mutant allele, indeed could only infer its very existence, by the observation of white-eyed flies. Likewise, one might argue that the (Dawkins B) Windsor knot meme is only identifiable by observation of Windsor knots in this present era of 'Mendelian' memetics, until we have sufficiently advanced in neurobiology to identify the true, internal, mental, neural configurational, meme, in itself without recourse to the external manifestation - just as we can now identify the white (*w*) mutant allele by DNA sequencing without looking at the flies' eyes at all. This argument fails however, for the following reason.

Recall the thought experiment with 100 individuals and the Windsor knot. The individual with the knowledge of how to tie the knot was able to produce 100 knots while there was still only

one meme (Dawkins B definition) for the Windsor knot. By contrast, in fly genetics, each white eyed fly corresponds to two white (*w*) mutant gene sequences (the allele is autosomal recessive - FlyBase 1998). There is no ambiguity; 100 white-eyed flies equals 200 white (*w*) mutant alleles. But in memetics, we cannot be sure how many Windsor knot memes (Dawkins B definition) were responsible for 100 Windsor knots. Thus the situation in memetics at present is not analogous to the Mendelian era of classic genetics. It is true to say that genes were merely abstractions in those days, but they were a very different kind of abstraction to the meme.

### **8.3 That All the Above is Inadmissible as it Constitutes an Attack on Memetics**

The above two theses, that there can be no population memetics (except in a strictly behaviourist sense) or thought contagion theory, and that memetics cannot be used to study belief states, are bound to cause some consternation within the memetics community. Many may take them as an attack on memetics as a discipline. This is *not* my intention. Rather I see myself as strengthening memetics by preventing its inflation beyond the boundaries of what is acceptable for a scientific field. Population memetics proper (ie. those authors who treat memes as culturally inherited traits) is mostly conducted in a behaviourist manner. Its best empirical results have been achieved when investigators cease attempting to assign memes to individuals and treat them as behaviours occurring within a cultural milieu (Guglielmino et al. 1995).

### **8.4 That Memeticists Don't Say that People Have Memes Anyway, We Say that Memes Have People**

Lynch (1996b) inverts the meme-host relation to give memes which have people, rather than the converse.

*'In the area of population psychology and psychohistory, memetics achieves just such a paradigm shift by inverting an everyday question. Instead of asking how people acquire ideas, the new paradigm asks how ideas acquire people.'*

This, however, does nothing to help. When does the meme 'Napoleon died in 1822' have me? Only when I am transmitting it? Or all the time, given that I transmitted it on one occasion and may or may not transmit it in the future?

### **8.5 That Memes are Only the Homoderivative Subset of Mnemon Replication Events**

Although Lynch (1998) has mnemons of the type 'awareness of *x*' or 'belief in *x*', not all of these mnemons are memes in his system. Only homoderivative mnemons are memes. By homoderivative, Lynch means that they are produced by replication of the same mnemon in another individual. In Lynch's own words, it is: 'A memory item, or portion of an organism's neurally-stored information, identified using the abstraction system of the observer, whose instantiation depends critically on causation by prior instantiation of the same memory item in one or more organism's nervous systems.' By contrast, heteroderivative mnemons are somehow produced in response to an attempted but unsuccessful transmission of a different mnemon, or as I think Lynch might say, whose instantiation is not a consequence of any such causation. Lynch is therefore postulating that only a subset of Dawkins B memes should be considered as true

memes. But this is not a fixed subset, as mnemons can be *simultaneously* memes *and* not memes. For instance, in Lynch's system, whenever I utter the phrase 'Napoleon died in 1821' (N1821) to my student, the following events occur.

Two mnemons are involved. The first is the mnemonic N1821, as a basic unit of information, and the second is 'belief in N1821'. Our mnemonic, N1821, which I believe and therefore, Lynch would say, constitutes a mnemonic 'belief in N1821', is a meme *only* if my student believes it, since only then does a homoderivative replication event take place. If I have a lecture theatre full of students of varying degrees of scepticism and gullibility, and I utter N1821 (which *I* believe), it is simultaneously a meme *and* not a meme depending on the belief reactions of the students. However, its alter ego, the simple mnemonic 'awareness of the concept N1821' *is* always a meme, provided the students comprehend it, regardless of whether or not any of them believe it. Aside from the awkwardness of N1821 being simultaneously a meme and not a meme, the problem for any empirical memeticist is that we cannot observe if belief has been replicated or not, so we cannot in any case decide in Lynch's system if the mnemonic 'belief in N1821' is a meme or not. Besides, Lynch's system still presents the same problems of defining when and how a mnemonic is instantiated in its host. Since mnemons can only be inferred on the basis of behaviour or artefacts, adoption of these complicated neologisms does not clarify any of the issues raised here (and surely the last thing memetics needs is yet more neologisms). Lynch also presents 'a symbolic calculus of mnemonic conjugations and replication events', but this is not usable in the elucidation of any of the problems here presented, since again the invalid assumptions are made at the outset that individuals *have* mnemons, and that belief states are transmissible mnemons.

## **8.6 That I Contradict Statements Made In My Earlier Work**

This is true, so I ought to be specific about how much backtracking I wish to do. In Gatherer (1998), p.209-210), I state:

*'Rather than mind viruses, religious memes, like scientific ones, may be considered as mind symbionts, replicating themselves through a positive contribution to the well-being of (most of) those who carry them.'*

I am now considerably less comfortable with the idea of individuals *carrying* any memes. Religious beliefs may be beneficial to (most of) those who profess them, but this is a psychological phenomenon rather than a memetic one. The whole mind-virus/symbiont controversy is based on an incorrect meme-host duality, as described in the present article. I still stand by my view of religions as 'large, integrated complexes of memes' (Gatherer 1998, p.205), but I should now say that beliefs are not among those memes. The memes in question are the observable, copiable aspects of religion, such as prayers, rituals, artefacts, liturgies, dogmas etc. In an earlier article (Gatherer 1997b), I discuss the merits of memetic diversity, and the problems arising from its global depletion. At the time I intended beliefs to be considered as part of such memetic diversity. However, this argument loses no force even when belief is left out of the issue. In fact, that much was already implicit in the original. We may preserve the memes of indigenous cultures, their practices, artefacts and general way of life, but we cannot preserve

their beliefs. I can dress like an Amazonian shaman, dance like him and perform his rituals, but I *cannot*, try as I might, *believe* the same things as he does; belief is simply not transmissible. By the time I came to write Gatherer (1997c), I was already effectively considering memes as observable artefacts (music, in the context of that paper). I had not then consciously come to reject a mentalistic notion of memes, but perhaps the rigours of considering a real cultural system were beginning to push me in that direction.

## 9 And Finally....Yet Another Meme Definition

One who spends so much effort criticising the meme definitions of others, ought to present a clear target for his own critics. The following is therefore offered:

*Meme: an observable cultural phenomenon, such as a behaviour, artefact or an objective piece of information, which is copied, imitated or learned, and thus may replicate within a cultural system. Objective information includes instructions, norms, rules, institutions and social practices provided they are observable.*

That will suffice for a bald definition. By way of supplementary clarification, one might say that this definition excludes behaviours which are genetically determined (which are left to the sociobiologists), although a behaviour which is genetically determined in one individual may be copied by another individual who is not so determined, and thus become a culturally propagated meme (eg. incest avoidance, see Durham 1991, Chapter 6). It also excludes those aspects of culture which are not observable, and not readily transmissible, such as beliefs (which are left to the cognitive psychologists). It also excludes behaviours which are learned via trial-and-error or Skinnerian reinforcement although, as above, once such behaviours are observable, they can be copied and thereby become memes. Individuals may be, and usually are, intimately involved in the production, transmission and copying of memes, but these memes cannot be said to belong to such individuals, or to parasitise such individuals.

I do not claim any originality for this definition, because it is essentially identical to that previously given by Benzon (1996):

*'.... I suggest that we regard the whole of physical culture as .... [memes]: the pots and knives, the looms and cured hides, the utterances and written words, the ploughshares and transistors, the songs and painted images, the tents and stone fortifications, the dances and sculpted figures, all of it. For these are the things which people exchange with one another, through which they interact with one another. They can be counted and classified and variously studied.'*  
(Benzon 1996, p.323)

One might ask: what difference would such a memetics have to already existing fields such as diffusion studies or social contagion theory? The answer would be: very little, except perhaps that a Darwinian streak would be more prominent; it would be an empirically-based evolutionary contagion/diffusion theory, which might eventually tell us something about the evolution of culture as a whole.

But this would be no bad thing. Memetics as it currently stands is a strange beast, an unrealistic psycho-epidemiology, laden with neologisms, neither theory of mind nor theory of culture. To make it into a science, we need to move forward into territory already occupied by diffusion sociologists and social psychologists. *We are not pioneers* (despite the 'new science of memetics' rhetoric that occasionally surfaces), but just *a new wave of settlers*. To survive in our new environment we need to learn from those who have been living there at least since the days of Gabriel Tarde at the turn of the century. The first step must be to shed the unnecessary baggage of 'thought contagion', 'memory abstractions' and meme-host duality. This is not just necessary in order to retain the goodwill and respect of our new neighbours, *but for the sake of intellectual survival*.

## Acknowledgements

I thank the members of the Bristol Meme Lab for kindly inviting me to visit them, and for many helpful comments. Similar thanks go to Stephen Bentley and his evolutionary discussion group at University College London. My colleagues Nick Owen (Liverpool Institute of the Performing Arts) and Holly Smith (School of Human Sciences, Liverpool JMU) have also suffered having my ideas bounced off them. None of these people, of course, necessarily agree with any of the above.

## References

- Ball, J.A. (1985) Memes as replicators. *Ethology and Sociobiology* **5**, 145-161.
- Barbrook, R. (1996) Never mind the cyberbollocks... A contribution to *Memesis* <http://web.aec.at/meme/symp/panel/msg00076.html>
- Benzon, W. (1996) Culture as an evolutionary arena. *Journal of Social and Evolutionary Systems* **19**, 321-362. <http://www.newsavanna.com/wlb/CE/Arena/Arena00.shtml>
- Brown, L. (undated, probably 1964) *A Bibliography on Spatial Diffusion - with Special Emphasis on Methodology and Theory*. Department of Geography, Northwestern University, Evanston, Illinois.
- Cavalli-Sforza, L.L. and Feldman, M.W. (1981) *Cultural Transmission and Evolution: A Quantitative Approach*. Princeton University Press: Princeton.
- Cavalli-Sforza, L.L., Feldman, M.W., Chen, K.H. and Dornbusch, S.M. (1982) Theory and observation in cultural transmission. *Science* **218**, 19-27.
- Cavalli-Sforza, L.L. and Feldman, M.W. (1983) Cultural versus genetic adaptation. *Proc. Natl. Acad. Sci. USA*, **80**, 4993-4996.
- Cavalli-Sforza, L.L. (1986) Cultural evolution. *American Zoologist* **26**, 845-855.
- Cohen, D. (1979) *J.B. Watson: the Founder of Behaviourism: a Biography*. Routledge, London.

- Crandall, C.S. (1988) Social contagion of binge eating. *Journal of Personality and Social Psychology* **55**, 588-598.
- Dawkins, R. (1976) *The Selfish Gene*. Paperback edition 1978. Granada Paladin: London.
- Dawkins, R. (1982) *The Extended Phenotype*. Paperback edition 1983. Oxford University Press: Oxford.
- Dawkins, R. (1993) Viruses of the mind. In: *Dennett and his Critics*. Dahlbom, B. (Ed.) Blackwell: Oxford. pp.13-27.
- Dennett, D.C. (1990) Memes and the exploitation of the imagination. *Journal of Aesthetics and Art Criticism* **48**, 127-135. <http://www.tufts.edu/as/cogstud/papers/memeimag.htm>
- Dennett, D.C. (1991) *Consciousness Explained*. Little Brown: Boston.
- Doran, J. (1998) Simulating collective misbelief. *Journal of Artificial Societies and Social Simulation* **1**, no. 1. <http://www.soc.surrey.ac.uk/JASSS/1/1/3.html>
- Durham, W.H. (1991) *Coevolution: Genes, Culture and Human Diversity*. Stanford University Press: Stanford.
- Erasmus, C.J. (1952) Changing folk beliefs and the relativity of empirical knowledge. *Southwestern Journal of Anthropology* **8**, 411-428.
- Feldman, M.W. and Cavalli-Sforza, L.L. (1984) Cultural and biological evolutionary processes - gene culture disequilibrium. *Proc. Natl. Acad. Sci. USA*, **81**, 1604-1607.
- Feldman, M.W., Cavalli-Sforza, L.L. and Peck, J.R. (1985) Gene culture coevolution - models for the evolution of altruism with cultural transmission. *Proc. Natl. Acad. Sci. USA*, **82**, 5814-5818.
- Ficken, M.S. and Popp, J.W. (1995) Long-term persistence of a culturally transmitted vocalisation of the black-capped chickadee. *Animal Behaviour* **50**, 683-693.
- FlyBase (1998) FlyBase Gene Report. FlyBase: a Database of the Drosophila Genome. Department of Biology, Indiana University, Jordan Hall, Bloomington, IN 47405. [http://flybase.bio.indiana.edu:82/exec+-vtext/html\\_-dR58262235-59314610-.data/genes/genes.txt:/](http://flybase.bio.indiana.edu:82/exec+-vtext/html_-dR58262235-59314610-.data/genes/genes.txt:/)
- Freedman, J.L. and Perlick, D. (1979) Crowding, contagion, and laughter. *Journal of Experimental Social Psychology* **15**, 295-303.
- Freud, S. (1959) *Group Psychology and the Analysis of the Ego*. Trans. and ed. J. Strachey. Norton: London.

Gabora, L. M. (1997) The origin and evolution of culture and creativity *Journal of Memetics - Evolutionary Models of Information Transmission* **1**. [http://jom-emit.cfpm.org/1997/vol1/gabora\\_1.html](http://jom-emit.cfpm.org/1997/vol1/gabora_1.html)

Gatherer, D. (1997a) Macromemetics: towards a framework for the re-unification of philosophy. *Journal of Memetics - Evolutionary Models of Information Transmission* **1**. [http://jom-emit.cfpm.org/1997/vol1/gatherer\\_dg.html](http://jom-emit.cfpm.org/1997/vol1/gatherer_dg.html)

Gatherer, D. (1997b) Feyerabend, Dawkins and the politics of cultural diversity. *Anarchist Studies* **5**, 23-43.

Gatherer, D. (1997c) The evolution of music: a comparison of Darwinian and dialectical methods. *Journal of Social and Evolutionary Systems* **20**, 75-92.

Gatherer, D. (1998) Meme pools, World 3 and Averroës's vision of immortality. *Zygon: Journal of Religion and Science* **33**, 203-219.

Gould, S.J. (1996) *Life's Grandeur: the Spread of Excellence from Plato to Darwin*. Cape, London.

Guglielmino, C.R., Viganotti, C., Hewlett, B. & Cavalli-Sforza, L.L. (1995) Cultural variation in Africa - role of mechanisms of transmission and adaptation. *Proc. Natl. Acad. Sci. USA*, **92**, 7585-7589.

Hamilton, J.D. and Hamilton, L.C. (1981) Models of social contagion. *Journal of Mathematical Sociology* **8**, 133-160.

Harms, W. (1996) Cultural evolution and the variable phenotype. *Biology and Philosophy* **11**, 357-375.

Hawley, F. (1946) The role of Pueblo social organization and the dissemination of Catholicism. *Applied Anthropology* **48**, 407-415.

Hewlett, B.S. and Cavalli-Sforza, L.L. (1986) Cultural transmission among Aka pygmies. *American Anthropologist* **88**, 922-934.

Laland, K.N. (1992) A theoretical investigation of the role of social transmission in evolution. *Ethology and Sociobiology* **13**, 87-113.

Lefkowitz, M., Blake, R.R. and Mouton, J.S. (1955) Status factors in pedestrian violation of traffic signals. *Journal of Abnormal and Social Psychology* **51**, 704-706.

Levy, D.A. and Nail, P.R. (1993) Contagion: A theoretical and empirical review and reconceptualization. *Genetic, Social and General Psychology Monographs* **119**, 235-285.

Lumsden, C.J. and Wilson, E.O. (1985) The relation between biological and cultural evolution. *Journal of Social and Biological Structures*. **8**, 343-359.

Lynch, A. (1996a) *Thought Contagion: How Belief Spreads through Society: The New Science of Memes*. Basic Books, New York.

Lynch, A. (1996b) Contribution to 'Memes: Self-replicants or Mysticism.' *Wired* debate October 1996. <http://www.wired.com/braintennis/96/43/bt0a.html>

Lynch, A. (1998) Units, events and dynamics in memetic evolution. *Journal of Memetics - Evolutionary Models of Information Transmission* **2** [http://jom-emit.cfpm.org/1998/vol2/lynch\\_a.html](http://jom-emit.cfpm.org/1998/vol2/lynch_a.html)

Moss, P.D. and McEvedy, C.P. (1966) An epidemic of overbreathing among schoolgirls. *British Medical Journal* **1966/2**, 1295-1300.

Pennebaker, J.W. (1980) Perceptual and environmental determinants of coughing. *Basic and Applied Social Psychology* **1**, 83-91.

Phillips, D.P. (1974) The influence of suggestion on suicide: substantive and theoretical implications of the Werther Effect. *American Sociological Review* **39**, 340-354.

Pinker, S. (1998) *How the Mind Works*. Allen Lane The Penguin Press, London.

Prete, A. and Miotto, P. (1997) Creativity, evolution and mental illnesses. *Journal of Memetics - Evolutionary Models of Information Transmission*, **1**. [http://jom-emit.cfpm.org/1997/vol1/preti\\_a&miotto\\_p.html](http://jom-emit.cfpm.org/1997/vol1/preti_a&miotto_p.html)

Rapaport, A. and Rebhun, L.I. (1952) On the mathematical theory of rumor spread. *Bulletin of Mathematical Biophysics* **14**, 375-384.

Rashevsky, N. (1951) Mathematical biosociology of beliefs and prejudices. *Bulletin of Mathematical Biophysics* **13**, 289-302.

Rashevsky, N. (1952) Prolegomena to a dynamics of ideologies. *Bulletin of Mathematical Biophysics* **14**, 95-118.

Richerson, P.J. and Boyd, R. (1978) A dual inheritance model of the human evolutionary process I: basic postulates and a simple model. *Journal of Social and Biological Structures* **1**, 127-154.

Rogers, E.M. and Shoemaker, F.F. (1971) *Communication of Innovations. A Cross-cultural Approach*. 2nd ed. Free Press: New York.

Rose, S., Kamin, L.J. and Lewontin, R.C (1984) *Not in Our Genes: Biology, Ideology and Human Nature*. Penguin, London.

Sigelman, C.K. and Sigelman, L. (1976) Authority and conformity: violation of a traffic regulation. *Journal of Social Psychology* **100**, 35-43.

- Skinner, B.F. (1972) *Beyond Freedom and Dignity*. Cape: London.
- Speel, H-C. (1997) A memetic analysis of policy making. *Journal of Memetics - Evolutionary Models of Information Transmission*, **1**. [http://jom-emit.cfpm.org/1997/vol1/speel\\_h-c.html](http://jom-emit.cfpm.org/1997/vol1/speel_h-c.html)
- Stahl, S.M. and Lebedun, M. (1974) Mystery gas: an analysis of mass hysteria. *Journal of Health and Social Behaviour* **15**, 44-50.
- Takahasi, K. (1998) Evolution of transmission bias in cultural inheritance. *Journal of Theoretical Biology* **190**, 147-159.
- Wheeler, L. and Caggiula, A.R. (1966) The contagion of aggression. *Journal of Experimental Social Psychology* **2**, 1-10.
- Wilkins, J. S. (1998) What's in a meme? Reflections from the perspective of the history and philosophy of evolutionary biology. *Journal of Memetics - Evolutionary Models of Information Transmission*, **2**. [http://jom-emit.cfpm.org/1998/vol2/wilkins\\_js.html](http://jom-emit.cfpm.org/1998/vol2/wilkins_js.html)
- de Winter, K.W. (1984) Biological and cultural evolution: different manifestations of the same principle. A systems-theoretical approach. *Journal of Human Evolution* **13**, 61-70.

## THE CASE FOR COMMENTARY

By

Derek Gatherer

School of Biomolecular Sciences

Liverpool John Moores University, Byrom St., Liverpool L3 3AF  
D.Gatherer@livjm.ac.uk

The article 'Why the Thought Contagion Metaphor is Retarding the Progress of Memetics' argues that the field of memetics took a wrong turning in the early 1980s, with the result that all subsequent memetic theorising has driven down the same blind alley. The wrong turning was to substitute the earlier, flexible definition of a meme (which in the article I term **Dawkins A**) with a far narrower and more specific definition as a unit of information in a mind or brain (termed **Dawkins B**).

Since science generally abhors vagueness and thrives on precise definitions, such a move naturally seemed like a step forward. In this case, however, it was counter-productive, since:

- a) Direct observation was relegated to a subsidiary role, while inference and speculation took centre stage. The derivation of internal memes from behaviour relies on acts of the most naive inference. This is not the same situation as that of a classical geneticist of the early 20<sup>th</sup> century, since genes were not naively inferred from phenotypes, but were indirectly demonstrable as segregation ratios, patterns of independent assortment or linkage. There is no equivalent phenomenology in memetics. Internal memes are not even indirectly demonstrable. Their existence cannot be independently verified in any way.
- b) Quantification became impossible. We cannot quantify any internal mental unit, and quantification is as essential as precise definitions (if not more so). While it is possible to quantify behaviour, either roughly in the field or more precisely in the laboratory, such quantification of behaviour does not translate into quantification of internal memes, unless the internal meme and the behaviour exist in a one-to-one correspondence. There is no evidence for this.
- c) The theory lacks explanatory power. To say that behaviour can be explained in terms of something which cannot be observed is unhelpful. In any case, there is also no evidence that replication of behaviour is necessarily correlated with replication of internal mental states. We have no grounds for believing that thoughts are contagious, and such a view trivialises psychology.
- d) Lynch's calculus of mnemonic conjugations is incompatible with any present thinking in either cognitive sciences or linguistics. The picture of the mind as a stack of 'awareness of statement x' and 'belief in statement x' would require the mind to store language in a manner reminiscent of a computer RAM. Such possibilities were effectively excluded by Chomsky in the early 1950s. I should therefore like to ask the commentators if they consider 'internal' memetics to be still tenable in the light of these criticisms. If so, how do they believe it should be constituted as an experimental science (as opposed to a pseudoscience dealing in postulated unobservable entities,

where neither verification nor falsification are possible)? If the commentators agree with me that 'internal' memetics is untenable, do they tend towards my quasi-behaviourist version of 'external' memetics, or have they a third alternative?

# A PLEA FOR METHODOLOGICAL DARWINISM: A COMMENTARY ON ROSE'S PAPER: CONTROVERSIES IN MEME THEORY

By

Derek Gatherer

School of Biomolecular Sciences, Liverpool John Moores University, Byrom St., Liverpool L3  
3AF  
D.Gatherer@livjm.ac.uk

*"We, alone on earth, can rebel against the tyranny of the selfish replicators."*  
(Dawkins 1978 [1976], p.215)

Free will and intentionality have bedevilled philosophers for centuries. "*We have free will and there's an end on't*", declared Samuel Johnson in the 18<sup>th</sup> century. Others, from Johnson's contemporary the Marquis de Laplace to B.F. Skinner in the 20<sup>th</sup> century, have, for a range of different reasons, been less certain. A middle road is often taken, which allows for strong exogenous influences on human behaviour while retaining a fundamental bedrock of human autonomy. The trick, apparently, is to learn to exercise that autonomy. Thus for Herbert Marcuse and his fellow Frankfurt School Marxists, the workers labour not merely under physical and economic repression, but under the hypnotic sway of a cultural hegemony that convinces them that they are not really repressed. For Freud's heterodox disciple Wilhelm Reich, repression is not so much economic as sexual, and indoctrination of inhibitions takes place at an early age. Either way, the solution is to recognise the source of the cultural 'programming' and the reward for casting it off is freedom. Jean-Paul Sartre by contrast, in his early philosophy, leans towards the Johnsonian end of the spectrum. For the early Sartre, the Marxist and Freudian accounts of human behaviour sound more like excuses than explanations. We are always free to choose, whether we like it or not, and to fail to recognise that is to commit the only cardinal sin recognised by the existentialist, that of mauvaise foi (Sartre 1969 [1943]).

Recognition that genes may have phenotypic effects that are not always to the benefit of the individual organisms that contain them, presaged by J.B.S. Haldane but beginning rigorously with the work of W.D. Hamilton in the 1960s, led to the body of evidence summarised by Dawkins (1978 [1976]) in *The Selfish Gene*. Dawkins, of course, added the meme as a second, non-genetic, replicator. Just as some genes are beneficial to their hosts, while others are more concerned with propagating themselves at their hosts' expense, so too are memes arranged on a spectrum from the unreservedly wonderful (eg. science in Dawkins' estimation) to the truly parasitic. Dawkins' intellectual heritage owes more to Dr Johnson than to Marx, but his picture of religious memes as the ultimate 'mutually compatible gang of viruses' (Dawkins 1993) has distinct affinities with Louis Althusser's 'false consciousness' working in favour of a sinister status quo (Althusser 1977). Darwinian algorithms have replaced dialectic, but the call to arms is the same: it's time to rebel against the selfish replicators.

Dennett (1987, 1991) takes the meme theory a stage further. Now the meme is not seen merely as a replicator within consciousness, but as the essence of consciousness itself. Of course, Dennett does not intend that memes should be units of phenomenology - qualia are disqualified

in his theory - but that the illusion of consciousness is produced through the combined action of myriads of behavioural and abstract memes. The higher functions of consciousness, the ability to choose, the ability to exercise our freedom, the ability to design intentionally, thus no longer require a homuncular agent in the brain, no longer require a Sartrean pour-soi or a Johnsonian free will. We are after all automata, but of an infinitely more sophisticated kind than the Marquis de Laplace could ever have envisaged. Our consciousnesses are so Machiavellian that they can even fool us into believing that we have qualia, that we have intentionality, that we have freedom. Clearly, in this view rebellion against the selfish replicators is impossible. We are made of selfish replicators, they do not merely colonise our otherwise healthy minds. They are our minds, healthy or otherwise.

Against this, Anthony O'Hear (1997) maintains that there is nevertheless still a problem with our free will (illusion or not, it's still a factor in which memes get selected). O'Hear would accept Nick Rose's proposition that attribution of variation and selection of memes to conscious foresight does undermine the claim of memetics to be an evolutionary project. In reply to Rose's question - why do we need an evolutionary theory of culture at all? - O'Hear might reply that we don't need one and furthermore we should not concern ourselves with attempting to find one.

So we have Dennett in one corner, updating Laplace and Skinner to produce a non-homuncular, non-Cartesian theory of consciousness as grand illusion, and in the opposite corner we have Dr Johnson, Sartre and O'Hear, maintaining that we really are free come what may and that there is no way any Darwinian algorithm is going to straighten out that one. Somewhere in the middle we have Dawkins, the Freudians and the Marxists, acknowledging our capacity for freedom but pessimistic concerning its ability to resist the exogenous forces that dictate to it. Rose stands over on the Dennettian side, and naturally sees the others as fudging the task of explaining the mind in a mechanistic evolutionary manner. Rose is right, if Universal Darwinism is the game, they have given up on it.

I cannot claim to contribute anything novel to the debate concerning free will or intentionality, but I suggest that part of the problem lies in the tendency to see memetics as a theory of the mind. Dennett, of course, is primarily interested in the mind, and much of the recent resurgence of memetics is owing to his sponsorship. But memetics, both in Dawkins' original version (Dawkins 1978 [1976]) and in its treatment by Cavalli-Sforza and Feldman (1981), was originally a theory of culture couched in fairly materialist terms. If we adopt a population-level approach to culture, if we proceed from meme as behaviour, artefact or other materially identifiable unit straight up to whole culture as meme pool, skipping the grey and murky intervening level of individual human consciousnesses, then the problem, if not exactly solved, becomes less serious. I concede that those who believe in intentionality as a real thing, and not merely an illusion, may still protest that the origins of memetic novelty are therefore not random, that we have a spanner in the works of the Darwinian algorithm. Darwinism would seem to require that variation be randomly generated and not in some way directed towards some goal, otherwise the ghost of Lamarck intervenes. However, to this I would reply that we can only really address this possibility by empirical observation. Analysis of novel behaviours in the context of the cultural systems in which they appear is the only way to determine if the non-randomising effect of intentionality in individuals will be sufficient to destabilise the Darwinian algorithm as applied to the 'big picture' of culture.

I thus plead for a methodological Darwinism in cultural evolution rather than a theoretical one, just as I plead for a methodological behaviourism rather than a Skinnerian one. O'Hear may be right, perhaps intentionality does rig the jury where cultural novelty is concerned. However, the big picture may still nevertheless be amenable to Darwinian analysis. Only empirical studies can answer this question. So do we need an evolutionary theory of culture? Perhaps not, but I'm going to continue on the assumption that we do.

## References

- Althusser, L. (1978). *Lenin and Philosophy, and Other Essays*. trans. Brewster, B., New Left Books: London.
- Cavalli-Sforza, L.L. and Feldman. M.W. (1981). *Cultural Transmission and Evolution: A Quantitative Approach*. Princeton University Press: Princeton.
- Dawkins, R. (1978 [1976]). *The Selfish Gene*. Granada Paladin: Oxford.
- Dawkins, R. (1993). Viruses of the Mind. In: *Dennett and his Critics*. Dahlbom, B. (Ed.) Blackwell: Oxford. pp.13-27.
- Dennett, D.C. (1987). *The Intentional Stance*. MIT Press: Cambridge.
- Dennett, D.C. (1991). *Consciousness Explained*. Allen Lane The Penguin Press: London.
- O'Hear, A. (1997). *Beyond Evolution: Human Nature and the Limits of Evolutionary Explanation*. Clarendon: Oxford.
- Sartre, J.-P. (1969 [1943]). *Being and Nothingness: an Essay on Phenomenological Ontology*. trans. Barnes H.E., Methuen: London.

## **REPLY TO THE COMMENTARIES ON MY PAPER: “WHY THE THOUGHT CONTAGION METAPHOR IS RETARDING THE PROGRESS OF MEMETICS”**

By  
**Derek Gatherer**  
[D.Gatherer@organon.nhe.akzonobel.nl](mailto:D.Gatherer@organon.nhe.akzonobel.nl)

Firstly I should like to thank all the contributors for taking the time to read my article and comment upon it. Secondly, I ought to apologise for the lateness of my reply. This is a consequence of my absence in South America for 6 weeks at the end of 1998, and also my abrupt and unexpected departure from academia at the beginning of February. Consequently I am still having problems with library access, and have been unable to do the reading required to address some of the points made by the commentators. Nevertheless, I hope I can use this opportunity to clarify a few ambiguities that have clearly arisen.

David Hales (1999) makes the point that the 'intentional stance' is a necessary feature of everyday life. I agree entirely, but I am unsure that this is really an argument against behaviourism in the laboratory or field. Dennett's work, such as the intentional stance, and (especially) his subsequent 'disqualification of qualia' are, as Dahlbom (1993) has pointed out, aspects of behaviourism in its modern guise. Behaviourism is not, contra Heylighen (1999), an outmoded school of psychology but one which continues to evolve and adapt to the changing empirical landscape. Dennett's work, including the 'intentional stance', is just one important manifestation of that process. The intentional stance is precisely that, a stance. It has much in common with what Paul Marsden has called the 'memetic stance'. In both cases we are required to analyse situations 'as if' intentions/memes (delete as applicable) were involved, and ask how such assumptions help to explain the behaviour we see. I would say that the intentional stance is best taken in what Hales (1999) calls 'its weakest form' as an 'instrumentalist device' - although I accept John Wilkins' (1999) caution that I should not rely overly on instrumentalism. Whether or not it also gives the user 'enhanced predictive power', as Hales says, is a matter for empirical demonstration. Hales (1999) holds out some hope yet for the Dawkins B model, in which case I suggest a further argument which may be used against it (Houghton's 'shopping list' thought experiment, see Houghton 1997, pp. 162-163). I can also point to the most recent reviews from those actually looking for memes in human brains. Their results are not encouraging to the Dawkins B school (Lounasmaa et al 1996, Gall et al. 1998).

Francis Heylighen (1999) correctly points out that a theory-free science would be scarcely scientific. My determination to throw out most of post-Dawkins B memetic theory may seem to give the impression that I am against theory in principle. It is true that I have no novel theory of my own to offer in its place, but I would insist that my own work is couched in the standard (neo-) Darwinian Theory which involves identification of variation, replication and selection. There are newer theories which may also be of assistance in the construction of a behaviourist memetics, for instance the recent work of Wallace et al (1996) and Wallace and Wallace (1998). Theoretical constructs are necessary, I agree. Our job is to find new ones that will succeed where older varieties have failed. Like most memeticists, I came to the subject through theoretical speculation, but I now see myself as a 'coal-face' empiricist. Coal-face workers are bound to

revolt occasionally against bad (theoretical) management. But that does not mean we are opposed to management in principle.

Heylighen (1999) refers to the interesting contagious suicide problem, and states that "*what was replicating was not the actual suicide behaviour, but the idea that suicide is a good way out of an apparently unsolvable problem*". I have to disagree with this. I fear it makes too many assumptions. Many of the people who succumb to the strange phenomenon of contagious suicide are precisely those who would seem to have the least reason to do so. 'Unsolvable problems' are rarely the issue. If they were, the phenomenon would scarcely be so macabre and fascinating, or so interesting a subject for memetics.

Paul Marsden (1999) suggests that the strategy should be a focus for memetic analysis. I am perfectly happy with strategies that can be analysed behaviourally. I have done a little empirical work on lecture attendance as a student learning strategy. In some cases, it seems to work quite well and in other cases does not seem to have any relevance to the postulated goal (see Gatherer and Manning 1998). Experimental design is difficult, and ethical considerations regarding students' careers do rather limit the extent to which the experimenter can manipulate the system. But I agree with Marsden, strategy-oriented memetics may represent the way forward. Hans-Cees Speel (1999) requests a clarification of my point regarding the meme-host duality. This is an important plank of my thesis, so I ought to try to improve on my original formulation, as follows.

Individuals have genes in the 'ordinary language' sense of the word 'have'. Molecular genetics has demonstrated that we do indeed have these genes in the same way that we have the grosser parts of our anatomy. The Dawkins B meme is also something which we could have in the same way. If there are replicating information structures in our brain, then the same ordinary language use of the word 'have' could apply to Dawkins B memes just as it does to genes. I would submit, however, that there are no such things. We could have Dawkins B memes, but we don't because there aren't any. Therefore, in ordinary language usage, we do not have Dawkins B memes.

What about Dawkins A memes? Well, these are a very different and much more heterogeneous class of entities, being behaviours and artefacts and all the other messy things that empirical work throws at us. Although there is another ordinary language use of the word 'have' which applies to artefacts, for instance I may have a remaindered copy of 'Thought Contagion', this is not the same sense in which I have my genes, or in which I could have the fictitious Dawkins B memes.

I speculate that there is perhaps a sense of unease among our non-Anglophone colleagues, concerning such fussiness about the use of words. It is true that ordinary language philosophy in the spirit of JL Austin is not something that has travelled well outside of the Anglo-American world. That honorary Anglo-American, Ludwig Wittgenstein, described ordinary language philosophy as a way of showing the fly out of the fly bottle. I maintain that the meme-host duality is a fly bottle in which the memetics movement has become trapped. Our escape from this particular trap also leads to the demonstration that we can have no true population memetics, as follows.

Population genetics is based conceptually on the assignment of genes to individuals, usually two alleles per locus in a diploid Mendelian population. Those individuals have those genes. They have them in the strong ordinary language sense of 'have'. This is crucial. If they ceased to have them in such a sense, population genetics would crumble; all the equations depend on the assignment of genes to individual members of the population. In order to have a population memetics, it would be necessary to assign Dawkins B memes to individuals with the same degree of rigour. But we can't. There are two reasons why not. The first is the trivial one, which is that there are no such things as Dawkins B memes. The second is exemplified in my 'Windsor knot' thought experiment, and also in Houghton's (1997) 'shopping list' thought experiment.

All we are left with is Dawkins A memes. These are infuriatingly transient entities which fail to satisfy the strong sense of the word 'have' which is necessary to a population science, be it population genetics or population memetics. Heylighen (1999) makes some additional point that genes can have multiple phenotypic effects, such as a feather gene whatever that may be, producing several hundred feathers. I do not think this is relevant. A feather gene may produce several hundred or thousand feathers, but they are all on the same individual. That individual has that feather gene, in the strong sense of the word. There is a definite relationship between the feather allele and the individual hosting it. How many feathers that gene produces on that individual is of no consequence to my argument.

This brings me on to another point, made by John Wilkins (1999) which is the hypothetical entity (HYPE) nature of the Dawkins B meme. HYPES are often discovered to be real entities; that which begins as a convenient theoretical construct can often turn out to have a genuine reality. For instance in the pre-Watson/Crick era, genes were HYPES. However, there is an important sense in which pre-Watson/Crick genes ('classical' genes we might say) were much less HYPE than Dawkins B memes. Wilkins alludes to 'Mendelian genes, observed by their effects'. They are of course observed by their effects, but not just by their effects. I well remember from my student days the long list of sceptical criteria which classical geneticists (my experience was with the Aspergillus school in Glasgow) applied to any claim for the discovery of a novel gene. A classical gene has to be true breeding, it has to demonstrate segregation in crosses, it has to consistently assort independently from, or map together with, other loci etc. The classical gene was not merely a HYPE lying behind the empirical reality of phenotype, but an empirical phenomenon in its own right, defined by all manner of rigorous criteria. The Dawkins B meme, by contrast, has no such criteria, no rigour of any description. It is simply a HYPE lying behind the empirical reality of behaviour.

Wilkins (1999) raises another interesting question: how, in a behaviourist framework can one distinguish between homology and analogy, eg. between little Charlie's accident and Prince Philip's butler? I cannot pretend to have an answer to this. However, I am consoled by the fact that it is equally difficult to distinguish between these things in genetic evolution, even at the molecular level.

Wilkins (1999) also refers to the informational aspects of memes. I think his argument here has affinities with a thought experiment presented to me by Sue Blackmore and friends at Meme Lab, which runs as follows: if I tell you a joke while I have a hoarse voice, you do not tell the joke to your friends using the same hoarse voice, or even using my normal voice. Rather you tell

the joke in your own normal voice. We are able to extract a certain informational content from a speech behaviour which does not involve all aspects of the behaviour.

Wilkins (1999) postulates a realm of 'culturally significant transmits' (in his Fig. 2) Having just about removed myself from the influence of Popper, I have to confess I feel a little frisson of fear at this spectre of World 3. Having said that, behaviourists as well, of course, have to come to terms with 'memes without a hosting subject' (to paraphrase Popper). Popper lurks behind so much of memetics, but his exact relationship to the discipline is still not completely defined. I hesitate to address the comments of Aaron Lynch (1999), for fear that I leave myself open to further accusations of willful misrepresentation. Lynch insinuates that my criticism of his work is religiously motivated. Elsewhere, referring to the change from Dawkins B in my earlier papers to my current behaviorist stance, he ruminates darkly "*Exactly what happened to change his mind is unclear*". I should have thought it was very clear - reflection on 20 years of internalist memetics persuaded me it was going nowhere. He also insinuates that I was '*recently still unaware of the evolutionary epidemiological ("thought contagion/mind virus") explanation for the prevailing forms of monotheism by natural selection*'. Alas, I was aware of them, but they are so vacuous as to be scarcely worthy of comment in a scholarly journal. Since Lynch insists, I deal with them below.

Lynch also repeats his peculiar allegation that I have misrepresented him, drawing attention to an 'if' which he claims that I have ignored, thus changing the sense of his argument. However, the 'if' in question makes no difference. Lynch's picture of menmons arranged ~PQ~TY etc is quite clearly a stack of memory bits, exactly like a computer RAM. His calculus of mnemon conjugations then deals with individuals as hosts of string of the said bits ~PQ~T~Y etc. Why he should now seek to deny this is unclear. It is in his paper.

Lynch also claims that I have a "*casual lack of understanding of the term infinite*". What can one say? This is nonsense. Of course I understand the meaning and use of the term infinite. And what is the basis for such a strange allegation? His subsequent sentences are unclear, but he appears to be arguing my case. It is precisely because as he says 'it is physically impossible for humans with finite brains to construct an infinite number of dates for Napoleon's death' that his calculus of mnemon instantiations is untenable, as developed in detail in my article.

Even more bizarrely, Lynch refers to my 'monotheism'. What monotheism? I am not a monotheist, or indeed a theist of any description. However, this outlandish ad hominem argument does present an opportunity to deal with Lynch's own specific theory concerning monotheism, which is that: 'In an ancient society, where people believed in numerous gods, a belief that there is only one God had a competitive propagation advantage - not as genetic information, but as mental information. It caused its hosts to devote all of their meme transmission efforts to just one god-meme, allowing this god-meme to out-propagate competing memes.' (Lynch 1998)

As with much of Lynch's work, it is difficult to know if this is a serious proposition or just some kind of witty self-parody. Unfortunately I must assume the former. If monotheism really requires less propagative effort than polytheism, and will out-compete polytheism on this basis, one must be able to explain why Christianity was the successful Trinitarian derivative of a not particularly

successful strictly monotheistic religion (Judaism). Christianity, with its almost incomprehensible concepts of 3-gods-in-1 and gods turning into lunchbox items, rapidly became even more polytheistic with the acquisition of a fourth member of the pantheon (the BVM) and a whole panoply of minor deities (the communion of saints). Monotheism, contra Lynch (1998) does not automatically outcompete polytheism, as the example of Christianity shows. Indeed, the incorporation of Roman polytheism in growing Christianity, taking it away from its monotheistic origins, was a crucial element in its spread. One wonders if Lynch thinks at all about his speculations before committing them to publication.

Similarly we have, concerning baseball 'Does internally stored information about numbers of players to recruit explain why baseball out-propagates tennis in the US? You have to read Lynch for yourself'. Rather than analyse such a laughable claim, I shall simply quote Caton (1997), who agrees with me that Lynch "*suppresses empirical psychology so that personal speculations may flourish.*"

## References

- Caton, H. (1997). Explaining culture: a naturalistic approach. *Politics and the Life Sciences*. September 1997:319-323.
- Dahlbom, B. (1993). *Dennett and his Critics*. Oxford: Blackwell.
- Gall, C. M., Hess, U. S. and Lynch, G. (1998). Mapping brain networks engaged by, and changed by, learning. *Neurobiology of Learning and Memory*, **70**:14-36.
- Gatherer, D. and Manning, F. C. R. (1998). Correlation of examination performance with lecture attendance: a comparative study of first-year biological sciences undergraduates. *Biochemical Education*, **26**:121-124.
- Hales, D. (1999). Belief has utility - an intentional stance. *Journal of Memetics - Evolutionary Models of Information Transmission*, **3**. [http://jom-emit.cfpm.org/1999/vol3/hales\\_d.html](http://jom-emit.cfpm.org/1999/vol3/hales_d.html)
- Heylighen, F. (1999). The necessity of theoretical constructs: a rebuttal of the behaviourist approach to memetics. *Journal of Memetics - Evolutionary Models of Information Transmission*, **3**. [http://jom-emit.cfpm.org/1999/vol3/heylighen\\_f.html](http://jom-emit.cfpm.org/1999/vol3/heylighen_f.html)
- Houghton, D. (1997). Mental content and external representations. *Philosophical Quarterly*, **47**:159-177.
- Lounasmaa, O. V., Hämäläinen, M., Hari, R. and Salmelin, R. (1996). Information processing in the human brain - magnetoencephalographic approach. *Proceedings of the National Academy of Science (USA)*, **93**:8809-8815.
- Lynch, A. (1998). Memetic monotheism. *Skeptic*, **6**:26-27.

Lynch, A. (1999). Misleading mix of religion and science. *Journal of Memetics - Evolutionary Models of Information Transmission*, **3**. [http://jom-emit.cfpm.org/1999/vol3/lynch\\_a.html](http://jom-emit.cfpm.org/1999/vol3/lynch_a.html)

Marsden, P. (1999). A strategy for memetics: memes as strategies. *Journal of Memetics - Evolutionary Models of Information Transmission*, **3**. [http://jom-emit.cfpm.org/1999/vol3/marsden\\_p.html](http://jom-emit.cfpm.org/1999/vol3/marsden_p.html)

Speel, H-C. (1999). On memetics and memes as brain entities. *Journal of Memetics - Evolutionary Models of Information Transmission*, **3**. [http://jom-emit.cfpm.org/1999/vol3/speel\\_h-c.html](http://jom-emit.cfpm.org/1999/vol3/speel_h-c.html)

Wallace, R., Fullilove, M. T. and Flisher, A. J. (1996). Aids, violence and behavioral coding: information theory, risk behavior and dynamic process on core-group sociogeographic networks. *Soc. Sci. Med.*, **43**:339-352.

Wallace, R. and Wallace, R. G. (1998). Information theory, scaling laws and the thermodynamics of evolution. *Journal of Theoretical Biology*, **192**:545-559.

Wilkins, J. (1999). Memes ain't just in the head. *Journal of Memetics - Evolutionary Models of Information Transmission*, **3**. [http://jom-emit.cfpm.org/1999/vol3/wilkins\\_j.html](http://jom-emit.cfpm.org/1999/vol3/wilkins_j.html)

# MODELING THE EFFECTS OF MEMETIC TABOOS ON GENETIC HOMOSEXUALITY

By

Derek Gatherer

derek-gatherer@usa.net

## Abstract

Simple computer simulations of the interaction of genetic factors and memetic taboos in human homosexuality are presented. These simulations clearly show that taboos can be important factors in the incidence of homosexuality under conditions of evolutionary equilibrium, for example states produced by heterozygote advantage. However, frequency-dependent taboos, i.e. taboos that are inversely proportional to the incidence of homosexuality, cannot produce the oscillating effect on gene frequencies predicted by Lynch (1999). Effective oscillation is only produced by rapid withdrawal and re-imposition of taboos in a non-frequency-dependent manner, and only under conditions where the equilibrium incidence of homosexuality is maintained by heterozygote advantage, or other positive selectional mechanism. Withdrawal and re-imposition of taboo under conditions where homosexuality is subject to negative selection pressure, produce only feeble pulses, and actually assist in the extinction of the trait from the population. Additionally it is shown that frequency-dependent taboos assist in a more rapid achievement of equilibrium levels, without oscillation, under conditions of heterozygote advantage. An attempt is made to relate the simulations to past and contemporary social conditions, concluding that it is impossible to decide which model best applies without accurate determination of realistic values for the parameters in the models. Some suggestions for empirical work of this sort are made.

**Key words:** gene-meme co-evolution, homosexuality, memetics

## 1 Introduction

Evidence for a genetic basis for homosexuality in humans remains controversial (Pillard and Weinrich 1986; Bailey and Pillard 1991; Bailey and Benishay 1993; Bailey et al 1993; Hamer et al 1993; Risch et al 1993; Haynes 1995; Hu et al 1995; Rice et al 1999). It is clear that psychological, sociological and environmental influences also play a part (e.g. Werner 1979). However it would be surprising if there were not at least some genetic predisposition, given the weight of evidence for biological contributory factors (e.g. LeVay 1991; Allen and Gorski 1992; Snyder et al 1994; Reite et al 1995; Swaab and Hofman 1995; Zhou et al 1995; Sanders and Wright 1997; McFadden and Pasanen 1999). Among the possible non-genetic influences may be counted the existence of taboos against homosexuality in many societies. It has been proposed that frequency-dependent taboos may interact with genetic influences to produce an oscillating frequency of homosexuality, specifically:

*"In a nutshell, ... [the theory]... says that adherents of the taboo out-procreated more tolerant people over the course of many generations in ancient times, leading to increased prevalence of*

*the taboo. (This does not require anything like a perfect correlation between morality and behaviour, or a perfect child inculcation rate, but only enough to increase taboo prevalence by several percent per generation over hundreds of generations.) Then horizontal transmission kicked in as people maligned homosexuality to "prove" their adherence to the taboo. As the taboo becomes extremely widespread, most homosexuals live heterosexual lives, leading them to reproduce any genes involved. As these genes gain prevalence, the rate of taboo dropout increases. Gene carriers who have dropped the taboo are more sexually and socially motivated to spread acceptance of homosexuality than are non-gene carriers who drop the taboo. So the rising gene prevalence can lead to a self-sustained propagation of pro-gay memes. That, in turn, can lead to lower gene prevalence in the next generation, and even lower prevalence of pro-gay memes. All of this leads to potential fluctuations over long time spans." (Lynch 1999)*

This paper investigates this hypothesis using a simple model incorporating Fisher's classic equations of natural selection, upon which is superimposed a model of a memetic taboo as a restriction on the freedom to openly practice a homosexual lifestyle. The taboo is assumed to be strictly memetic, i.e. there is no genetic predisposition to intolerance, and there is no tendency for individuals of any particular genotype to exhibit one meme as opposed to another, i.e., there is no gene-meme linkage disequilibrium (Laland 1992). Various computer simulations were scripted in Perl, and used to analyse the effects of taboos and heterozygote advantage on the incidence of genes predisposing to homosexuality. The Perl scripts are available on [http://www.geocities.com/derek\\_gatherer/supp.htm](http://www.geocities.com/derek_gatherer/supp.htm).

## **2 Methods and results**

There is no clear indication of the location in the human genome, of any genes predisposing to homosexuality. The best candidate is the tip of the long arm of the X chromosome, but the evidence for this is still not absolutely convincing (Hamer et al 1993, Hu et al 1995; Rice et al 1999). In the present paper, homosexuality is modelled as an autosomal recessive trait, primarily for ease of simulation.

### **2.1 Model of selection against homosexuality as an autosomal recessive trait** **In this section the following assumptions are made:**

- a) predisposition to homosexuality is a single-locus, biallelic autosomal recessive;
- b) heterozygotes are heterosexual;
- c) the homosexual phenotype results in a lower level of sexual reproduction than the heterosexual phenotype;
- d) a large population is modelled, in which there is no significant genetic drift, and no immigration of groups with different genetic compositions.

Following Fisher's classic model of selection against an autosomal recessive trait (given in Ayala and Kiger 1984), a hypothetical population can be modelled, in which  $a/a$  is the genotype of individuals predisposed to homosexuality,  $A/a$  is the genotype of heterozygous individuals, and

$A/A$  are individuals carrying no copies of the gene predisposing to homosexuality. A small illustration follows of the way that selection pressures are calculated in the classic methodology. Following the tradition in classical genetics, these individuals  $A/A$  are referred to as the 'wild-type', and  $A$  as the 'wild-type allele'. H-W is the expected proportion of each genotype under the Hardy-Weinberg equilibrium.

	$A/A$	$A/a$	$a/a$
<b>H-W</b>	0.25	0.5	0.25
<b>Parents</b>	2500	5000	2500
<b>Progeny</b>	7500	15000	750
<b>Av./indiv</b>	3	3	0.3
<b>Fitness</b>	1	1	0.1
<b>s</b>	0	0	0.9

**Table 1.** H-W is the expected frequency of genotypes under the Hardy-Weinberg equilibrium,  $s$  is the selection pressure. 'Av./indiv.' Is the average number of progeny per individual

Table 1 shows an illustrative population of 10,000 individuals in an initial Hardy-Weinberg equilibrium. Thus, if  $q$  is frequency of the allele  $a$  predisposing to homosexuality and  $p$  is the frequency of the wild type allele  $A$ , and  $p + q = 1$  (i.e. a biallelic locus), then the frequencies of our phenotypes are:

$$\begin{array}{ll} A/A & p^2 \\ A/a & 2pq \\ A/a & q^2 \end{array}$$

In table 1,  $q = 0.5$ , thus the proportions of  $A/A:A/a:a/a$  are 1:2:1.

The population in Table 1 is growing rapidly, purely for illustrative purposes. The frequencies of the progeny in the next generation are counted. (The progeny of each male and female is counted separately, so each offspring is effectively counted twice - this however is of no bearing on the result, as the *relative* quantities of progeny are the important issue in the calculation of fitness and selective pressure).

In the example in Table 1  $A/A$  and  $A/a$  individuals produce an average of 3 progeny each, but  $a/a$  individuals only produce 0.3 offspring per individual. Fitness ( $w$ ) is the relative reproductive success of each genotype compared to the most successful genotype. By this criterion, genotypes  $A/A$  and  $A/a$  have fitness  $w = 1$ , and genotype  $a/a$  has fitness  $w = 0.1$ . Selection pressure,  $s = 1 - w$ , so for genotype  $a/a$  this is 0.9.

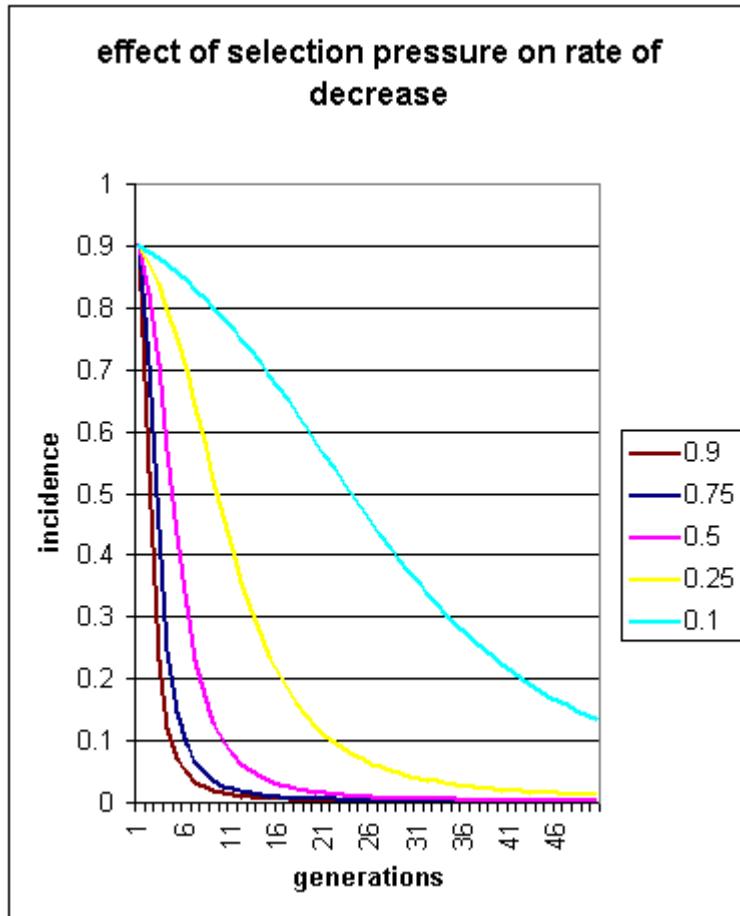
One may then proceed to Fisher's model of change in allele frequency per generation for a disadvantageous autosomal recessive trait (given in Ayala and Kiger 1984)

$$(1) \quad \Delta q = \frac{-spq^2}{1-sq^2}$$

## 2.2 Simulation 1: Effect of differing value of $s$ on the incidence of homosexuality

The initial incidence of homosexuality in the population is taken to be 0.9, i.e. 90% of the population is homosexual. This kind of extreme *reductio ad absurdum* is used purely for the purposes of making the clearest illustration of the trends predicted by the model. Later, more realistic parameters are discussed. This simplest simulation in which negative selection pressure gradually eliminates a gene from a population, is, of course, trivially true from the viewpoint of geneticists. It is included however, in order that other readers may see the most basic framework of the Fisherian analysis.

5 different values of  $s$  are chosen ranging from 0.9 down to 0.1. The results are plotted below (Figure 1):



**Figure 1:** incidence of homosexuality plotted over 50 generations, at varying levels of  $s$ , from 0.1 to 0.9.  $s$  has an effect on the rate of decrease, but the overall trend is always downwards

It can be seen that in all cases, the incidence of homosexuality declines. At the highest value of  $s$ , 0.9, only 1.5% of the population is homosexual after 10 generations. Even at the relatively mild level of  $s = 0.1$ , corresponding to a reproductive deficiency of only 10% per generation, 10 generations produces a decline in homosexuality from 90% of the population to 78.4%.

It is thus clear that if homosexuality is due to recessive genetic factors, and if homosexuals are less successful at sexual reproduction than heterosexuals, (i.e. if  $s > 0$ , even by a small value), homosexuality will eventually disappear from the population (except for new mutations). This disappearance can be astonishingly rapid, e.g. at  $s = 0.9$  homosexuality falls from 90% of the population to 23% in a mere 3 generations.

### 2.3 Simulation 2: Introduction of taboo

Let the taboo,  $b$ , be quantified as the proportion of homosexual individuals who lead heterosexual lives.

$$b = 1 - (h_{\text{vis}}/h_{\text{total}})$$

where  $h_{\text{vis}}$  are 'visible' homosexuals, leading openly homosexual lives and  $h_{\text{total}}$  is the total number of predisposed homosexuals of genotype  $a/a$ .

It is assumed that those individuals of genotype  $a/a$  who lead heterosexual lives have rates of sexual reproduction equal to individuals of genotypes  $A/A$  and  $A/a$ .

Therefore  $b$  reduces  $s$  as follows:

$$s_{\text{tab}} = s(1-b)$$

where  $s_{\text{tab}}$  is the selection pressure in the presence of taboo  $b$ , and  $s$  is the selection pressure in the absence of any taboo.

$s_{\text{tab}}$  can then be substituted for  $s$  in equation (1) above.

Which can also be reexpressed using Laland's (1992) nomenclature, as:

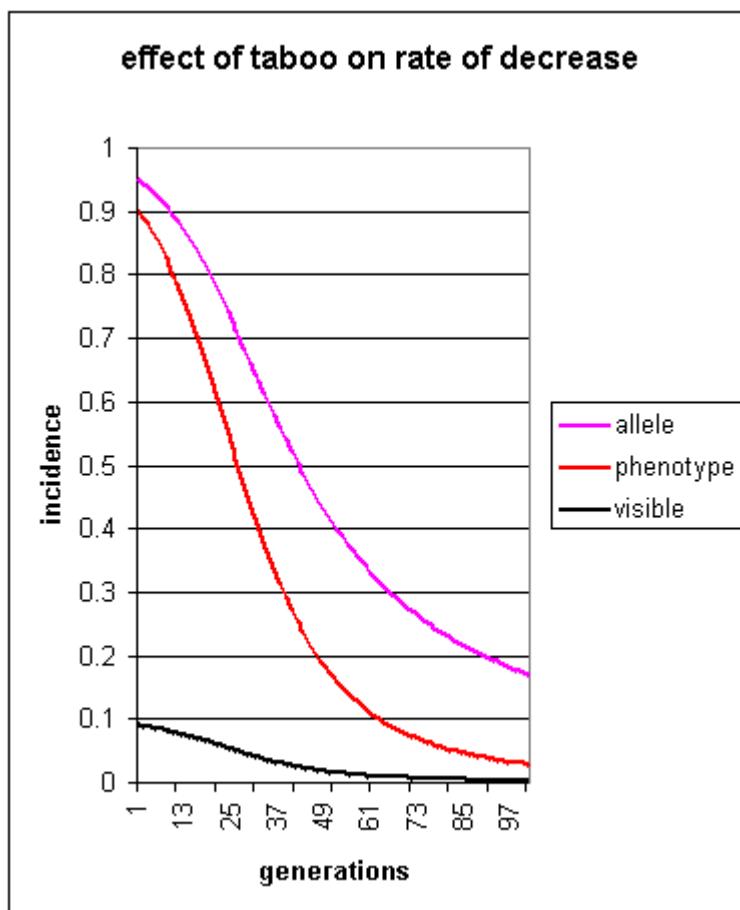
$$(2) \quad \Delta q = \frac{-r_2 spq^2}{1 - r_2 sq^2}$$

where there are two cultural traits,  $r_1$  and  $r_2$ , and only the phenogenotype  $a/a$ .  $r_2$  is deleterious.  $b$  is kept constant by requiring that Laland's (1992) expression for the horizontal transmission of a meme within a single generation, result in no change, i.e.. that  $N_d(i)$  should always equal  $N_d(i-1)$ :

$$(3) \quad N_d(i) = \frac{N_e}{1 + \left( \frac{N_e}{N_o} - 1 \right) e^{-\ln(1-a)i}}$$

where  $N_e$  is the total population,  $N_0$  is the initial population exhibiting the meme,  $N_d(i)$  is the population exhibiting the meme after  $i$  cultural generations (and there are  $i$  cultural generations per biological generation), and  $a$  is the rate at which the meme tends to be replicated horizontally within that cultural generation.  $N_e$  and  $N_0$  can be assumed to be constants for the present purposes, so the crucial variables are  $i$ , which is arbitrarily set at 1, and  $a$ . Therefore simply by adjusting  $a$ , the taboo,  $b$ , can be kept constant within the population. Adjustments in which  $b$  is allowed to change owing to changing values of  $a$ , are of course, possible. However,  $b$  is kept constant again here for the purposes of simplicity.

Let  $s$  and  $b$  both be 0.9. Let the initial frequency of individuals of genotype  $a/a$  be 0.9 (again an absurdly high starting value is taken for the purposes of illustration).



**Figure 2:** Frequencies of alleles predisposing to homosexuality, the predisposed homosexual phenotype and visible homosexuality, plotted over 100 generations under conditions of taboo,  $b = 0.9$

As can be seen, under conditions of taboo of 0.9 the number of visible homosexuals in the population is initially only 9%. The frequency of visible homosexuality, and of predisposed individuals  $a/a$ , falls. After 28 generations the number of visible homosexuals has fallen to 4.5%. Under such circumstances, it can be seen that taboos alone cannot act to increase or even maintain levels of homosexuality, but can only slow the rate at which it decreases (as is generally predicted for any situation of this kind by Laland 1992, Figure 2 of that article).

Lynch (1999) states "*As the taboo becomes extremely widespread, most homosexuals live heterosexual lives, leading them to reproduce any genes involved. As these genes gain prevalence...*". It can be seen from the above simulation that this is impossible. A taboo of 0.9 has been chosen, and yet genes predisposing to homosexuality still decrease rapidly. They certainly do not 'gain prevalence'.

#### 2.4 Simulation 3: Fluctuating levels of taboo

Let initial levels of homosexuality be 0.9, as before. Let the taboo initially be 0.9, but then be decreased to 0.1 after 3 generations, simulating a major social attitude change. Let the taboo increase again to 0.9 after a further 3 generations, and return again to 0.1 after 3 more etc. Again this is a crude simplification of what in real life are bound to be complex political, social and anthropological factors affecting taboos. Once more, these assumptions are made purely for the purposes of clarity in the simulation. Since it is clear from simulation 2, that a steady taboo will not make genes predisposing to homosexuality 'gain prevalence', the purpose here is to investigate what effect a fluctuating taboo will have.

This situation is produced by again manipulating Laland's equation (3) above; such that  $a$  is alternately extremely high for the taboo, then extremely high for the non-taboo in 3-generational blocks. Why  $a$  might vary in such a way is irrelevant to present purposes.

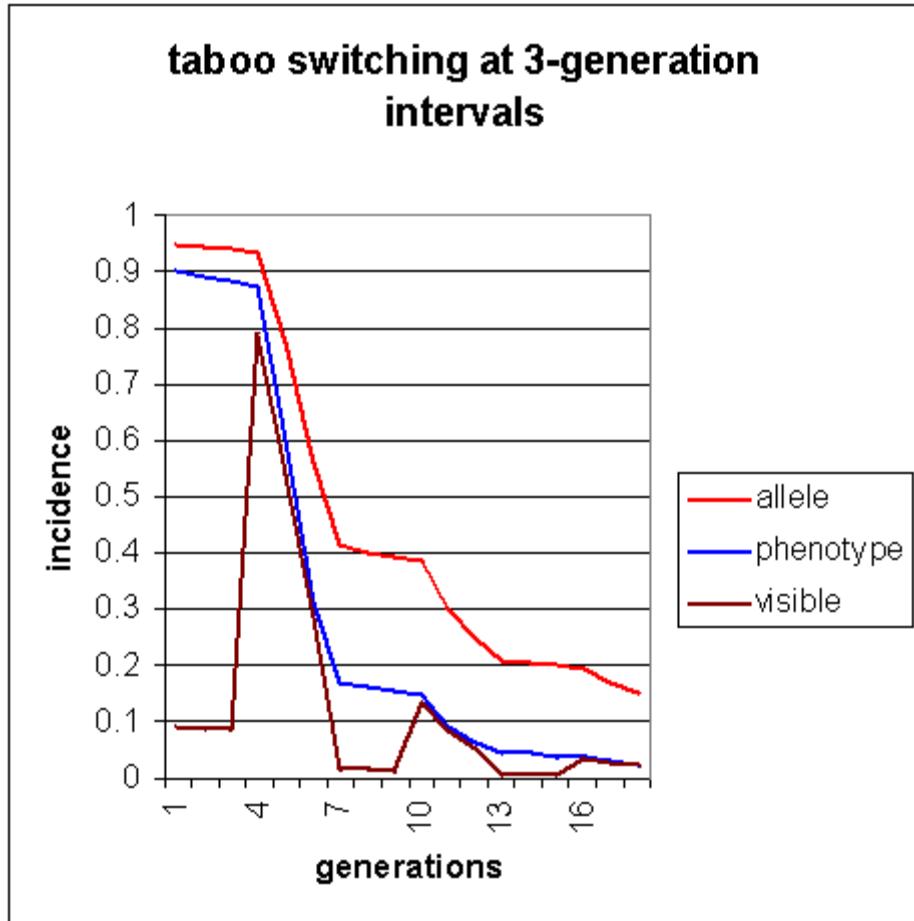
The withdrawal and imposition of the taboo in 3-generation cycles produces a faster overall decline than when the taboo is constant at 0.9. Compare Figure 2 in simulation 2, where  $s = 0.9$ ,  $b = 0.9$ , in which the same interval of 12 generations only produces a decline in visible homosexuality from 9% to 7.7%, with Figure 3 above, in which the withdrawal and reintroduction of the taboo every 3<sup>rd</sup> generation produces a reduction in 12 generations from 9% to 5.4%

Whenever the taboo is withdrawn, as in generations 4 and 10, there is an increase in visible homosexuality. However, the second sudden increase is much lighter than the first, and further pulses of this sort, such as that caused by the release of the taboo in generation 16, are scarcely visible.

#### 2.5 Simulation 4: Incorporating heterozygote advantage

Among the most potent forces for the maintenance of allele frequencies in populations is heterozygote advantage, or heterosis. This phenomenon is established as an important factor in sickle cell anaemia, since heterozygotes are malaria resistant (reviewed by Durham 1996). There is some evidence for cholera resistance in cystic fibrosis heterozygotes (Gabriel et al 1994), and for ochratoxin resistance in phenylketonuria heterozygotes (Woolf 1986). Tuberculosis resistance has also been hypothesised in Tay-Sachs disease heterozygotes (see Spyropoulos *et al* 1981 and Diamond 1988 for both sides of the debate). Possible heterozygote advantage in homosexuality is discussed by Ruse (1981).

Again following Fisher (as reproduced in Ayala and Kiger 1984), it is possible to produce an equation for the change in allele frequency in the case of heterozygote advantage. In this situation, there is a selection pressure against  $a/a$ ,  $s$  as before, but also a selection pressure  $t$  against wild-type homozygotes  $A/A$ .



**Figure 3:** Incidence of allele predisposed to homosexuality, the predisposed phenotype and visible homosexuality over 20 generations under conditions where the taboo is withdrawn and imposed at 3 generation intervals.

$$(4) \quad \Delta q = \frac{pq(tp - sq)}{1 - tp^2 - sq^2}$$

Under such conditions, an equilibrium level is achieved at:

$$(5) \quad q = \frac{t}{s+t}$$

Let  $s$  be 0.9 as before.

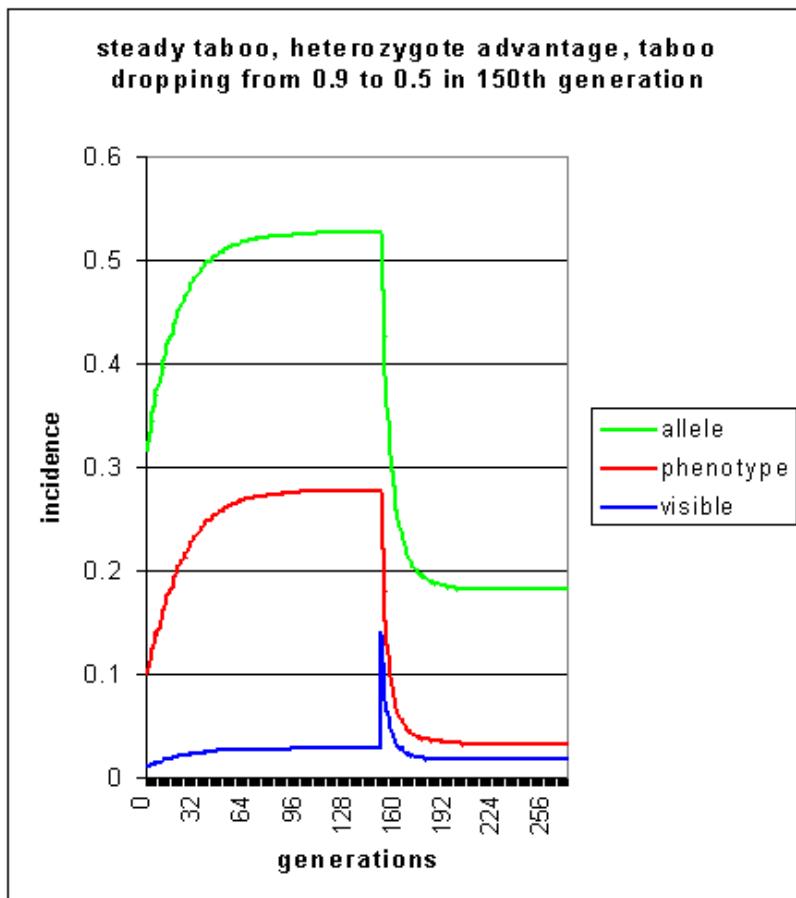
Let  $t$  be 0.1 This means that wild-type homozygotes  $A/A$  have a reproductive or survival deficit of 10% relative to heterozygotes.

Let the taboo,  $b = 0.9$ .

Let the initial frequency of homosexuality in this case be 0.1.

Let these conditions apply for 150 generations, after which the taboo decreases to  $b = 0.5$  (again by manipulating the variable  $a$  in Laland's equation (3) above to suit our purposes)

This population initially has a level of visible homosexuality of 1%, but the heterozygote advantage allows this to increase to 2.8% over 95 generations. This is virtually an equilibrium level, which is sustained for a further 55 generations. When the taboo is suddenly diminished after 150 generations, the level of visible homosexuality peaks suddenly to 14% of the population. However, with  $b$  now at 0.5,  $s_{\text{tab}}$  is now 0.45. The population begins to move towards a new equilibrium level of visible homosexuality of approx 1.7% achieved after a further 43 or so generations.

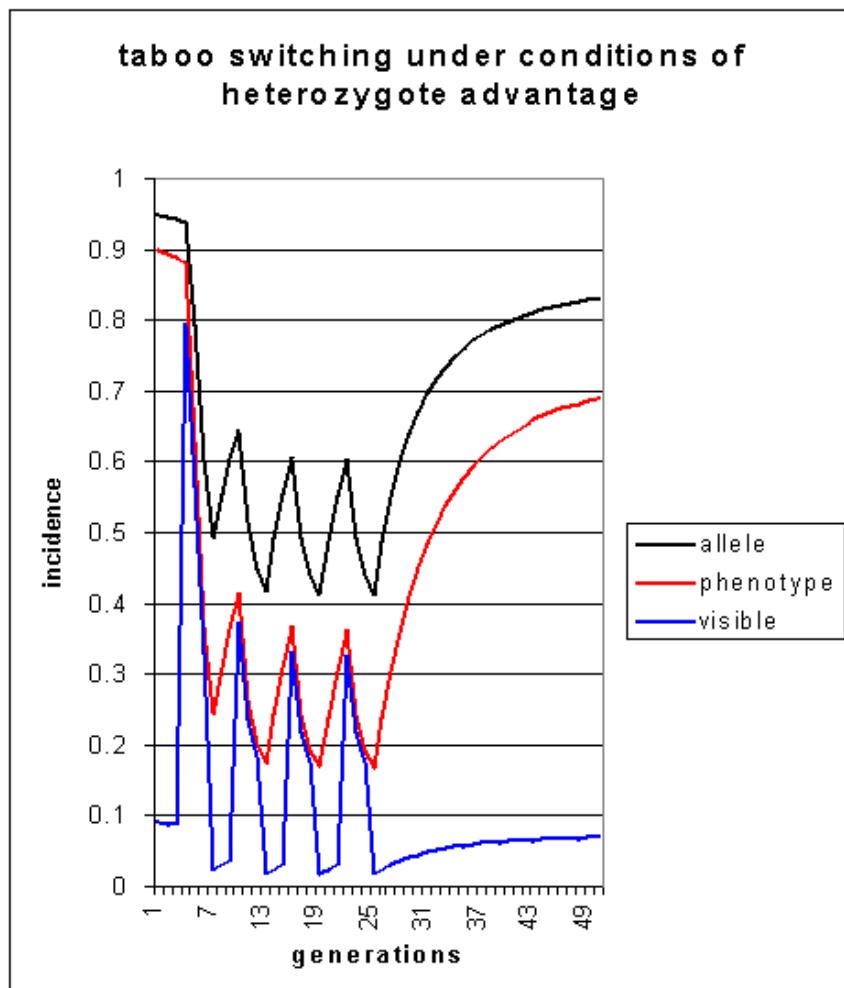


**Figure 4:** Incidence of allele predisposing to homosexuality, the homosexuality phenotype and visible homosexuality, under conditions of heterozygote advantage, over 300 generations. The taboo drops from 0.9 to 0.5 in generation 150

Heterozygote advantage thus allows homosexuality to achieve equilibrium levels, whereas simulations 1-3 above all result in downward trends. As was pointed out before in simulation 2, this result is textbook material for geneticists. However, it is included here in order that non-genetical readers may clearly see the effects of heterozygote advantage on a deleterious allele.

## 2.6 Simulation 5: Heterozygote advantage with fluctuating taboo

Let the heterozygote advantage be heavy,  $t = 0.5$ . Let  $s$  again be 0.9, and the initial incidence of homosexuality be 0.9. The taboo,  $b$ , is initially 0.9 and switches between 0.9 and 0.1 every third generation until the 25<sup>th</sup> generation when it returns permanently to  $b = 0.9$ . Once again, it is not suggested that such arbitrary fluctuations are in any way a realistic portrayal of any current or past political situation. The purpose is merely to illustrate how heterozygote advantage may interact with fluctuating taboos.



**Figure 5:** Incidence of allele predisposing to homosexuality, the homosexually predisposed phenotype and visible homosexuality, under conditions of heterozygote advantage, over 50 generations. The taboo switches between 0.9 and 0.1 every third generation until the 25<sup>th</sup> generation when it returns permanently to taboo = 0.9

Again the alternating release and reimposition of the taboo cause pulses in the level of visible homosexuality. These pulses are much stronger than when there is no heterozygote advantage (see simulation 3), and when the taboo is permanently reapplied, the level of visible homosexuality stabilises at around 7% in the 50<sup>th</sup> or so generation. Only an excessively implausible level of heterozygote advantage, at  $t = 0.5$ , can produce such a striking effect.

## 2.7 Simulation 6: Introducing a frequency-dependent taboo

The frequency-dependent taboo is a taboo which diminishes as the frequency of homosexual individuals rises ("the rising gene prevalence can lead to a self-sustained propagation of pro-gay memes" Lynch 1999). Note that this is not 'Frequency-Dependence at the Memetic Level' (Laland 1992), since the frequency of the taboo is dependent on the frequency of the genetically determined phenotype, not on the frequency of other memes.

Therefore in each generation, the frequency dependent taboo  $b_{fd}$

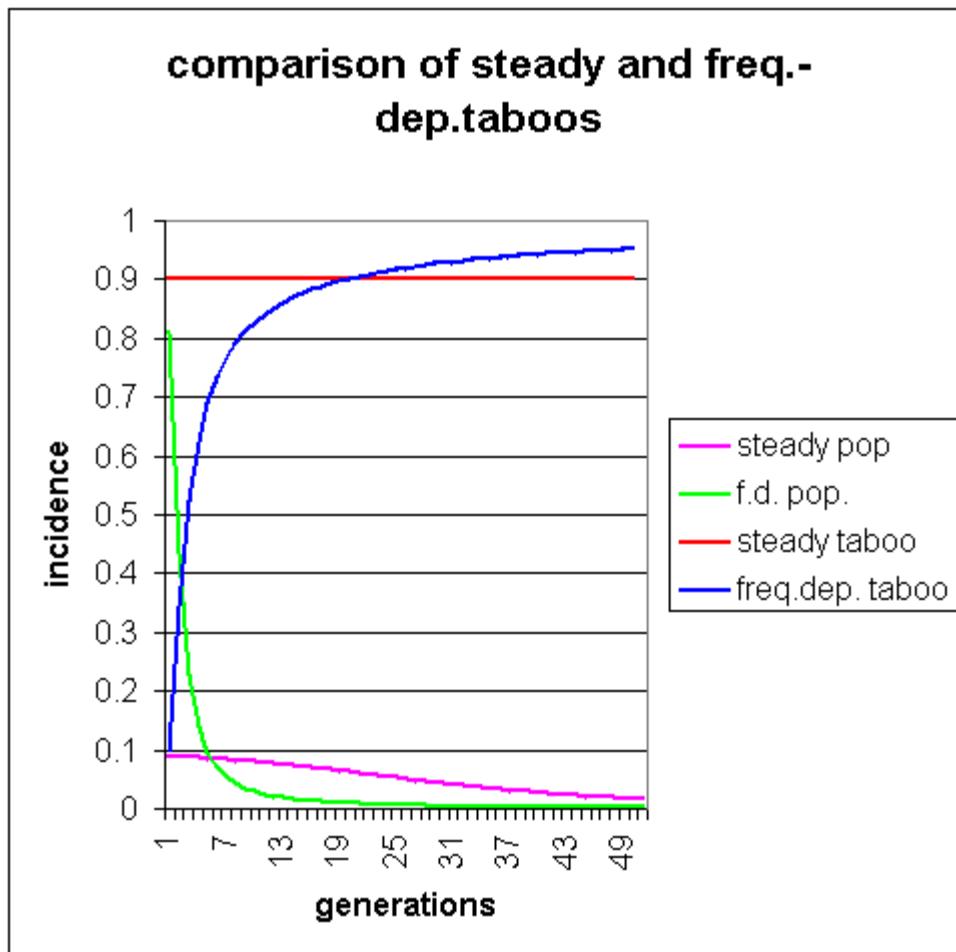
- $b_{fd} = 1 - q^2$
- and the selection pressure,  $s_{fd}$  is then
- $s_{fd} = s_{init} (1 - b_{fd})$
- where  $s_{init}$  is the selection pressure in the absence of a frequency-dependent taboo.

Let us compare 2 populations, both starting with incidences of homosexuality of 0.9, and starting with  $s = 0.9$ , but one population having a steady taboo of 0.9, and the other exhibiting a frequency-dependent taboo, where the  $b_{fd}$  changes in each generation as above.

The steady taboo population has a slow decrease in the levels of visible homosexuality from 9% to 4.5% in 23 generations. The frequency-dependent taboo population initially has high levels of visible homosexuality and consequently a low taboo. However, as levels of visible homosexuality drop dramatically under the influence of  $s = 0.9$  in the absence of heterozygote advantage, the taboo climbs, and in generation 20 exceeds the level in population 1. Such high levels of taboo, achieving 0.95 by generation 48, means that visible homosexuality is virtually eliminated from population 2. This shows that frequency-dependent taboos do not produce the oscillating effect that Lynch (1999) claims. On the contrary, a taboo of this sort would speed the eventual elimination of the genes predisposing to homosexuality from the population.

## 2.8 Simulation 7: Frequency-dependent taboo combined with heterozygote advantage

Let us compare 2 populations as above, but incorporating heterozygote advantage. Both populations have  $s = 0.9$ , and  $t = 0.1$ . One population has a constant taboo 0.9 and the other has a frequency dependent taboo.



**Figure 6:** Incidence of visible homosexuality and level of taboo, over 50 generations. 'Steady pop.' is the frequency of visible homosexuality in the population with the steady taboo. 'F.d. pop.' is the frequency of visible homosexuality in the population with the frequency-dependent taboo

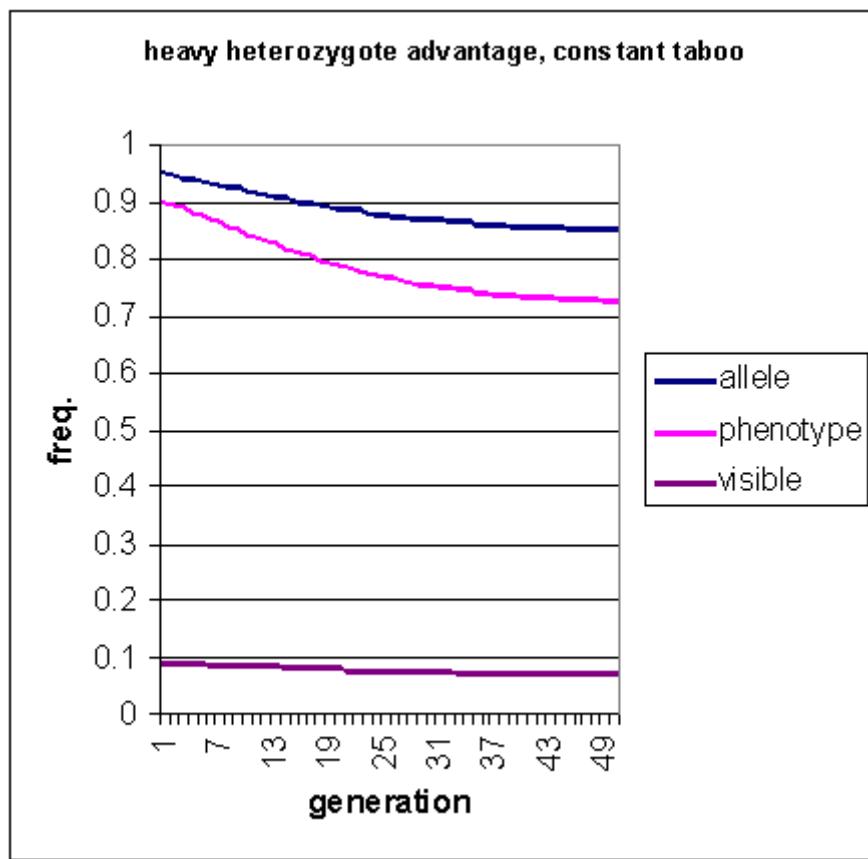
The population with the frequency -dependent taboo rapidly achieves an equilibrium level after only 12 generations, whereas the population with constant taboo takes 50 generations. Note that there are no pulses in the frequency of visible homosexuality. This demonstrates that neither constant nor frequency-dependent taboos can produce the oscillating effect predicted by Lynch (1999).

### 3 Discussion

The above models have dealt with extreme and implausible values of incidence and selection pressure, simply in order to make the trends on the graphs more obvious. In this section an attempt to simulate a realistic set of variables will be attempted.

The fundamental assumption of this paper is that the predisposition to homosexuality is primarily genetic, but that the behavioural expression of that predisposition is determined by memetic factors such as taboos. These models therefore only represent a limited subset of possible

models. For instance one might construct a model in which homosexuality is entirely culturally determined, or in which homosexuality is related to environmental factors which are neither

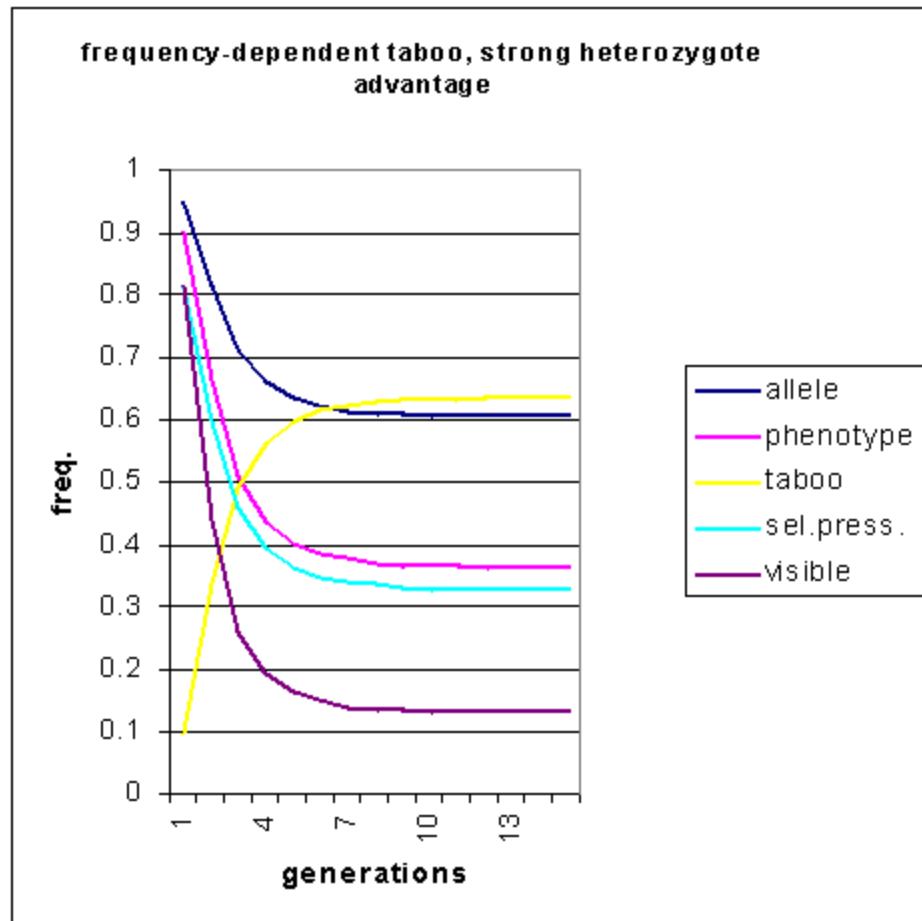


**Figure 7:** Incidence of allele predisposing to homosexuality, the homosexually-disposed phenotype and visible homosexuality, under conditions of heterozygote advantage, over 50 generations. The taboo is constant.

genetic nor cultural. Nevertheless, there are compelling reasons to suggest that the gene-meme interactionist approach is at least plausible, given the weight of empirical research that has accumulated on the biological basis of homosexuality and on how cultural attitudes affect it.

Exact quantification of the incidence of male homosexuality in the past and present is a highly controversial area (Gonsiorek et al 1985, Lhomond 1993; Weinrich et al 1993), but an estimate of a current incidence of visible homosexuality of some 5% seems not unreasonable (Johnson 1994). Waugh et al (2000)'s intriguing 'lost letter' experiment suggests that the general public are only 49% as likely to behave altruistically to an overtly homosexual as to a non-homosexual person, which means that one might measure taboos accordingly. This would bring the score of *a/a* individuals up to 10% of the population which is still within some of the upper-end estimates (e.g. Thin and Smith 1976). It is clear that visible homosexuality has recently increased, at least in Western (post)-industrialised societies. Downey (1980) surveyed over 5 generations, and concluded that homosexuality has increased overall among males but that it has decreased among married males, over this time. The decrease in the number of married homosexual males

is strongly suggestive of a decrease in taboo. However, the *overall* increase may suggest that there has *also* been an actual increase in the frequency of *a/a*, perhaps due to a heterozygote



**Figure 8:** Incidence of allele predisposing to homosexuality, the homosexually-disposed phenotype and visible homosexuality under conditions of heterozygote advantage over 15 generations. The taboo is frequency-dependent.

advantage in the process of achieving equilibrium (e.g. simulation 4). It is possible that the release of previous taboos has been superimposed onto this. If taboos are frequency-dependent (Lynch 1999), then an increase in numbers of *a/a* individuals due to sex-specific advantage, or a heterozygote advantage in the process of achieving equilibrium, would also trigger decrease in the taboo.

It is clear from the simulations presented here that frequency-dependent taboos do not result in the oscillations predicted by Lynch (1999). By contrast, oscillations can only be produced by imposition and release of *frequency-independent* taboos, and *even then* only when phenomena such as heterozygote advantage or sex-specific selection are *also* operating.

Despite the difficulty of disentangling genetic and cultural factors in the short-term, it may be possible to gather some empirical facts that might help in choosing the most likely model. The models for heterozygote advantage and sex-specific advantage (the latter not explicitly discussed

here, but see Ruse 1981; Gallup and Suarez 1983; Dickemann M 1995; Miller 2000) require a fairly extreme set of parameters in order for them to be able to sustain high levels of homosexuality in populations. In the case of the heterozygote advantage model presented here,  $A/a$  individuals should have a greater reproductive success.  $A/a$  individuals represent two-thirds of the siblings of  $a/a$  individuals. If the current level of homosexuality is approximately 10%, i.e.  $q^2 = 0.1$  (Thin and Smith 1976), and homosexuals have currently half the level of reproductive success of heterosexual males, i.e.  $s = 0.5$  (this is pure conjecture - this parameter is one which needs empirical research), then the heterozygote equilibrium equation (equation 5, above): trivially solves to:

$$t = 0.24$$

Therefore, if the above estimates for  $q$  and  $s$  are correct, one requires  $t = 0.24$  to sustain this in a system of heterozygote advantage. This implies that the  $A/a$  siblings of homosexuals have a strong reproductive/survival advantage over wild-type  $A/A$  individuals, of some 24%. The magnitude of the required  $t$  is therefore scarcely plausible, and one might be tempted to exclude the heterozygote advantage hypothesis on this basis. However, empirical data is need to ascertain if the above estimates for  $q$  and (especially)  $s$  are accurate. The 1994 National Survey of Sexual Attitudes and Lifestyle (Johnson 1994) estimated the incidence of homosexuality to be between 1.1% and 6.4% of men and 0.4% and 3.4% of women. If the most conservative estimate of 1.1% is taken, then  $t$  under heterozygote advantage where  $s$  remains 0.5, becomes  $(0.10 \times 0.5)/0.9 = 0.06$ , which is perhaps a more plausible value. If  $s$  is overestimated, then  $t$  reduces even further.

A 1970 survey (Fay et al 1989) placed the level of homosexual contact in the previous 12 month period as being some 1.6 to 2.0% of males. This gives an upper bound to  $q$  of 0.14. Furthermore, half of these males were found to be 'currently or previously married'. If we assume these males to have been equally reproductive as heterosexual men within their marriages, and if we assume never-married homosexual males to be non-reproductive, then we have  $s = 0.5$ . If we assume that non-reproductive homosexuality is predominantly male in its expression, then  $s$  reduces further to 0.25. We can then estimate  $t$  at  $(0.14 \times 0.25)/0.86 = 0.04$ . Heterozygote advantage need only be 4% if these values are accurate.

Furthermore, it is clear that although heterozygote advantage can (at least theoretically) sustain levels of homosexuality, and female-specific selection is even capable of increasing it, both of these mechanisms are unstable and vulnerable to sudden fluctuations in taboo (but *not* to frequency-dependent taboos). It is also clear that taboos alone in the absence of a positive selective mechanism such as heterozygote advantage, are insufficient to account for oscillating patterns in visible homosexuality. For instance compare simulations 2 and 3 with simulation 5. Simulation 5 has a strong oscillation *only* because there is *also* strong heterozygote advantage in the system.

One must be careful to distinguish current selective pressures from those operating in the past. It is possible that genes predisposing to homosexuality were elevated in the past through heterozygote advantage or female-specific selection, but that those selective pressures no longer apply. Therefore genes predisposing to homosexuality would now be in the model described in simulation 1, and be in a steady decline. The apparent 20<sup>th</sup> century increase in visible

homosexuality (Downey 1980) might just be a blip caused by a release in a taboo, which will be followed by a greater decrease in future generations. Alternatively, there may be selective mechanisms, such as heterozygote advantage or female-specific selection which are still in operation.

In summary, it is not possible to make accurate predictions concerning the application of these models to the real world without data concerning the reproductive success of siblings of homosexual males, and indeed of homosexuals themselves. If such data could be obtained, an estimate could be made for the parameters of the heterozygote advantage and female-specific advantage models. If either of these models is found to be plausible in view of this empirical data, estimates could be made for future levels of homosexuality subject to any fluctuations caused by changes in taboo which, as shown, are at their most important when strong biological selection is occurring.

If the empirical data does not support either of the strong selection models above, it must be concluded that visible homosexuality will decline. On the other hand, it may be that gene-meme interaction models of the sort presented here, miss some crucial environmental, psychological or sociological element that can explain why homosexuality is maintained even when gene-meme analysis might suggest it should not be.

## References

- Allen LS, Gorski RA (1992) Sexual orientation and the size of the anterior commissure in the human brain. *Proc Natl Acad Sci U S A* **89**, 7199-202.
- Ayala FJ and Kiger JA (1984) *Modern Genetics 2<sup>nd</sup> ed.* Benjamin/Cummings, Menlo Park.
- Bailey JM and Pillard RC (1991) A genetic study of male sexual orientation. *Arch. Gen. Psychiat.* **48**, 1089-1096.
- Bailey JM and Benishay DS (1993) Familial aggregation of female sexual orientation. *Am. J. Psychiat.* **150**, 272-277.
- Bailey JM, Pillard RC, Neale MC and Agyei Y (1993) Heritable factors influence sexual orientation in women. *Arch. Gen. Psychiat.* **50**, 217-223.
- Diamond JM (1988) Tay-Sachs carriers and tuberculosis resistance. *Nature* **331**, 666.
- Dickemann M (1995) Wilson's Panchreston: the inclusive fitness hypothesis of sociobiology re-examined. *J. Homosex.* **28**, 147-183.
- Downey L (1980) Intergenerational change in sex behavior: a belated look at Kinsey's males. *Arch Sex Behav* **9**, 267-317.
- Durham WH (1991) *Coevolution: genes, culture and human diversity.* Stanford University Press, Stanford.

Fay RE, Turner CF, Klassen AD and Gagnon JH (1989) Prevalence and patterns of same-gender sexual contact among men. *Science* **243**, 338-48.

Gabriel SE, Brigman KN, Koller BH, Boucher RC and Stutts MJ (1994) Cystic fibrosis heterozygote resistance to cholera toxin in cystic fibrosis mouse model. *Science* **266**, 107-109.

Gallup GG and Suarez SD (1983) Homosexuality as a by-product of selection for optimal heterosexual strategies. *Perspect Biol Med* **26**, 315-22.

Gonsiorek JC, Sell RL, Weinrich JD (1995) Definition and measurement of sexual orientation. *Suicide Life Threat Behav* **25** Suppl, 40-51.

Hamer DH, Hu S, Magnuson VL, Hu N and Pattatucci AML (1993) A linkage between DNA markers on the X chromosome and male sexual orientation. *Science* **261**, 321-327.

Hu S, Pattatucci AML, Patterson C, Li L, Fulker DW, Cherny SS, Kruglyak L and Hamer DH (1995) Linkage between sexual orientation and chromosome Xq28 in males but not in females. *Nature Genet.* **11**, 248-256.

Haynes JD (1995) A critique of the possibility of genetic inheritance of homosexual orientation. *J Homosex* **28**, 91-113.

Johnson AM (1994) (ed.) *Sexual attitudes and lifestyles*. Blackwell Scientific Publications, Oxford.

Laland KN (1992) A theoretical investigation of the role of social transmission in evolution. *Ethology and Sociobiology* **13**, 87-113.

LeVay, S (1991) A difference in hypothalamic structure between heterosexual and homosexual men. *Science* **253**, 1034-1037.

Lhomond B (1993) Between man and woman: the character of the lesbian. *J Homosex* **25**, 63-73.

Lynch A (1999) Thought Contagion: How Belief Spreads through Society. A response to Paul Marsden. *Journal of Artificial Societies and Social Simulation* **2** (<http://jasss.soc.surrey.ac.uk/2/3/lynch.html>).

McFadden D and Pasanen EG (1999) Spontaneous otoacoustic emissions in heterosexuals, homosexuals, and bisexuals. *J Acoust Soc Am* **105**, 2403-13.

Miller EM (2000) Homosexuality, birth order, and evolution: toward an equilibrium reproductive economics of homosexuality. *Arch Sex Behav* **29**, 1-34.

Pillard RC and Weinrich JD (1986) Evidence of familial nature of male homosexuality. *Arch. Gen. Psychiat.* **43**, 808-812.

Reite M, Sheeder J, Richardson D and Teale P (1995) Cerebral laterality in homosexual males: preliminary communication using magnetoencephalography. *Arch Sex Behav* **24**, 585-93.

Rice G, Anderson C, Risch N and Ebers G (1999) Male homosexuality: absence of linkage to microsatellite markers at Xq28. *Science* **284**, 665-667.

Risch N, Squires-Wheeler E and Keats BJB (1993) Male sexual orientation and genetic evidence. *Science* **262**, 2063-2065.

Ruse M (1981) Are there gay genes? Sociobiology and homosexuality. *J Homosex* **6**, 5-34.

Sanders G and Wright M (1997) Sexual orientation differences in cerebral asymmetry and in the performance of sexually dimorphic cognitive and motor tasks. *Arch Sex Behav* **26**, 463-80.

Snyder PJ, Weinrich JD and Pillard RC (1994) Personality and lipid level differences associated with homosexual and bisexual identity in men. *Arch Sex Behav* **23**, 433-51.

Spyropoulos B, Moens PB, Davidson J and Lowden JA (1981) Heterozygote advantage in Tay-Sachs carriers?. *Am. J. Hum. Genet.* **32**, 375-380.

Swaab DF and Hofman MA (1995) Sexual differentiation of the human hypothalamus in relation to gender and sexual orientation. *Trends Neurosci* **18**, 264-70.

Thin RN and Smith DM (1976) Some characteristics of homosexual men. *Br J Vener Dis* **52**, 161-4.

Waugh IM, Plake EV and Rienzi BM (2000) Assessing attitudes toward gay marriage among selected Christian groups using the lost-letter technique. *Psychol Rep* **86**, 215-8.

Weinrich JD, Snyder PJ, Pillard RC, Grant I, Jacobson DL, Robinson SR and McCutchan JA (1993) A factor analysis of the Klein sexual orientation grid in two disparate samples. *Arch Sex Behav* **22**, 157-68 .

Werner D (1979) A cross-cultural perspective on theory and research on male homosexuality. *J Homosex* **4**, 345-62.

Woolf LI (1986) The heterozygote advantage in phenylketonuria. *Am. J. Hum. Genet.* **38**, 773-775.

Zhou JN, Hofman MA, Gooren LJ and Swaab DF (1995) A sex difference in the human brain and its relation to transsexuality. *Nature* **378**, 68-70.

# THE SPREAD OF IRRATIONAL BEHAVIORS BY CONTAGION: AN AGENT MICRO-SIMULATION

By  
**Derek Gatherer**  
gatherer@biotech.ufl.org

## Abstract

A micro-simulation is described, for rational and irrational strategies in human mating behaviour. The spread of irrational behaviour through a population from a single initial individual, the 'contagionist paradigm', is shown to be highly unlikely in most realistic circumstances. An exception to this rule is shown to be where the 'self-fulfilling prophecy' phenomenon is exhibited, i.e. the irrational meme affects the outcome of the mating. Additionally, where the irrational strategy, under conditions of self-fulfilling prophecy, is allowed to co-exist with a rational strategy (i.e. a strategy based on factual information), both can proceed to fixation, resulting in a population of individuals exhibiting both rational and irrational memes simultaneously. However, where successful pairs are removed from the population, there is a tendency for neither behaviour to persist. Maintenance of either behaviour in the population under circumstances of removal of successful pairs requires a cultural information system, i.e. one where a common pool of information may be accessed without a requirement for contagion. This implies that contagionist explanations of culture may be strictly limited in their application. Some attempt is then made to generalise the conclusions to financial systems.

## 1 Introduction

Human behaviour, at both the individual and social levels, is often apparently inexplicable. By contrast, the behaviour of software agents is, by definition, determined by their programmer, at the individual level. Nevertheless, at the social level, collections of software agents may behave as surprisingly as their real world counterparts. The appeal of agent-based social simulation is that one may design agents, of varying degrees of simplicity, that, when allowed to interact with each other, reproduce some of the identifiable behaviour of real human social systems. Some of the more puzzling aspects of human social systems may then be illuminated in the light of the computer simulation.

This paper investigates an agent population exhibiting two kinds of behaviour:

- Rational - defined as behaviour based on factual information, that maximizes the long-term material welfare of the agent, and
- Irrational - defined as behaviour based on incorrect premises, that is at best neutral to that agent's interests, and in some cases potentially detrimental.

In a simulation, it may be decided with certainty what constitutes 'factual information' and the best interests of the agent. In real life, of course, these can be more difficult to determine with

any reliability. It should also be stated that rationality in real organisms is a complex function that requires sophisticated cognitive functions. The agents described here have no AI component, but merely a series of rule-based functions. A considerable literature exists on the circumstances in which maladaptive, or apparently maladaptive, traits may become prevalent in human populations (e.g. Boyd & Richerson 1985, Durham 1991, Blackmore 1999). In this paper, the behaviours designated 'rational' and 'irrational' do correspond in many ways to 'adaptive' and 'maladaptive', respectively. However, the terminology 'rational' and 'irrational' is preferred since the 'irrational' behaviour in this simulation may coincidentally act in the agent's interests, and therefore is not always maladaptive.

These two kinds of behaviour are replicated in 2 ways:

- Contagion - where behaviour can only be copied directly from one agent to another via social contact (simple 'echo' contagion of the behavioural variety - Levy & Nail 1993; those authors' requirement that there should be non-intentionality is not relevant to software objects).
- Via a common pool of cultural information - an agent can select a behaviour from a set of social norms, without coming into direct contact with another individual exhibiting that behaviour.

The essential concern of this paper is related to that of Doran (1998), in that an effort is made to produce mass behaviour arising from a set of simple behavioural rules. Another methodological inspiration is Ophir (1998), in that pairing events within the agent population are followed by the execution of a variety of calculations affecting the subsequent state of the agents. The agents are reflex agents, as defined by Doran (1998), in that they have attributes which are modifiable according to a set of 'condition-action rules', and also trigger actions that modify the attributes of other agents. The model presented is, in Edmonds' (1998) terms, an 'abstract, forward' model, in that the interest lies in what kinds of unpredicted group phenomena may emerge out of strictly regimented individual behaviours.

The system simulated is that of human mate pairing. It is clear from biological studies that many species are very selective in mate choice, and that there may also be vigorous competition between individuals for mates. As Darwin (1872) recognised, these two factors have resulted in the phenomenon of sexual selection for a wide variety of traits in species ranging from insects to mammals. In some species, mate choice behaviour also has a cultural dimension which may amplify, or conflict with, any underlying innate biological tendencies (Wade & Pruett-Jones 1990; Laland 1994; Freeberg *et al* 1999). The simulation presented here attempts to dissect how a 'rational' mating behaviour, arbitrarily taken to be selection of a mate of same status as oneself, conflicts with an 'irrational' mating behaviour, for instance the use of astrology to select a mate. Note that such 'astrological' behaviour need not necessarily always be maladaptive, since a good mate may coincidentally be chosen. It is, however, always 'irrational' regardless of its adaptiveness. Although astrology is apparently believed by 25% of US adults (Blackmore 1999, p184), it may not necessarily figure very prominently in real human partner choice (but perhaps in other aspects of reproduction, see Goodkind 1993). However, the example is given merely to contrast irrational and rational memes in a simple contagion-based agent-pair interaction system.

Generally similar dynamics may occur wherever agents are exchanging information repeatedly in a closed system. Additionally, it should be pointed out that the mated pairs are not engaged in reproduction, or even necessarily in any physical copulation, but rather in partner choice for potential reproduction. In other words, this model takes place within a single generation.

Moreover, no genetic components are envisaged, and the behaviours are taken to be totally cultural. Finally, the population is not spatially segmented, and any individual can potentially mate with any other of the opposite sex. Ways in which the model could be expanded to incorporate a combination of genetics, multiple generations, and spatial heterogeneity are considered in the discussion.

## 2 Methods

The agents have the following attributes:

- Gender - male or female
- Status - high or low
- 'Star sign' - A or B
- Seek status information, i.e.. act 'rationally' - yes or no
- Seek astrology information, i.e. act 'irrationally' - yes or no

The status and astrology attributes (4 and 5) are thus different memes and not mutually exclusive 'allomemes' (Durham 1991). With 5 binary attributes, there are 32 different potential types of agent in the population, that is 16 different types for each gender. Since all pairings in all simulations are between members of opposite genders, there are 120 possible kinds of pairings, i.e.  $n(n-1)/2$ , for which rules need to be constructed. Some of the preliminary simulations only use a subset of the attributes and have very simple rules. The patience of the reader is requested for the apparent triviality of the first few simulations, as they are designed to clarify the basic dynamics of the system. From simulation 3 onwards, the system becomes increasingly complicated, as more attributes and complex rules of interaction are progressively applied.

Gender, status and 'star sign' (attributes 1 -3) are considered fixed for any individual and cannot be changed. The behavioural states (attributes 4 and 5) relating to status and astrology may be changed under certain circumstances, representing a behavioural change in that agent. A 'yes' state for astrology means that the individual will seek to ascertain the star sign of the other individual. Likewise a 'yes' state for status results in behaviour to determine the status of the potential partner. The status-'yes' state is 'rational' in that status objectively influences the likelihood of success of the pairing (in the simulation, whether or not it does so in real life is quite another matter - see Brown *et al* 1988; Pizzari 2001). By contrast, the astrology-'yes' state is 'irrational' in that star sign has no objective effect on success (see Goodkind 1993, for an empirical example of how astrology may influence human reproductive behaviour).

Agents seek a mate at random from among the opposite sex in the population (as in Ophir 1998). The outcome of the mating (successful or unsuccessful) is determined by a combination of random chance, a background a priori likelihood of success, and various rules dependent on the types of agent involved. Having mated every individual in the population, the cycle repeats. Each iteration, the proportions of the population exhibiting the enquiry behaviours, i.e.. the 'yes' states,

for status (rational) and star sign (irrational) are recorded. In some later simulations, successfully mated pairs are withdrawn from the mating pool and replaced with naïve individuals, as would be the case if mated pairs were to concentrate on reproduction and cease to look for further partners

Each simulation is produced using Perl scripts which, along with the full list of rules used in the more advanced simulations, are available from  
[http://www.geocities.com/derek\\_gatherer/supp.htm](http://www.geocities.com/derek_gatherer/supp.htm).

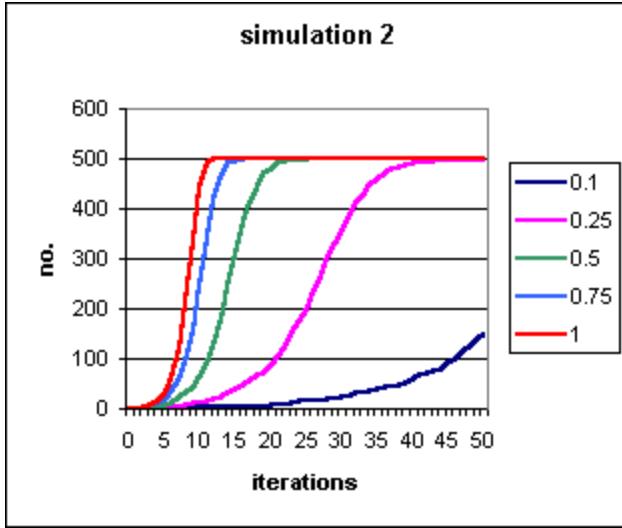
## 3 Results

### **3.1 Simulation 1 - Highly contagious irrational behaviour in a highly susceptible population**

The population is initialised with an equal number of individuals of each gender, each of which has randomly assigned status and star signs. Initially, all individuals are in the 'no' state for both behaviours. In this first simulation, a single individual is released into the population, exhibiting the 'yes' state for the irrational behaviour. A single simple rule is made that individuals only convert from 'no' to 'yes', and never in the reverse direction (in other words, this is the 'biased cultural transmission' of Takahasi 1998, 1999; or the 'cultural selection' of Cavalli-Sforza & Feldman 1981). As might be expected, under such a rule, the 'irrational' behaviour spreads rapidly throughout the population, essentially doubling in frequency for the first six cycles of pairing. After that point, its rate of increase slows a little as 'yes' individuals become more likely to encounter prospective partners who are already 'yes' in phenotype. In 3 simulations conducted in a population of 500 individuals, the entire population exhibits the irrational behaviour by cycle 12 in all 3 cases (data not shown). The dynamics are precisely the same as those of a highly infectious, incurable, sexually transmitted disease in a promiscuous population with no immunity. The total takeover of a population by a behaviour is called 'fixation', borrowed from population genetics - the opposite state being more obviously termed 'extinction'. The sigmoid curve of uncontrollable epidemic spread from a single individual to encompass the entire population is referred to here as the 'contagionist paradigm', and has been a regular feature of models of social processes at least as far back as Rashevsky (1949).

### **3.2 Simulation 2 - Contagious irrational behaviour with poorer replication rates**

Let the irrational behaviour now depend on an element of chance in its replication. Once again the population is seeded with a single individual exhibiting the irrational behaviour. With lower chances of replication at any particular pairing event, the progress to fixation is delayed. However, since an individual, once exhibiting the irrational behaviour, is permanently in that state, eventual fixation is inevitable, even when transmission rates are very low.



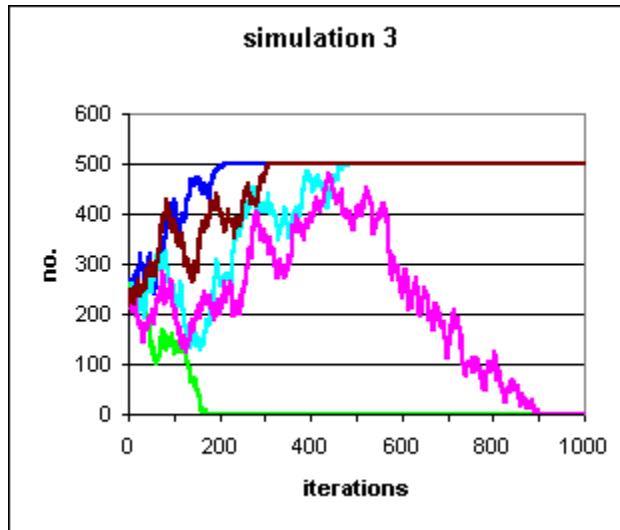
**Figure 1:** Spread of a contagious trait introduced by a single agent into a population of 500 agents, at differing values of probability of transmission per contact,  $p$ . Where  $p = 1$ , i.e. the trait is always transmitted, it approximately doubles every iteration, and a population of 500 is almost always fixed by generation 12. These S-shaped curves represent the most trivial case of contagion, and are presented as a reference point for the curves in the subsequent simulations.

Although these curves are sometimes fairly flat, their shape is still sigmoid, and the 'contagionist paradigm' still applies. The only variable is whether the contagion is slow or fast. For instance in Fig.1, for  $p = 0.1$ , i.e. 10% chance of transmission per contact, it still takes 45 iterations for 20% of the population to have been infected. Thus, for cases of 'biased transmission'/ 'cultural selection', the contagiousness of a trait is not important to the eventual outcome, *ceteris paribus*. It is however, precisely this *ceteris paribus* assumption that is violated in all but the most trivial cases.

### 3.3 Simulation 3 - Contagious irrational behaviour with possibility of loss of irrational behaviour

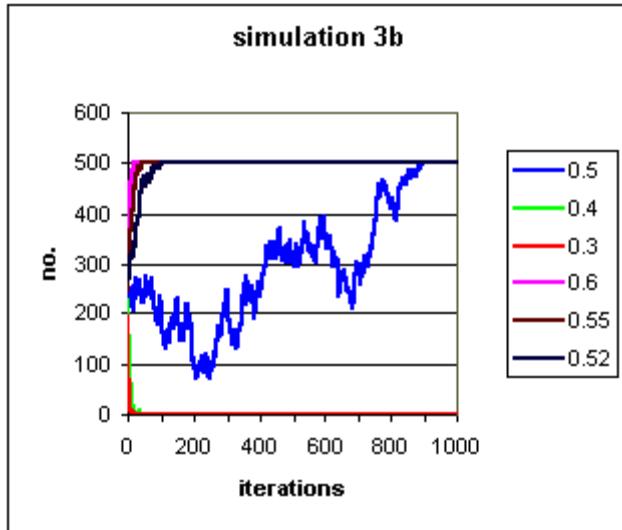
Let the 'no' state for the irrational behaviour now also be transmissible. This means that in any pairing event, it is possible that individuals exhibiting the irrational behaviour may be converted to the rational behaviour. Here, the analogy with an infectious disease begins to break down, as contact between two individuals is as likely to 'cure' the one as 'infect' the other, at  $p = 0.5$ . A better analogy is to the random drift of a rare neutral allele frequency in a genetic system (Kimura & Ohta 1978, Kimura 1979). It is clear that under such circumstances, with two mutually exclusive behaviours vying for predominance in the population, the eventual victor will be the one with the greater tendency to convert the other. This is merely a confirmation of intuition, but it is perhaps less intuitive that a corollary of this is that, at  $p = 0.5$ , about half of novel contagious behaviours will face immediate extinction in the first generation. In order to illustrate the 'random walk' tendencies of the population under such circumstances, the simulation is started with equal numbers of 'yes' and 'no' individuals. The conversion likelihood is equal in both directions, and 1000 iterations are used. Of 5 populations of 500 individuals

beginning with 50% in each behavioural category, after 1000 iterations, 3 of the 5 populations have become fixed for the irrational behaviour and 2 have lost the irrational behaviour entirely.



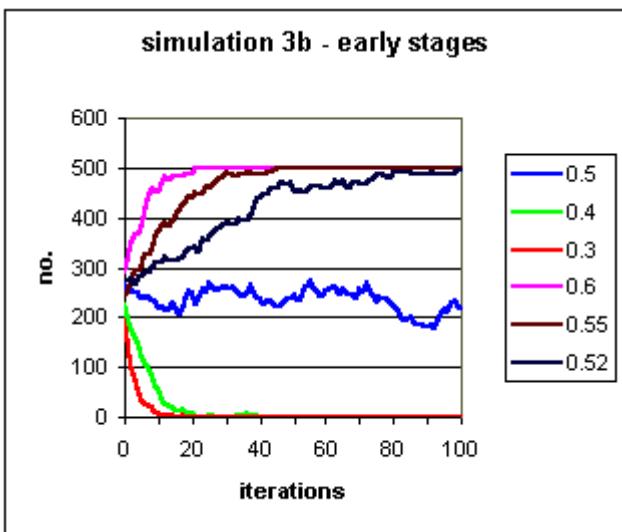
**Figure 2:** 5 runs of a simulation in which the likelihood of conversion to a trait and conversion away from a trait are equal in each contact. In a population of 500, each starting with 250 individuals exhibiting the trait, 3 populations become fixed for the trait (blue, brown and turquoise) and the trait becomes extinct in 2 (green and pink). These are the most trivial examples of random walks, and like the S-shaped curves in Figure 1 are given for the purposes of reference.

Random walks are totally random, by definition, but may appear directional. Take for example the pink population in Fig. 2. After a period of quasi-periodic movement upwards, it suddenly turns and declines to zero. As this is an agent simulation, it is known that this is simply a random walk. However, an empirical worker approaching evidence in the other direction might easily conclude, wrongly, that there must be non-random influences on the trend. In a random walk, at  $p = 0.5$  as given here, any trajectory is equally likely and the odds of any trajectory occurring are  $0.5^n$  (i.e.  $p^n$ ) for any particular trajectory of  $n$  steps. Since  $n$  is 1000 in simulation 3, there are approximately  $10^{301}$  different possible trajectories, all equally likely, and each of which has an approximate probability of  $10^{-301}$ . Over an infinite number of  $n$ , all trajectories will eventually terminate in fixation or extinction, with equal probability (Kimura & Ohta 1978, Kimura 1979). The sensitivity of such a situation to variation in  $p$  may be illustrated by considering what happens when  $p$  deviates from 0.5. Effectively, even small deviations away from 0.5 will pull the random walk rapidly to fixation or extinction.



**Figure 2b:** 5 runs of a simulation in which the likelihood of conversion to a trait and conversion away from a trait are equal in each contact. Again a population of 500 is used, starting with 250 individuals exhibiting the trait. Varying values of  $p$ , the probability of conversion to the trait, are given.

The rapidity of fixation and extinction is almost too fast to see over a timescale of 1000 iterations in Fig. 2b, so the early stages of the 5 simulations are shown again in Fig. 2c. Where  $p \leq 0.4$  the trait barely survives 20 generations.



**Figure 2c:** The early stages of the simulations given in Fig. 2b above.

In general, the probability of a random walk, where  $p$  is not equal to 0.5 are  $p^n + (1-p)^m$ , where  $n$  is the number of incremental turns and  $m$  is the number of decremental turns. Therefore where  $p$  is not equal to 0.5, not all walks are equally probable, and the walks are quasi-random. This much is trivial, once again, but illustrates how easily cultural traits may exhibit chaotic

dynamics, especially if  $p$  is variable from generation to generation (dependent perhaps on other cultural factors). Under such circumstances, a predictive memetics would become virtually impossible, as cultural traits would swing wildly in frequency with apparently no rationale (perhaps a familiar scenario for empirical observers of culture). However, the situation does become more predictable as other factors are added to the simulation.

### 3.4 Simulation 4 - Making contagion dependent on the outcome of an event

So far, the irrational behaviour considered has been purely (and arbitrarily) contagious. Now a complication is introduced in that propagation of the irrational contagious behaviour depends on the perceived success of the mating by the pair involved. As described above, the population is divided into two arbitrary star signs. These are random labels with no function other than to divide the population into two random groups. Therefore any behaviour based on such divisions is by definition irrational, for example, the discussion of star sign with prospective partner. See Lindeman (1998) and Delfabbro & Winefield (2000), for some real life examples of irrationality in event interpretation.

Let the outcome of pairings in each iteration be recorded. Let also the chance of success in each pairing be  $s$  ( $0 \leq s \leq 1$ ), the a priori success probability. A random number generator is used to generate a number between 0 and 1. If this random number is less than  $s$ , the pairing is deemed to have been successful. The following rules then apply:

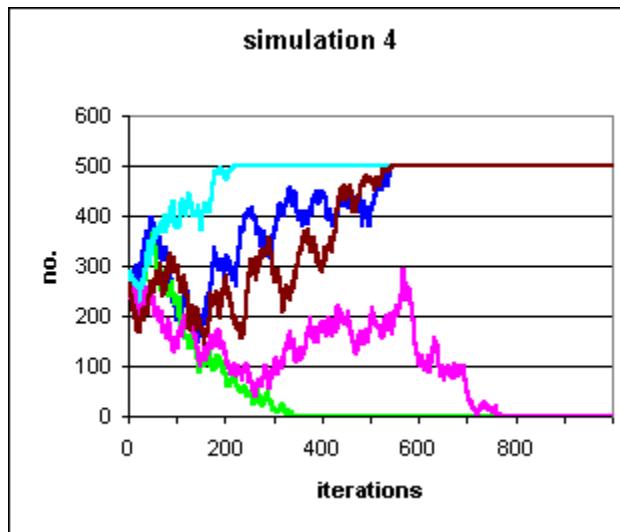
If an 'astrological' individual pairs with a 'non-astrological' individual, and they are of divergent star signs, and the pairing is unsuccessful, the 'non-astrological' individual will convert to 'astrological' (because that is what the irrational hypothesis predicts, and spurious 'evidence' has apparently been provided).

If an 'astrological' individual pairs with a 'non-astrological' individual, and they are of divergent star signs, but the pairing is successful, the converse will ensue, i.e. the 'astrological' individual will convert to 'non-astrological' (because the irrational hypothesis has been falsified in this instance).

If an 'astrological' individual pairs with a 'non-astrological' individual, and they are of compatible star signs, and the pairing is successful, the 'non-astrological' individual will convert to 'astrological' (again because the irrational hypothesis has been spuriously 'verified').  
If an 'astrological' individual pairs with a 'non-astrological' individual, and they are of compatible star signs, but the pairing is unsuccessful, the 'astrological' individual will convert to 'non-astrological' (because the irrational hypothesis has been falsified in this instance).

In other words, the outcome of the event will either be as predicted by the irrational individual or will not be so. The success or failure of the prediction determines the likelihood that one or the other individual will be converted to the behaviour pattern of the other. Note that pairings between two irrational, or two rational, individuals are not considered, as contact is required with an individual of the opposite behavioural type in order to produce behavioural modification. This is what is meant by a purely 'contagionist' mechanism, since contact between individuals is required for 'transmission' of behaviour.

Under such circumstances, it is again, as in simulation 3 above, virtually impossible for the irrational behaviour to colonise the population from a single individual. That is because the rules above are balanced, i.e. the number of events causing a conversion in one direction is equal to the number of events causing conversion in the opposite direction (rules 1 & 3 versus 2 & 4). As in the previous simulation, 5 populations are allowed to run for 1000 iterations with the initial frequency of the irrational behaviour at 50%. The chance of success for each pairing is also set at 50%.



**Figure 3:** 5 populations of 500 individuals, with initially 250 individuals exhibiting the irrational behaviour, and the transmission of the behaviour dependent on the outcome of pairings. Although the dynamics of this simulation are strictly rule-based, and those of simulation 3 dependent on binomial chance, the population trajectories produced by the two systems are indistinguishable.

### 3.5 Simulation 5 - Introducing the 'self-fulfilling prophecy'

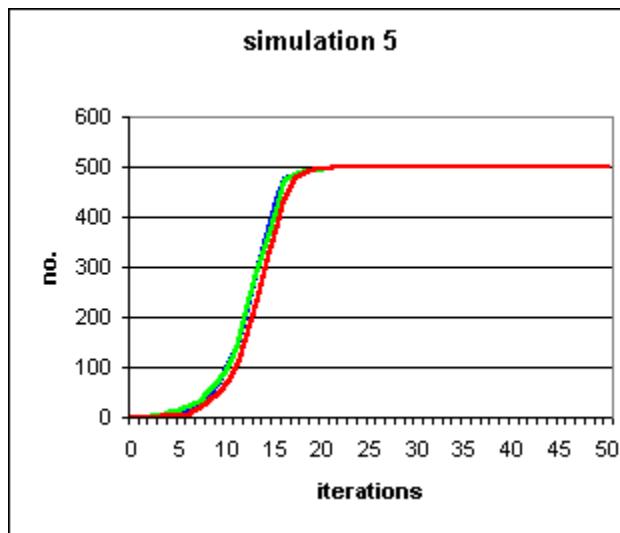
The above example demonstrates that an irrational behaviour, even when based on 'evidence', is likely to follow a random walk within a population. Such a randomly walking trait is very unlikely to be able to colonise a population if there is initially only a single individual manifesting it. However, in simulation 3 above, it is assumed that the success or otherwise of every pairing was due to random chance. On the basis of this random chance, various rules determined the tendency of individuals to convert from one behaviour to the other. Let it now be imagined that the irrational behaviour affects the outcome of the mating (see Downey *et al* 1998 for a real-life example). The rules remain the same, but likelihood of success of a pairing is not equal for all pairs. If those individuals enquiring about the star sign of a prospective partner are less likely to have a successful pairing with individuals of incompatible star signs, and more likely to have successful pairing with those of compatible star signs, then this constitutes a 'self-fulfilling prophecy'.

Bias is defined as the extent to which astrology will interfere with the potential success of a relationship on a scale of 0 to 1, where the pair has incongruent star signs, and one or both of the

pair exhibit the astrology behaviour. This is programmed in the following manner. The a priori likelihood of success of a relationship is  $s$ , and bias is  $b$ , both are between zero and 1, and both are set in advance. A random number generator is again used to generate a number between zero and 1. Without bias, if that random number is less than  $s$ , the pairing is successful. With bias,  $b$  is subtracted from  $s$ . This reduces the likelihood that the random number will be less than  $s$ , and thereby 'loads the dice' against success. Where the pair have congruent star signs, and one or both of the pair exhibit the astrology behaviour, bias acts in the other direction, to increase the chance of success. This is programmed by adding  $b$  to  $s$ , rather than subtracting it as above, and therefore increasing the likelihood of success. Where  $b-s$  is less than zero, the relationship is doomed to failure, regardless of the number produced by the random number generator.

Likewise where  $b+s$  is greater than one, the relationship is guaranteed to succeed.

Where  $b = 0.5$ , populations run to fixation fairly smoothly for most values of  $s$ , provided the novel irrational meme can survive the first few iterations. This 'self-fulfilling prophecy' thus changes the irrational behaviour from a randomly walking trait into a colonizing one. The 'contagionist paradigm' of steady sigmoid curves, reappears.



**Figure 4:** An irrational trait originating in a single individual, colonising a population of 500, running to fixation in under 25 generations in 3 separate runs (red, green and blue). 5 runs of this simulation were performed, of which the other 2 resulted in immediate extinction of the trait at the first iteration. This is clearly an S-shaped curve, demonstrating that, once the novel meme has survived the first few iterations, the 'self-fulfilling prophecy' mechanism turns a random walk into an irresistible tendency. Parameters were  $s=0.5$ ,  $b=0.3$ .

### 3.6 Simulation 6 - Introducing rational behaviour

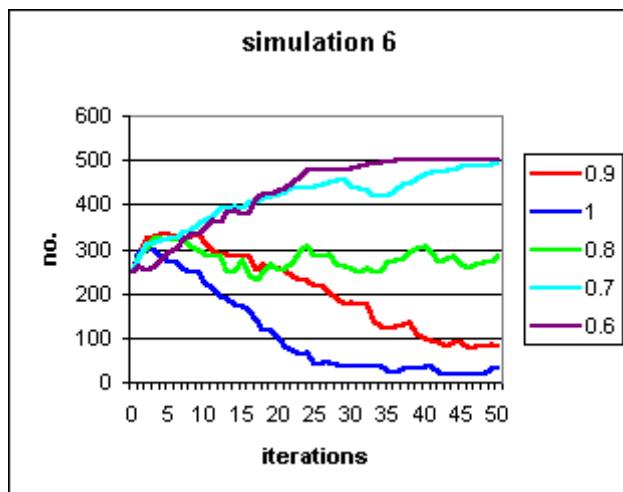
Up to now, only the presence or absence of the irrational astrology-based behaviour has been considered. Now, let individuals also assess potential partners according to the rational criterion of 'status'. It is unimportant how this is defined, except that it is necessary that this be a genuine factor contributing to the success of pairings. Whereas the success of pairings of partners of differing or compatible astrological sign was purely a function of chance alone, or chance coupled to the 'self-fulfilling prophecy' phenomenon, the success of pairings of different status

groups has a fixed contribution to overall pairing success, independent of star sign or any behaviours. Greater pairing success for compatible status groups is therefore objectively true. That is why the status behaviour is regarded as rational.

Status effects can co-exist with irrational bias effects. In the interaction of the 2 strategies, the following general rules apply. The full set of precise rules is available on [http://www.geocities.com/derek\\_gatherer/supp.htm](http://www.geocities.com/derek_gatherer/supp.htm).

- The status of paired agents is compared.
- The star sign of paired agents is compared.
- The rational and irrational behaviours for the paired agents are compared.
- If status is compatible, a priori likelihood of success is doubled (irrespective of whether or not either agent is rational)
- If at least one agent is irrational, the 'self-fulfilling prophecy' phenomenon interferes with the a priori likelihood of success up or down, accordingly.
- The success of the pairing is determined, using the random number generator.
- Based on success or failure, the agent states for the rational and irrational behaviours are altered.
- However, behaviours only change if an agent is paired with an agent exhibiting the appropriate behaviour (contagion still rigidly applied).

These rules allows both rational behaviour (relating to status) and irrational behaviour (relating to star sign) to enter the population. The two are not mutually exclusive and therefore separate memes rather than 'allomemes' (Durham 1991). Which meme propagates faster depends on the relative degrees of bias (the motor of the self-fulfilling prophecy phenomenon for irrational behaviour) and objective status-related increase in the likelihood of successful pairing. Indeed where a priori likelihood of success,  $s$ , is low and bias,  $b$ , is high, the rational behaviour tends to extinction.

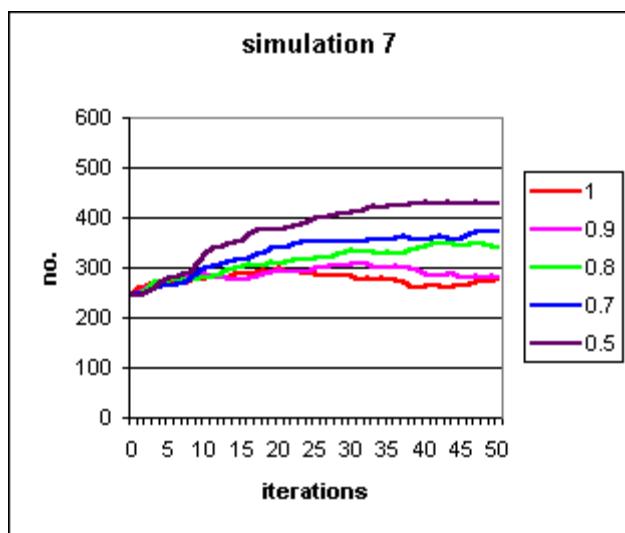


**Figure 5:** Progress to fixation or decline to extinction of rational behaviour under varying conditions of 'self-fulfilling prophecy' bias,  $b$ , from 0.6 to 1, with an a priori likelihood of success of pairings of  $s = 0.5$ . In all runs, the irrational behaviour ran to fixation in 10 iterations.

### 3.7 Simulation 7 - Giving agents the benefit of experience

The previous simulations 3 to 6, have assumed that conversion events are dependent on the success of immediately preceding pairings. Perhaps more realistically in terms of human society, individuals will withhold judgment until several experiences have accumulated. Let the agents now have a threshold,  $t$ . Experiences liable to cause conversion to rationality or irrationality are accumulated, and the appropriate conversion takes place when the threshold is exceeded. The experience counter is then reset to zero. In other words, at a threshold of, say, 3, an agent will convert to rational behaviour only after 3 appropriate contributory experiences.

This decreases the volatility of the system, and promotes the settling of the population settles into a stable state. The equilibrium position is determined by the relative values of  $s$  and  $b$ , and the speed with which it is obtained is dependent on the threshold for conversion,  $t$ .



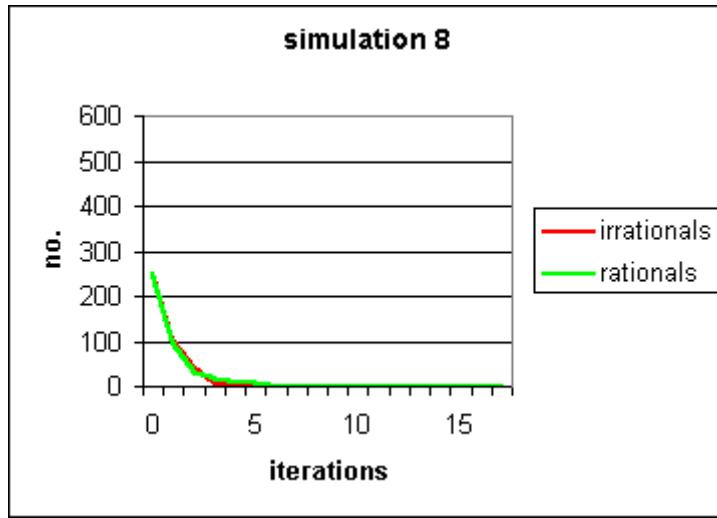
**Figure 6:** Progress of rational behaviour in the population under the same conditions as in figure 5 above ( $s$  constant at 0.5 and  $b$  run at 0.5 to 1), but with a threshold of 3 introduced. This delays fixation of rational behaviour under conditions of low bias, but also prevents its immediate extinction under conditions of high bias.

### 3.8 Simulation 8 - Removing successful pairs

All the previous simulations recycle the same individuals in a vulgar imitation of Schnitzler (1982). If successful pairs no longer seek further mates, as seems more realistic in human terms, let successful pairs now be replaced in the population by naïve individuals. Note that this is done based on simulation 6 - with possibility of individual expressing both rational and irrational attitudes, and without allowing any accounting of previous experience (naïve individuals have, by definition, no experience).

Where the a priori likelihood of success,  $s$ , is high, individuals are removed so rapidly from the population that neither the rational nor irrational behaviour is very persistent. Lowering  $s$  to 0.01,

and with high self-fulfilling prophecy bias,  $b$ , it is possible to retain them for some longer time. However, the influx of new individuals wins out eventually.



**Figure 7:** Replacement of successful pairs with naïve individuals rapidly removes both behaviours, rational and irrational, from the population. Compare with Figure 4. The same parameters are used,  $s = 0.5$ ,  $b = 0.3$ , which would normally generate a powerful "self-fulfilling prophecy" for the fixation of irrational behaviour. However, here the arrival of waves of naïve agents overwhelms the capacity of the contagion system.

### 3.9 Simulation 9 - Removing successful pairs from a population with memory

Simulation 7 showed that allowing agents the benefit of experience before changing strategy, caused changes to be subdued, both up and down. Simulation 8 showed that removal of successful pairs extinguishes both behaviours. This simulation combines 7 and 8 to see if memory can maintain the behaviours in such a population. However, the dynamics are the same as in simulation 8 (data not shown). This is only to be expected, as when individuals rely on their own experience, but cannot share it with the rest of the population, that experience is no longer available when the individual leaves the population.

### 3.10 Simulation 10 - Introducing a communal cultural information pool influencing conversion decisions

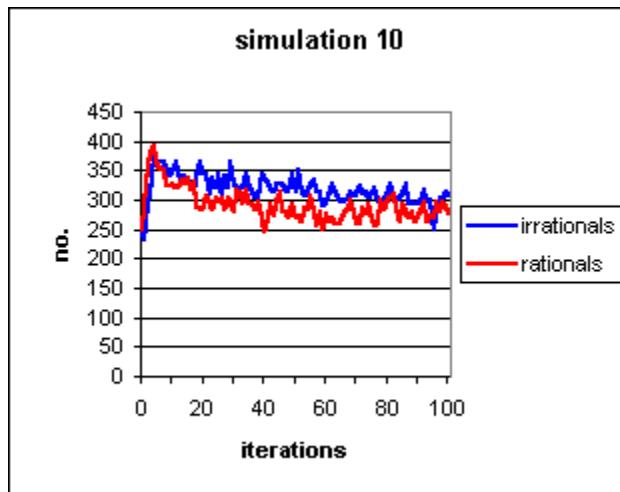
All previous simulations 1 to 9 have only allowed individuals to change their behaviour based on the behaviour of individuals with whom they come in contact. These models have thus been strictly contagionist models. Let the individuals now draw on a communal store of previous experience. This communal reference is built by recording the successes of pairings of different types. This is slightly different to simulation 8, in that there individuals were allowed to make decisions on the basis of their own experience only. Now recourse is permitted to knowledge of the experiences of other individuals. Individuals can convert without contact with another individual, if the cultural assessment is that rational or irrational behaviour is correct. This is similar, but not exactly identical to, the 'conformist transmission' of Henrich & Boyd (1998).

Thus rules 7 and 8 from simulation 6, above, are now altered to:

- Based on success or failure, a contribution to the communal information pool is made. The agent states for the rational and irrational behaviours are still altered in the usual way if an agent is paired with an agent exhibiting the appropriate behaviour.
- 8. But now additionally, behaviours can change even if an agent is not paired with an agent exhibiting the appropriate behaviour, subject to consulting the communal information pool for the latest consensus.

Thus the communal information pool does not replace contagion, but supplements it. This simulates an element of 'persuasion'. Agents still convert under the influence of other agents, both rationally and irrationally, if conditions are correct. However, where no persuasion is exercised, agents can still alter their behaviour if conditions are correct, and the cultural norm suggests they should. (Perhaps this violates Levy & Nail (1993)'s requirement for an absence of intentionality in a true contagion system, but I am unsure of what intentionality means in terms of software agents.)

This helps to maintain steady random walks very near the starting values, after an initial movement upwards or downwards (dependent on starting parameters) as the contagion effect begins, before the communal information pool helps to correct it.

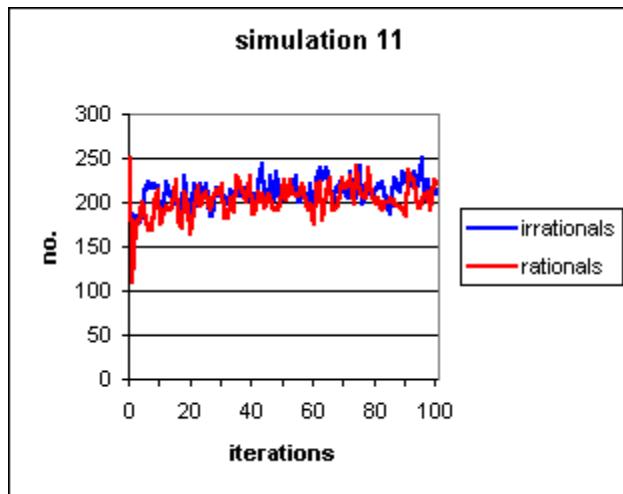


**Figure 8:** Non-contagion-dependent cultural system. The self-fulfilling prophecy maintains some degree of irrationality, but the ability of agents to convert away from irrationality without contact with a rational agent prevents it from running to fixation. Again,  $s=0.5$ ,  $b=0.3$ .

Note that in Fig.8 the starting parameters are  $s=0.5$ ,  $b=0.3$ . By reference to Fig. 5, it can be seen that these would normally cause both memes to run to fixation. This is indeed the initial tendency (note the initial rapid upsurge in both blue and red lines), but the communal information then begins to pull the meme frequencies down towards a quasi-random walk around an equilibrium level.

### 3.11 Simulation 11- Cultural information pool with replacement of successful pairs

This is a combination of simulation 10 with simulation 8. In simulation 8 both behaviours were lost as naïve individuals took over the system. However, with the common cultural information pool, the behaviours persist.



**Figure 9:** Same conditions as simulation 10, but allowing replacement of successful pairs. The initial dip is caused by the arrival of the first set of naïve agents in the first cycle or so before the cultural system has adequate data concerning the likely success of various pairs.

As in simulation 10, the cultural pool requires a few iterations to establish its influence, this time after a very short initial spell (almost invisible on Fig. 9) where the downward tendency of simulation 8 is repeated. This precipitate drop recovers almost immediately as the common cultural pool begins to exert its influence.

## 4 Discussion

The main conclusions of the paper are that:

- The 'contagionist paradigm', i.e. the common sigmoid curve representation of a memetic epidemic, is a very special case, requiring a fairly contrived set of system parameters in order to be produced.
- The 'random walk', i.e. an apparently stochastic meandering of meme frequency over time, is the more likely situation, even when the underlying parameters are far from random - a frequent variation is a pseudo-random walk around an equilibrium level.
- A population with a high turnover of agents (and hence a high proportion of naïve agents at any one given time) cannot maintain either of the described behaviours, rational or irrational, i.e. those meme frequencies drop to zero, without recourse to the use of a cultural information pool.

Many additional variables could be added to the simulation. For instance, Ophir (1998) gives his agents age attributes, mates them at age 30, and has them expire at a random age between 60 and 100, after producing a single family of progeny. This allows discrete iterations to be replaced with more realistic overlapping generations. However, Ophir's agents are monogamous, and reproduce only once, the emphasis of Ophir's work being the distribution of economic resources rather than mate pairing behaviour. Also, Pearson & Boudarel (2001) have recently devised a matrix based system for agent pair interactions, which could be incorporated into the pairing process described here.

The strictly contagionist person-to-person replication of both the rational and irrational behaviour is shown by simulation 8 to be unstable in the face of removal of individuals from the population. Indeed, even when all individuals are retained in the population, irrational behaviour requires the 'self-fulfilling prophecy' phenomenon to spread through the population, as illustrated in the difference between simulations 4 and 5. Self-fulfilling prophecies have been empirically demonstrated in human mate pairing (involving 'rejection-sensitivity' rather than astrology, but the basic principle is the same - see Downey *et al* 1998).

The only way to maintain the irrational behaviour in the population in the face of high levels of replacement is to use some kind of common information pool system. This, however, requires that the agents choose to refer to the information system for guidance. If, for some reason, the population in simulation 11 neglects to maintain the information system, the dynamics of simulation 8 will begin to take over and both memes will slide towards extinction.

It is clear that a 'random walk' pattern of meme incidence over the iterations is more common than the 'contagionist paradigm' sigmoid curve, which only occurs in rather 'pure' situations.

Random walk-like effects may be genuinely random because of the stochastic nature of the system (simulation 3), because of balanced rules in a deterministic system (simulation 4), or be anchored pseudo-random walks because of fluctuations around an equilibrium (simulations 10 and 11). Good 'epidemiological' sigmoid curves in culture seem to require either arbitrary contagiousness (simulations 1 and 2), or a powerful selective force (the unopposed self-fulfilling prophecy of simulation 5). Some aspects of human life may be arbitrarily contagious, for example crowd hysteria (see reviews by Levy & Nail 1993; Marsden 1998). Several authors have worked within the 'contagionist paradigm', producing models of traits that spread in an epidemic manner. For instance, Takahasi (1998, 1999) refers to 'biased cultural transmission', Cavalli-Sforza and Feldman (1984) to 'cultural selection', and Sperber (1985, 1996) to the 'epidemiology of representations'. In these cases, it is taken for granted that the epidemic trait is either arbitrarily contagious or, in the case of Cavalli-Sforza and Feldman's 'cultural selection', that some powerful, and usually unspecified, psychological factor is promoting the spread of the trait. The self-fulfilling prophecy phenomenon of simulation 5 would presumably therefore be a case of 'cultural selection'. Another example might be 'prestige-based transmission' (Heinrich 2001) where certain individuals are far more widely copied than others, or 'emotional selection' (Heath *et al* 2001). Blackmore (1999) gives extensive further examples. Cavalli-Sforza and Feldman (1984) also refer to 'natural selection' on culture, where a cultural trait spreads through genuine adaptiveness in the absence of either arbitrary 'biased cultural transmission' or psychological 'cultural selection'. Powerful natural selective forces of this kind do indeed also

operate in culture, resulting in some sigmoid incidence curves for possession of, in modern times, mobile phones and other useful paraphernalia (Rogers 1995), or, in the Bronze Age, for bronze knives (Renfrew 1987). Hewlett and Cavalli-Sforza (1983) describe an epidemic spread of the use of the crossbow among Pygmies, conferring a great hunting advantage over the traditional bow and arrow.

Granted that cultural epidemics can, in theory, occur where there is natural or cultural selection (and certainly do occur for cases of natural selection), the question here is: to what extent could the contagionist paradigm be applied to (non-hysterical) irrationality? If most irrationality is disadvantageous (and therefore unlikely to be 'naturally selected', sensu Cavalli-Sforza & Feldman, as an adaptation), and it is not the result of hysteria (and therefore not likely to have greatly biased transmission over a short time period), and the human society concerned has ample recourse to a pool of experience of previous members, then the spread of irrationality in an epidemic manner would seem to be unlikely. The contagionist paradigm therefore seems to be an inadequate model for the long term persistence of irrational behaviour in human populations.

For instance, in a stock market system, agents could predict the future price of shares, and then bid for them irrationally (perhaps again based on astrology) or rationally (based on profit indicators). If the purchase of shares causes their price to be elevated, this constitutes a kind of 'self-fulfilling prophecy' phenomenon. This would tend to result in irrational bidding spreading throughout the entire investment community. The fact that this does not happen too often, is perhaps due to there being a well established communal information pool about the success of various previous investments, on which current investment decisions can be based. As in simulation 10, this would prevent even a strong 'self-fulfilling prophecy' of irrational share inflation from taking over entirely. Interestingly, Caginalp et al (2001) have identified lack of open factual information as a contributory element in 'bubbles' created during lab simulations of stock market trading. Such a situation would promote information exchange on a person to person basis and therefore lead to a greater susceptibility to contagion effects. Study of real markets has also led to the conclusion that the "level and nature of information available to dealers, and social communication networks" are important (reviewed in Marsden 1998, section 2). Empirical data shows that random walking does tend to predominate over long periods of time (Malkiel 1985). It is also interesting that although the common information pool helps to prevent irrationality running to fixation, it does not eliminate it. Just as evolution in financial markets does not lead to maximal market efficiency (Frank 1999), evolution in the mating game, or in any other area of human conduct, does not lead to maximal rationality. Heinrich (2001b) presents an interesting example of the persistence of bottle-feeding in parts of the world where breast-feeding would result in lower infant mortality. One might also cite Rogers and Shoemaker's classic example of the determination of Peruvian villagers not to boil drinking water despite extensive health education and the established use of boiling for other food purposes (Rogers & Shoemaker 1972), or recent discussions of 'altruistic punishment' (Fehr & Gachter 2002) and the revival of the 'tragedy of the commons' debate (Ostrom *et al.* 1999, Heinrich *et al* 2001a). Heinrich (2001b) finds that the persistence of maladaptive behaviour is most likely when: a) there is high probability that environmental information is inconclusive, and b) there is biased transmission of the maladaptive trait. This is entirely consonant with the simulations presented here. Interestingly and rather controversially, Heinrich goes a step further and suggests that even adaptive traits require biased transmission in order to produce S-shaped

curves. This is a radical departure from the fundamental theoretical observation that natural selection of an advantageous trait is analogous in form to an epidemic, and produces an S-shaped curve (Fisher 1930). It is not appropriate to discuss this any further here other than to note that the present observations concerning the cultural information pool and the relative rarity of situations in which the contagionist paradigm applies, are consonant with the first part of Henrich's (2001b) thesis.

One point not covered in the simulations is the issue of spatial heterogeneity within the population. The simulations could be modified to give each individual a geographical location. Partner choice, and the extent of the cultural information pool, could thus be limited by geographical isolation. One might also permit the high status individuals greater geographical mobility than those of the low status. Multi-generational effects might allow some of the successfully mated individuals to produce offspring for reintroduction into the population at a later date. Incorporation of a genetic component might also be permitted, with some individuals being inherently more or less sceptical than others.

In summary, the contagionist paradigm is a special case scenario, only applicable to situations of:

- hysteria (Levy and Nail 1993),
- strong psychological distortion of transmission mechanisms (Cavalli-Sforza & Feldman 1981; Takahasi 1998, 1999), or
- powerful natural selection of an advantageous cultural trait (Cavalli-Sforza & Feldman 1981).

According to Henrich (2001b), the last of these three may also be suspect unless 2) also applies. Since 1) is rare, and 2) is often difficult to identify empirically (and in the worst cases is often no more than a post hoc explanation when no other can be found), it is submitted that contagionism has limited range as an explanatory tool of human culture.

## References

- Blackmore S (1999) *The Meme Machine*. Oxford University Press, Oxford.
- Boyd R & Richerson PJ (1985). *Culture and the Evolutionary Process*. The University of Chicago Press, Chicago.
- Brown PS, Humm RD & Fischer RB (1988) The influence of a male's dominance status on female choice in Syrian hamsters. *Hormonal Behavior* **22**, 143-149.
- Caginalp G, Porter D & Smith V (2001) Financial Bubbles: Excess Cash, Momentum, and Incomplete Information. *Journal of Psychology and Financial Markets* **2**.  
([http://www.psychologyandmarkets.org/journals/abstracts/issue\\_v2\\_2.html](http://www.psychologyandmarkets.org/journals/abstracts/issue_v2_2.html)).

Cavalli-Sforza LL & Feldman MW (1981) Cultural Transmission and Evolution. Princeton University Press, Princeton.

Darwin C (1871) The Descent of Man, and Selection in Relation to Sex, 2 volumes. John Murray, London.

Delfabbro PH & Winefield AH (2000) Predictors of irrational thinking in regular slot machine gamblers. *Journal of Psychology* **134**, 117-128.

Doran J (1998) Simulating Collective Misbelief. *Journal of Artificial Societies and Social Simulation* **1** (<http://www.soc.surrey.ac.uk/JASSS/1/1/3.html>).

Downey G, Freitas AL, Michaelis B & Khouri H (1998) The self-fulfilling prophecy in close relationships: rejection sensitivity and rejection by romantic partners. *Journal of Personality and Social Psychology* **75**, 545-560.

Durham WH (1991) Coevolution: Genes, Culture and Human Diversity. Stanford University Press, Stanford.

Edmonds B (1998) On Modelling in Memetics. *Journal of Memetics - Evolutionary Models of Information Transmission* **2** ([http://jom-emit.cfpm.org/1998/vol2/edmonds\\_b.html](http://jom-emit.cfpm.org/1998/vol2/edmonds_b.html)).

Fehr E & Gachter S (2002) Altruistic punishment in humans. *Nature* **415**, 137-40

Fisher RA (1930) The Genetical Theory of Natural Selection. Clarendon Press, Oxford.

Frank J (1999) Applying Memetics to Financial Markets: Do Markets Evolve towards Efficiency? *Journal of Memetics - Evolutionary Models of Information Transmission* **3** ([http://jom-emit.cfpm.org/1999/vol3/frank\\_j.html](http://jom-emit.cfpm.org/1999/vol3/frank_j.html)).

Freeberg TM, Duncan SD, Kast TL & Enstrom DA (1999) Cultural influences on female mate choice: an experimental test in cowbirds, *Molothrus ater*. *Animal Behavior* **57**, 421-426.

Goodkind DM (1993) New zodiacal influences on Chinese family formation: Taiwan, 1976. *Demography* **30**, 127-42.

Heath C, Bell C & Sternberg E (2001) Emotional selection in memes: the case of urban legends. *Journal of Personality and Social Psychology* **81**, 1028-1041.

Henrich J & Boyd R (1998) The evolution of conformist transmission and the emergence of between-group differences. *Evolution and Human Behavior* **19**, 215-241.

Henrich J, Boyd R, Bowles S, Camerer C, Fehr E, Gintis H & McElreath R (2001a) In Search of Homo Economicus: Behavioral Experiments in 15 Small-Scale Societies. *American Economic Review* **91**, 73-78.

Henrich J (2001b) Cultural Transmission and the Diffusion of Innovations: Adoption Dynamics Indicate That Biased Cultural Transmission Is the Predominate Force in Behavioral Change. *American Anthropologist* **103**, 992-1013.

Hewlett BS & Cavalli-Sforza LL (1983) Cultural transmission among Aka pygmies. *American Anthropologist* **88**, 922-934.

Kimura M (1979) The neutral theory of molecular evolution. *Scientific American* **241**, 98-108.

Kimura M & Ohta T (1978) Stepwise mutation model and distribution of allelic frequencies in a finite population. *Proceedings of the National Academy of Sciences USA* **75**, 2868-2872.

Laland KN (1994) Sexual selection with a culturally transmitted mating preference. *Theoretical Population Biology* **45**, 1-15.

Levy DA & Nail PR (1993) Contagion: a theoretical and empirical review and reconceptualization. *Genetic, Social and General Psychology Monographs* **119**, 233-284.

Lindeman M (1998) Motivation, cognition and pseudoscience. *Scandinavian Journal of Psychology* **39**, 257-265.

Malkiel BG (1985) *A Random Walk Down Wall Street*. 4<sup>th</sup> ed. Norton, New York.

Marsden P (1998). Memetics and Social Contagion: Two Sides of the Same Coin? *Journal of Memetics - Evolutionary Models of Information Transmission* **2** ([http://jom-emit.cfpm.org/1998/vol2/marsden\\_p.html](http://jom-emit.cfpm.org/1998/vol2/marsden_p.html)).

Ophir S (1998) Simulating Ideologies. *Journal of Artificial Societies and Social Simulation* **1** (<http://www.soc.surrey.ac.uk/JASSS/1/4/5.html>).

Ostrom E, Burger J, Field CB, Norgaard RB & Policansky D (1999) Revisiting the commons: local lessons, global challenges. *Science* **284**, 278-282.

Pearson DW & Boudarel M-R (2001) Pair Interactions : Real and Perceived Attitudes. *Journal of Artificial Societies and Social Simulation* **4** (<http://www.soc.surrey.ac.uk/JASSS/4/4/4.html>).

Pizzari T (2001) Indirect partner choice through manipulation of male behaviour by female fowl, *Gallus gallus domesticus*. *Proceedings of the Royal Society of London Series B, Biological Sciences* **268**, 181-186.

Rashevsky N (1949) Mathematical biology of social behaviour I. *Bulletin of Mathematical Biophysics* **11**, 105-113.

Renfrew C (1987) *Archaeology & Language: The Puzzle of the Indo-European Origins*. Jonathan Cape, London

Rogers EM & Shoemaker FF (1971) Communication of Innovations. A Cross-cultural Approach  
2<sup>nd</sup> ed. Free Press, New York.

Rogers EM (1995) The Diffusion of Innovations 4<sup>th</sup> ed. The Free Press, New York.

Schnitzler A (1982) The round dance and other plays. Arthur Schnitzler. trans. and intro.  
Osborne, C. Carcanet New Press, Manchester.

Sperber D (1985) Anthropology and psychology: towards an epidemiology of representations.  
Man, 73-89.

Sperber D (1996) Explaining Culture: A Naturalistic Approach. Blackwell, Oxford.

Takahasi K (1998) Evolution of transmission bias in cultural inheritance. Journal of Theoretical  
Biology **190**, 147-59.

Takahasi K (1999) Theoretical aspects of the mode of transmission in cultural inheritance.  
Theoretical Population Biology **55**, 208-25

Wade MJ & Pruett-Jones SG (1990) Female copying increases the variance in male mating  
success. Proceedings of the National Academey of Sciences USA **87**, 5749-5753.

## Acknowledgment

The author thanks the staff of the Sid Martin BDI, Alachua, and the Univ. of Florida Library for  
their kind hospitality and assistance during his stay.

# FINDING A NICHE FOR MEMETICS IN THE 21<sup>ST</sup> CENTURY

By  
Derek Gatherer  
[dgatherer@talk21.com](mailto:dgatherer@talk21.com)

Many scientific disciplines have a ‘paradigmatic experiment’, a simple example of how the discipline is applied to a problem. Usually, the paradigmatic experiment is like a light that, if switched on, suddenly illuminates the rest of the field. Mendel’s pea crosses are the paradigmatic experiment for classical genetics. Sometimes, for disciplines where direct experimentation is a little harder, there is not so much a paradigmatic experiment as a paradigmatic example. For instance, the body of data relating to the distribution of sickle cell anemia is a paradigmatic example of the process of molecular evolution in human populations. Such paradigmatic instances are often found in textbooks, and with good reason, given the purpose of textbooks. Sometimes they also constitute the founding experiment in the field, but they need not do so; Mendel’s pea crosses were the founding work in genetics, but the sickle cell anemia story only gradually emerged from a large body of research on biochemical variation in human populations.

So what is the paradigmatic example in memetics? My own favourite is the ‘homebodies vs. the hellraisers’ model used as an introductory vehicle by Boyd & Richerson (1985). Homebodies raise large families of more homebodies. Hellraisers by contrast, tend to have fewer progeny, themselves also of a hellraising tendency. The hellraisers would be heading straight for the evolutionary graveyard, if it were not for one factor: put a homebody and a hellraiser together for any length of time, and the homebody can convert into a hellraiser. Thus, depending on the frequency and strength of social interaction, and the relative reproductive differential of the two groups, hellraising can persist in an often fragile equilibrium.

This simple model is attractive because it encapsulates so many of the themes of the memetic approach to cultural evolution. Notice that the nature of the tendency to homeliness or hellraising is left open. It is possible to conceive at least part of the model in genetic terms: each group has progeny of its own kind because of an innate disposition. Alternatively, one might see homeliness or hellraising as purely cultural attributes, in which case the tendency of progeny to acquire the same phenotype as their parents is merely due to prolonged contact. What is certain is that the conversion of homebodies into hellraisers must be purely cultural. Thus the model contains what Cavalli-Sforza & Feldman (1981) defined as ‘horizontal’ and ‘vertical’ transmission of traits. The unidirectional nature of the horizontal transmission also illustrates ‘cultural selection’. The greater fecundity of homebodies illustrates ‘natural selection’. By subtly varying the parameters of each of these processes, Boyd & Richerson (1985) display several equilibrium states.

I call this a paradigmatic example, because if one can fully understand this simple model, and appreciate its sensitivity to parameter changes, the essence of most current work in cultural evolution and gene-culture co-evolution can be grasped. The problems also become apparent. Just what is the relevance of such a model to the real world? For those who take a co-

evolutionary approach, i.e., those who try to see such problems as involving genes as well as memes, there remains the problem of how, at a mechanistic level, genes influence behaviour. Evolutionary biology and sociobiology are not yet fully integrated into molecular genetics. To express it in Aristotelian terms, genes may, for some, provide a convincing ‘material cause’ for behaviour, but they have not yet been directly demonstrated to provide an ‘efficient cause’. Even for those who are happy to see the example as one that involves exclusively cultural traits, there remains the problem of how to ascertain the types for the individuals. In the model, the phenotypes are defined solely in terms of the tendency to have a certain number of progeny. In real life, the number of progeny is a normal distribution, and discrete categorization may be difficult. A further problem is why certain traits are culturally selected over others. Even if one abandons that concern to the psychologists, and simply accepts cultural selection as a given, what grounds are there for assuming that it will remain at a constant level, or not even go into reverse if other conditions change?

Because of these kinds of difficulties, the main stage for memetics research has been on the computer rather than in the wild. Software objects, or ‘agents’, can be programmed to behave in ways that make it possible to investigate the basic propositions of memetics. They can be given, in the jargon of object-oriented programming, attributes which are discrete values, and methods of manipulating those attributes, and the attributes of their fellows, in ways that correspond to strict rules set by the programmer. The homebodies and hellraisers of the model can be exactly reconstituted as software objects, with their parameters of reproduction and interaction precisely defined. The programmer can then easily expand the model in various ways; perhaps the external environment in which the agents exist can be partitioned into regions, each agent allowed a geographical location, and the parameters of reproduction and interaction varied in different regions. The agents might then be allowed to move, to seek out others of their own type as neighbours, or to culturally interact over long distances. One can even introduce elements of game theory into their actions, to allow them to negotiate and to deceive each other. When compared with the messy reality of human social life - difficulties of accurate categorization of individuals, uncertainty about what is cultural and what is genetic, even a lack of agreement about what culture is (eg. see Kuper 2000) - it is easy to see the appeal of artificial life as a laboratory for memeticists.

However, not everything is orderly in cyber-societies. Each small modification to the basic model adds another layer of complexity, and it is often easy to find oneself with a model that exhibits behaviour that one does not understand. Artificial societies can be just as perplexing as real ones. In the real world, we don’t completely understand individuals, so we are not particularly surprised when we don’t understand the societies they form. But in the artificial world, we do understand the individuals. We understand them perfectly, since we designed them. Yet, put them into social groups and inexplicable things can start to happen.

These caveats are not just peculiar to the creation of memetic models in artificial societies, but are well recognised in the larger social simulation community. They do not constitute an objection to the method in principle, but merely a warning against placing too much faith in it. So, in summary, what methodology can be used for memetics in the 21<sup>st</sup> century? I would recommend the following.

Find a problem where there is already a good body of quantitative data. It is impracticable for memeticists to contemplate collecting our own empirical data. Sociologists, anthropologists and experimental psychologists are better qualified to do the fieldwork, and we should leave it to them.

Search the literature to find out what has already been written on that problem. In particular, try to find any suggestions that processes relevant to memetics have played a part. For instance, what do social psychologists have to say on the matter? Has an evolutionary system of some kind previously been postulated?

Try to formulate an object-oriented analysis (OOA) of the problem. In particular, think about how the quantitative data relates to the agents in the OOA. How would you set up the system so that an analogous body of data could be generated?

Code a prototype of the system. Make sure you have clearly defined the social attributes of each of your agent objects and, if genetic attributes are also required, that you make a clear distinction between what is social transmission and what is genetic transmission. If there has previously been an analysis of the problem by social psychologists, try to incorporate some of the postulated social psychological mechanisms into the simulation.

Choose a starting set of parameters and run the simulation. Compare the data from your artificial society with the empirical data set in the literature. Vary the starting parameters and see how robust the outcome is to such variation.

If your artificial society has adequately reproduced what is seen in the real world, then you are justified in declaring that a set of agents with the set of attributes and methods that you have created, within the range of parameters that you have defined, is sufficient to achieve that outcome. Therefore, it could be that the mechanism in the real world is similar. If the outcomes in the artificial society are completely different to those in the real data set, then you may declare that your system is insufficient to reproduce the real world situation. In that case, the mechanism in the real world is likely to be different to that in your artificial society.

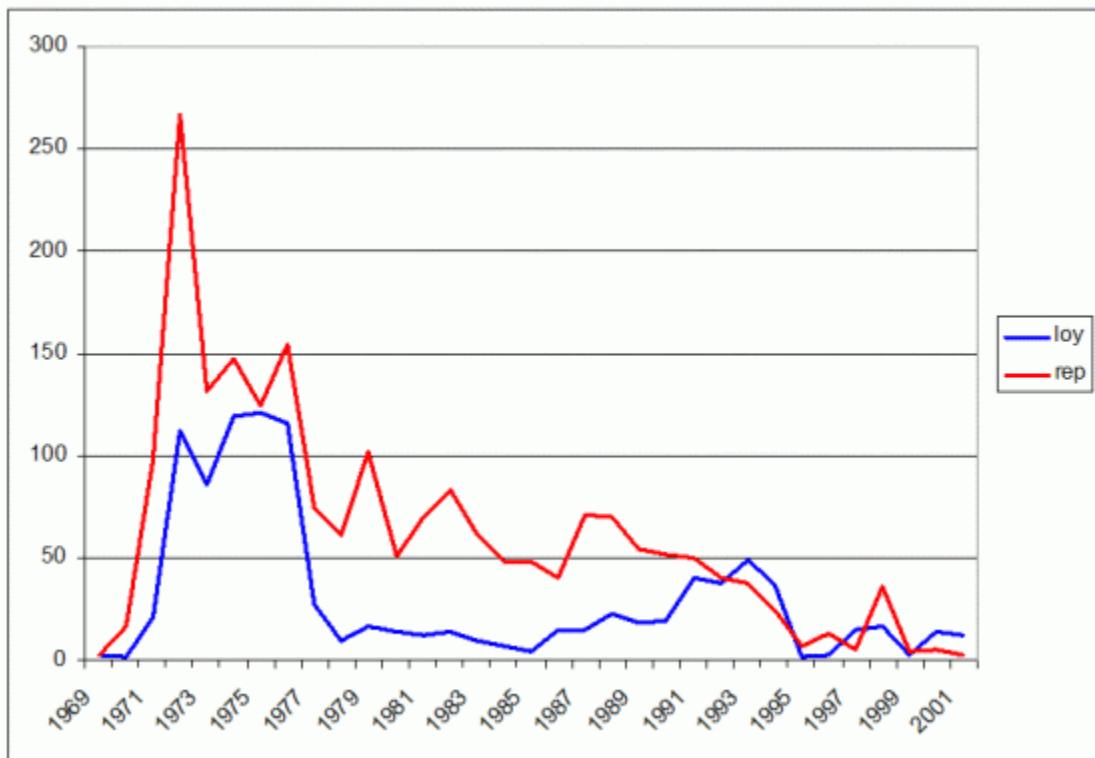
Since positive results only give indications that the real world could be as you have modeled it, whereas negative results give much stronger indications that the real world is not the same as the model, it seems that this methodology is actually quite Popperian. The function of the memeticist is to cast doubt on theories rather than to confirm them.

Is this not however, just an argument in favour of casting off the label of memeticists, and plunging into the social simulation community? Not quite; I propose that memeticists can form a recognizable sub-community by our adherence to the following requirements:

- **Always try to approach problems with a memetic or gene-meme co-evolutionary model in mind.**

- Always work on problems which already have a good body of empirical data, preferably quantitative, against which one can falsify the theories that have been used to build the software agent models.

What are those areas which have reasonable bodies of quantitative data? One obvious one is politics or any situation where human choices are recorded by ballots (e.g. Gatherer 2004). Another one is anthropology where there are extensive ethnographic databases (e.g. Gatherer 2002b). A third is demography where censuses of population movement are available. There are many more. The best way to find one is to ask a social scientist, geographer or historian. As well as my two published efforts in this direction (Gatherer 2002b, 2004), let me give one other preliminary example before finishing. There has been a lot of interest within the memetics community in the issue of terrorism (e.g. Marsden 2001; Lynch 2002). It was even briefly alluded to by Dawkins (1976). Memeticists often seem convinced that this is an area where the approach can make some kind of contribution. There is also a wealth of literature on the subject from other academic disciplines, including social simulation (Raczynski 2004; Eidelson & Lustick 2004), and terrorist behaviour has been variously assessed as a form of reactive warfare, organised crime, or psychopathology. There is, in at least one case, an impressive body of quantitative data: the CAIN database on the conflict in Northern Ireland from 1969-2001, available from the University of Ulster (<http://cain.ulst.ac.uk>). This database contains details on every politically motivated killing (over 3600 of them) during the 32 year period of its coverage, as well as data on the political and economic background. Figure 1 shows the number of murders attributed per year to Loyalist and Republican paramilitaries.



**Figure 1:** Number of killings per year attributed to Loyalist and Republican paramilitary organizations from 1969 to 2001

How do existing memetic models of terrorism stand up against such a data pattern? Is it possible to design a set of agents who would attack each other in such a way? Alternatively, is it possible to design sets of agents based on existing models who then fail to attack each other in such a way, thus casting doubt upon the model? Lynch's model (2002) only really goes as far as describing a variety of reasons, of varying degrees of cogency, why terrorism is likely to spread. One might say that Lynch is simply positing that terrorism is, in the standard terminology, culturally selected. Unfortunately, under simulation conditions, a model that simplistic merely leads to terrorism running to fixation within the population under the dynamics of a sigmoid curve (the "contagionist paradigm", see Gatherer 2002a). In effect, the agents simply kill each other off completely. The data in Figure 1 require dynamics that would explain the waxing and waning of the Republican kill rate at approximately 2-year intervals, at least during the period 1972-1989, while the opposing Loyalist kill rate was relatively unchanged year-on-year from 1979-1991. Looking not just at the overall distribution of murders, but also at their individual timing relative to each other, using a Run Test, suggests that a tit-for-tat mechanism may have been in place at some phases (data not shown). Thus a good memetic model for terrorism would not merely concentrate on trying to explain cultural selection, but would also require some game theory to be incorporated.

In summary, memetics can survive as a data-oriented branch of the social simulation field specializing in the evaluation of co-evolutionary or cultural models. Most of that evaluation will be negative, as the nature of the real world can never be confirmed using a computer model. Nevertheless, it may help us to identify cases where our a priori thinking about a cultural phenomenon is inadequate. Karl Popper would have been pleased with this.

## References:

- Boyd R & Richerson PJ (1985) Culture and the Evolutionary Process, The University of Chicago Press.
- Cavalli-Sforza LL & Feldman MW (1981) Cultural Transmission and Evolution. Princeton University Press, Princeton.
- Dawkins, R. (1976) The Selfish Gene. Paperback edition 1978. Granada Paladin: London.
- Eidelson BM & Lustick I (2004) VIR-POX: An Agent-Based Analysis of Smallpox Preparedness and Response Policy. Journal of Artificial Societies and Social Simulation 7, no. 3 <<http://jasss.soc.surrey.ac.uk/7/3/6.html>>
- Gatherer D (2002a) The Spread of Irrational Behaviours by Contagion: An Agent Micro-Simulation. Journal of Memetics - Evolutionary Models of Information Transmission, 6. <[http://jom-emit.cfpm.org/2002/vol6/gatherer\\_d.html](http://jom-emit.cfpm.org/2002/vol6/gatherer_d.html)>
- Gatherer D (2002b) Identifying cases of social contagion using memetic isolation: comparison of the dynamics of a multisociety simulation with an ethnographic data set. Journal of Artificial Societies and Social Simulation 5, no. 4 <<http://jasss.soc.surrey.ac.uk/5/4/5.html>>

Gatherer D (2004) Birth of a meme: the origin and evolution of collusive voting patterns in the Eurovision Song Contest. Journal of Memetics - Evolutionary Models of Information Transmission, **8**. <[http://jom-emit.cfpm.org/2004/vol8/gatherer\\_d\\_letter.html](http://jom-emit.cfpm.org/2004/vol8/gatherer_d_letter.html)>

Kuper A (2000). If memes are the answer, what is the question? In: Aunger R (ed) Darwinizing Culture: The Status of Memetics as a Science. Oxford: Oxford University Press

Lynch A (2002) Thought contagion in the dynamics of mass conflict. Swedish Defence Research Agency base data report FOI-R—0745—SE. Edited by Dahlberg M.

Marsden, P. (2001). Copycat Terrorism: Fanning the Fire. Journal of Memetics - Evolutionary Models of Information Transmission **5**. <[http://jom-emit.cfpm.org/2001/vol5/marsden\\_p\\_let.html](http://jom-emit.cfpm.org/2001/vol5/marsden_p_let.html)>

Raczynski S (2004) Simulation of The Dynamic Interactions Between Terror and Anti-Terror Organizational Structures Journal of Artificial Societies and Social Simulation **7**, no. 2 <<http://jasss.soc.surrey.ac.uk/7/2/8.html>>

# THE INHERENT INSTABILITY OF MEMETIC SYSTEMS: USE OF A GENETIC ALGORITHM TO SOLVE A PARAMETER OPTIMIZATION PROBLEM IN A MEMETIC SIMULATION

By

Derek Gatherer

[dgatherer@talk21.com](mailto:dgatherer@talk21.com)

## Abstract:

'Memetic isolation' is a situation in which a society or culture exhibits a cultural trait not found in any neighbouring cultures (Gatherer 2002b). A previously developed simulation, consisting of a grid of connected societies of cultural agents, is further analysed to define the circumstances under which memetic isolation is maximised and minimized. Parameters varied include migration, and cultural interaction within and between societies. Some stereotypical societies are simulated, and the differences in outcomes are analysed statistically. A genetic algorithm is then used to discover the combinations of parameters that produce extreme results. Minimization of memetic isolation appears to be simply a matter of cultural or natural selection on the meme concerned. By contrast, maximization of memetic isolation requires an intuitively unlikely combination of low cultural interaction, high migration and no selection. The classic memetic theoretical result, that horizontally transmitted memes tend to be more spatially clustered than vertically transmitted memes or genes, is shown to depend on the existence of cultural bridges, or 'weak ties', between societies and also to be at the mercy of migration rates and selective forces.

**Key words:** Contagion, evolutionary epidemiology of culture, cultural evolution, cultural selection, meme, allomeme, cultural trait, genetic algorithm, globalisation.

## 1. Introduction

A large literature exists on the relationship between genetic isolation and cultural isolation (e.g. Sokal *et al.* 1989; Barbujani & Sokal 1990; Vona *et al.* 1996; De Silvestri & Guglielmino 2000; Gonzalez Jose *et al.* 2002). Genetic isolation may be the result of a 'founder effect', where a population expands rapidly from a small set of original progenitors. This often happens in colonial situations, e.g. in South Africa (Torrington & Viljoen 1991), Venezuela (Ramirez-Duque *et al.* 1982; Vivenes De Lugo *et al.* 2003), Canada (Scriver 2001) or Finland (Kittles *et al.* 1999). It may also follow a sudden 'population crash' caused by a natural disaster or epidemic disease, followed by regeneration of the population from the survivors (Ambrose 1998; Harpending *et al.* 1998). All of these factors may also cause cultural isolation, which can then contribute to further genetic isolation if interbreeding with geographically neighbouring populations is prevented or reduced by cultural factors. However, the two phenomena are not necessarily related, and one may exist without the other.

Cultural isolation may be seen as an extreme case of 'memetic isolation' (Gatherer 2002b), defined here as the situation in which any cultural trait of a society is different to the corresponding mutually exclusive cultural trait, the 'allomeme' (Durham 1991), of any of its surrounding cultures. A true cultural isolate will be memetically isolated for an extensive selection of cultural traits.

In classic mathematical memetic theory, 'vertical transmission' refers to traits that are passed on from one generation to the next, specifically from parents to progeny (a variant is 'oblique transmission' in which traits are passed within families, but not necessarily from parent to child). By contrast 'horizontal transmission' refers to transmission within a single generation, either between siblings or other genetic relatives, or within a non-related peer group (Cavalli-Sforza & Feldman 1981). One of the results of this body of theory is that traits with more than one means of transmission have a greater tendency to homogeneity within populations, and also that horizontally transmitted traits are more likely to be spatially clustered (Cavalli-Sforza & Feldman 1973, Uyenoyama *et al.* 1979, reviewed by Cavalli-Sforza 1979). 'Spatial clustering' in this context means that individuals exhibiting the trait tend to be found in closer geographical proximity than those who do not exhibit the trait.

The present paper further investigates this area of theory using agent-based simulation. The simulation was previously described in detail in Gatherer (2002b), and may be considered as a variant of the Social Interaction Model (SIM) simulation of Axelrod (1997). Two questions are asked:

- 1) Under what combination of parameters is memetic isolation maximised?
  - 2) Under what combination of parameters is memetic isolation minimised?
- The answers provided enable us to refine the predictions of the original theoretical treatments of transmission and its relation to spatial clustering of traits (reviewed by Cavalli-Sforza 1979).

## 2. The Model

Since frequent comparison is made to the Social Interaction Model simulation of Axelrod (1997), this is described first. Imagine a set of distinct societies or cultures arrayed in geographical space. Let each culture consist of a small number of cultural traits, in the order of 10 or so.

Adjacent cultures may interact if they have a sufficient number of cultural traits in common. This situation is illustrated in Figure 1, where a minimum of four out of six cultural traits in common are required in order for the societies to interact with each other.

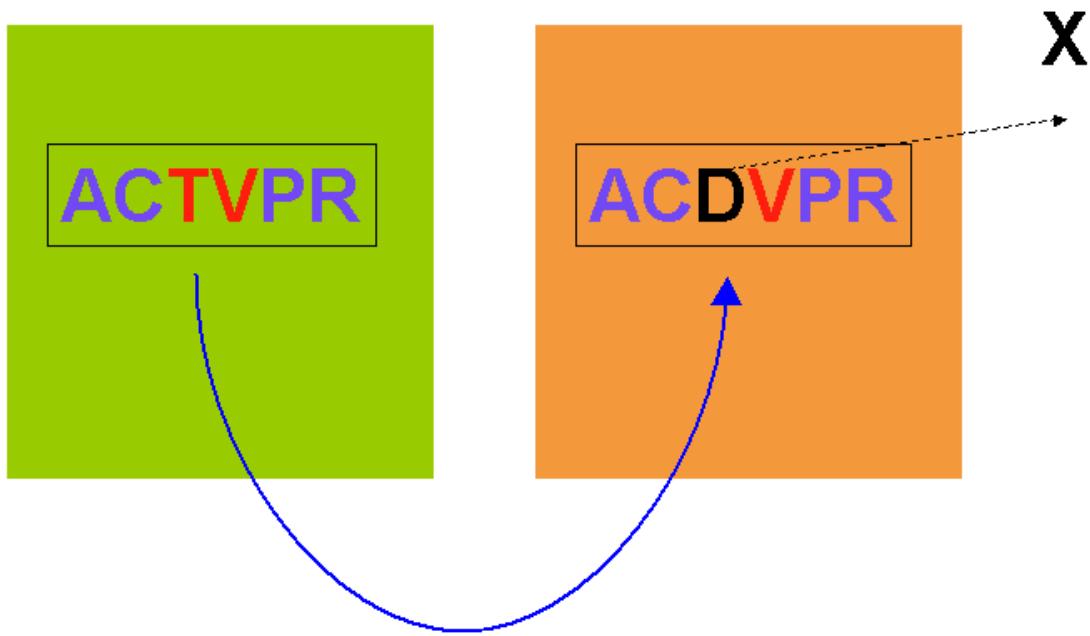
## Axelrod SIM interaction rule



**Figure 1:** Rules for possibility of interaction in Axelrod's SIM. Cultural traits are represented by letters, and coloured to show which are found in common between pairs. Since the upper pair has 4 out of 6 cultural traits in common, these two societies can interact. However, the lower pair only has 2 traits in common and therefore cannot interact.

Once the two societies have established an interaction, one society may randomly adopt a cultural trait from the other. This is illustrated in Figure 2. At the end of the interaction, one of the societies has become slightly more similar to its neighbour. It is easy to see how this simple situation mirrors the process of cultural transfer in real societies. Cultures with very differing languages, religions or habits may fail to make the necessary cultural contact. However, once a certain threshold has been crossed and communication is established, cultural interaction can commence. This model makes the assumption that there is no selection for or against any cultural trait, or that traits are imposed wholesale, e.g. by conquest.

# Axelrod SIM cultural adoption

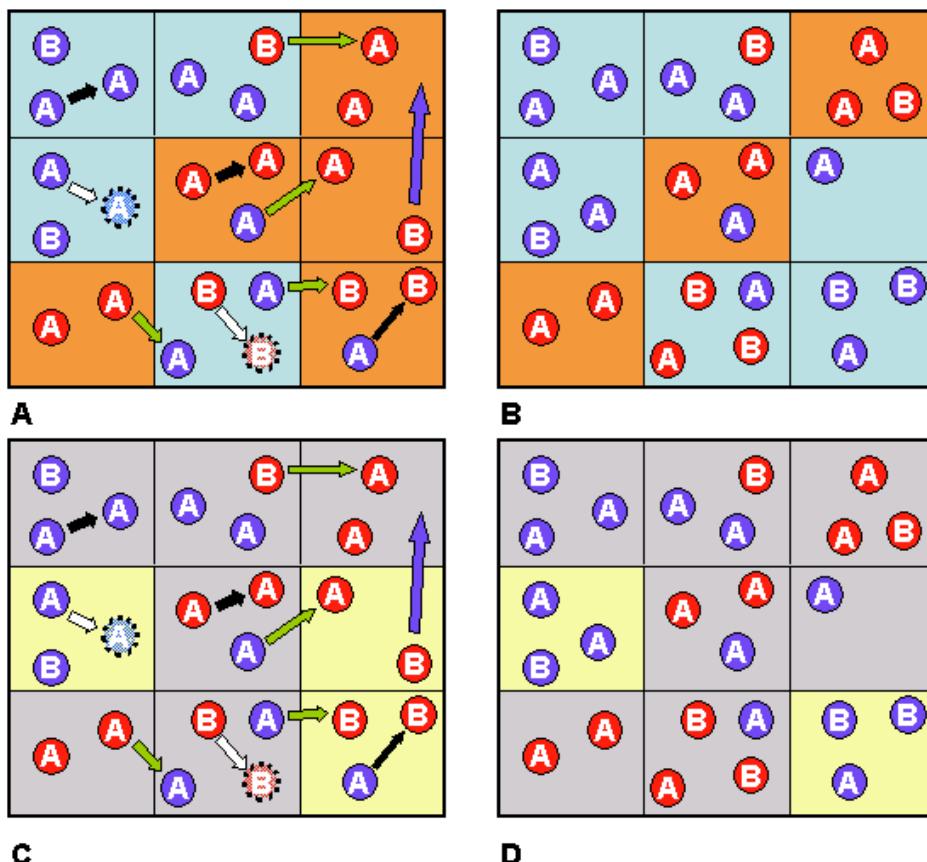


**Figure 2:** Adoption of one cultural trait from adjacent society in interacting SIM pair. Trait X in the right society is replaced by trait V from the left society. Having had 4 traits in common, the pair of societies now has 5 in common.

This process is repeated over several cultural generations in a grid of several cultures, typically 10 by 10, starting from a random distribution of traits. One of the principal results of SIM is that cultural 'blocs' are formed, i.e. areas on the cultural landscape with several near-identical cultures. This is especially the case if the probability of interaction is varied according to the relatedness of a pair of societies, i.e. if a pair of societies with 5 traits in common has more chance of interaction than a pair with only 3 in common, and so on. Intuitively, it can be seen that this tends to mean that societies that have previously interacted will be increasingly likely to interact further in future generations, thus producing a positive feedback.

The variation on SIM presented previously by Gatherer (2002b) uses an identical concept of a grid of societies in geographical space. The size of grid presented in the scenarios below is 10 x 10, the same order of magnitude used by Axelrod (1997). However, the focus is switched to the individual agent with the society rather than the society as a whole. SIM assumes that each society is culturally homogeneous, insofar as each society has a single set of cultural traits. Figure 3 shows how this may be adapted to allow individuals within societies to differ in their repertoire of memes and genes. The word 'genes' in this case is used as a shorthand for 'strictly vertically transmitted memes', since in haploid organisms (i.e. those with a single copy of each

gene in their genomes) these categories are formally equivalent (Cavalli-Sforza & Feldman 1981). In panel **A** of Figure 3, a grid of nine societies may be seen. Individuals are indicated as circles with their vertical memetic component as the letter **A** or **B**, and their horizontal memetic component as the colour red or blue. These individuals may perform three kinds of lifecycle activity: reproduction, migration and cultural interaction. In Figure 3**A** and 3**C**, cultural interaction is indicated with a black arrow if it occurs within societies (e.g. in the bottom right society of panel **A**, an individual with blue meme teaches it to an individual with red meme, thus resulting in that individual having blue meme in panel **B**), and a green arrow if it occurs across inter-societal boundaries. Reproduction is indicated with a white arrow and the newly produced individuals are surrounded by a dotted line in 3**A** and 3**C**. Migration is indicated by a blue arrow, and can only occur to neighbouring societies (in other words this is a 'two-dimensional stepping stone' model of migration - Rychkov & Sheremetyeva 1977). Panels 3**B** and 3**D** illustrate the end result after the processes displayed in 3**A** and 3**C**. The societies in panels **A** and **B** are coloured blue or orange according to the majority of blue or red memes respectively. In panels **C** and **D** the societies are coloured yellow or grey according to the majority of genes (or strictly vertical memes) **A** and **B**, respectively.



**Figure 3:** Reproduced from Gatherer (2002b). **A** and **B** labels are strictly vertical memes, red and blue colours are horizontal memes. Thus a blue individual labelled **B** has the **B** vertical meme and the blue horizontal meme. Black arrows: horizontal cultural interaction within societies. Green arrows: horizontal cultural interaction between societies. Blue arrows: migration

of individuals. White arrows: reproduction (new progeny are identical to their parents and are shown with dotted outlines). Orange squares: societies with majority red memes. Light blue squares: societies with majority blue memes. Grey squares: societies with majority of **A** vertical memes/genes. Yellow squares: societies with majority of **B** vertical memes/genes. Fuller explanation given in text.

In summary, the simulation consists of a grid of connected societies populated by cultural agents, its dynamics being defined with a set of four simple parameters:

- the rate of population growth, **r**
- the rate at which teaching/learning occurs, **o**
- the rate at which teaching/learning occurs between societies, **n**
- the rate at which individuals migrate to neighbouring societies, **m**

These parameters are designated mnemonically as **r**, **o**, **n** and **m**, respectively (i.e. reproduce, own-teach, neighbour-teach and migrate). All take values between zero and one, and represent probabilities that are either user-defined or created using a random number generator. The parameters **n** and **o** are interdependent, with **o** always evaluated first, and then **n** evaluated depending on the result of the evaluation of **o**. For example when **o** is 0.5, and **n** is 0.6, an agent will exhibit successful cultural transmission during its lifetime at probability 0.5. If that agent does exhibit cultural transmission, there is then a probability of 0.6 that that cultural transmission will be to a member of a neighbouring society, and 0.4 that it will be to a member of the same society. This could be re-expressed in terms of the notion of 'strong' and 'weak' ties, where **o** is the probability of existence of a tie, and **n** the probability that such a tie is 'weak' (sensu Granovetter 1973). A pseudocode representation of the process is as follows:

#### INITIALIZATION:

For each of the 100 cells in the 10-by-10 array:

```
{
    place two Agents in each cell
    randomly assign one of four Gene attributes to each individual
    randomly assign one of four Meme attributes to each individual
}
```

#### ITERATION:

For 100 generations do the following:

```
{
    For each Agent:
    {
        If Agent Meme attribute = "A" and CultSel = "Y"
        {
            Double parameters n and o for that agent
        }
        If Agent Meme attribute = "A" and NatSel = "Y"
```

```

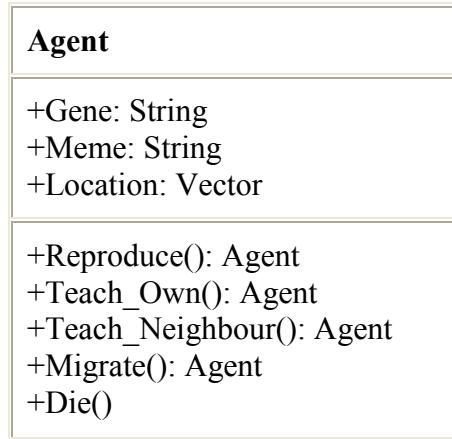
{
    Double parameter r for that agent
}
If a random number x < r
{
    Reproduce Agent
}
If a different random number y < o
{
    If same random number y is also < n
    {
        Transmit Meme to any Agent in any adjacent cell
    }
    else
    {
        Transmit Meme to any Agent in same cell
    }
}
If a different random number z < m
{
    Migrate Agent to adjacent cell
}
}
}

```

END.

Additionally, 'cultural selection' or 'natural selection' may be applied to one of the memes in the population. Here, these terms are used in the sense of Cavalli-Sforza & Feldman (1981), as follows: a culturally selected meme is one that has an increased probability of transmission relative to its allomemes; a naturally selected meme is one that provides a survival or reproductive advantage to those individuals that exhibit that behaviour or cultural trait.

Figure 4 provides a diagrammatic representation of the agent in UML (Universal Modelling Language). The simplicity of this agent (all attributes and methods public, no constructors etc) is due to its implementation as a Perl record rather than a true object in Java or C++. Perl records are fairly analogous to struct in C. Agents have one of four Meme attributes ("A", "B", "C", "D") and one of four Gene attributes ("1", "2", "3", "4").



**Figure 4:** UML diagram for the Agent class. Attributes are above the line and Methods below.  
Public scope is represented by a "+" (all are public)

Perl scripts are provided in the Appendix that:

- run the simulation with graphical output
- implement a genetic algorithm (GA) to optimise the parameters.

Both scripts require Perl, and the first also requires the Tk graphics library. A recent version of ActivePerl (Build 6 series or more recent, freely available from <http://www.activestate.com>) provides all the necessary library components. All scripts provided have been tested on RedHat Linux 6 running Gnome, and on Microsoft Windows 98.

### 3. The Effects of Varying Parameters on Memetic Isolation

The 'memetic state' of a society for a certain cultural trait, is here defined by the 'first-past-the-post' voting method. The Meme attributes of each agent in the society are scored, and if allomeme "A" scores more than any of the other three Meme attributes, the memetic state of that society is "A". If allomeme "A" is not the highest scoring attribute, the memetic state of that society is "not-A". 'Memetic isolation' occurs when that society's 'memetic state' is "A" and that of all its immediate neighbours is "not-A". The overall memetic isolation is the percentage of societies that are in this state. It is admitted that 'first-past-the-post' is potentially a flawed means of scoring a society, but it is often used in sociological contexts. For instance, statements such as the "Belgium is a Catholic country" are simply expressions of such head counts. Qualifying those situations where the 'first-past-the-post' voting produced a narrow majority, perhaps by the use of a 'marginal' label, could further refine the model. However, this would be complex enough to merit another paper, and so is not dealt with here.

'Genetic isolation' can similarly be calculated using the Gene attribute "1". Since the agents are haploid (i.e. they have only a single Gene attribute), the Gene is equivalent in terms of its transmission to a strictly vertically transmitted meme, as demonstrated by Cavalli-Sforza & Feldman (1973). It is thus possible to contrast 'memetic isolation' and 'genetic isolation' in terms

of the likelihood of isolation for horizontal and vertical memes. Therefore, 'genetic isolation' is used hereafter simply as a shorthand for 'memetic isolation for strictly vertical traits'.

### **3.1 Scenario 1: The Random Background - the "strictly vertical model"**

The scenario is initialised with two individuals per cell, and allowed to run for 100 generations with a rate of population growth (**r**) of 0.01 per generation. Migration (**m**), and the two social interaction variables (**o** and **n**) are set at zero. This simulates a situation where societies grow slowly in size, but there is no 'horizontal transmission' (Cavalli-Sforza & Feldman 1981), or movement of individuals across society boundaries. It is essentially a non-cultural population, or perhaps more exactly a population where culture is contained strictly within vertical lineages. This is given merely to demonstrate the properties of the system in a background state, where as little as possible is happening except the effects of randomness over a large number of cycles. It is difficult to think of a real-world example of this kind of cultural situation; even non-human species with aspects of culture, such as song birds and great apes, do not limit cultural transmission within genetic lineages (e.g. Burnell 1998; Lynch et al. 1989; Reader & Laland 2000; van Schaik et al. 2003). The scenario was repeated 100 times and the average memetic and genetic isolation measured. These were found to be 3.23% and 3.00% respectively. Statistical significance was assessed using a paired t-test. In this and all subsequent scenarios, t-tests are two-tailed, as there is no a priori expectation that parameters will increase or decrease levels of isolation. A paired t-test between the genetic and memetic isolation values for each of the 100 runs, was not significant, thus indicating that both horizontal and vertical memes behave identically in strictly vertical populations. In effect, the 'horizontal memes' are forced to become vertical memes by the absence of cultural interaction in the population, and thus this scenario constitutes a sort of *reductio ad absurdum*. However, it is necessary to demonstrate the basic background behaviour against which other scenarios will be measured.

### **3.2 Scenario 2: Introducing Cultural Exchange within Societies - the "desert island model".**

Social interaction within societies is now permitted, by raising parameter **o** to 0.5. This means that each agent will transmit a horizontal meme with probability 0.5 within its lifetime, in other words half of all agents transmit horizontally per generation, and half do not. All other parameters are the same. This recreates the situation of societies between which there is no contact, but which have a flourishing social life of their own, such as might occur on extremely isolated oceanic islands. It should be noted at this point, and for all subsequent scenarios, that it is recognised that real societies are infinitely more complex than any simulation can reproduce. The purpose of adding labels such as 'desert island' to the simulations, is really to serve as a mnemonic shorthand, and also to perhaps act as an 'intuition pump' for practical applications. After 100 generations, over 100 runs of the scenario, the average genetic and memetic isolation values are 3.16% and 4.08% respectively. A paired t-test shows these to be statistically significantly different at  $p<0.0004$ . Thus, against a background of no migration and no cultural exchange between societies, social interaction within our artificial societies increases the likelihood that those societies will be memetic isolates. This is because the rate of the random walk amongst competing allomemes in those societies is speeded (see Gatherer 2002a, Fig. 2). If **o** is raised to 1, allowing each agent a successful cultural transmission event per lifetime, the figures are essentially unchanged. This is called a 'desert island' model, since it would apply only

to societies that were completely cut off from the rest of the outside world. Some examples are known, of which the best studied is perhaps that of the Bounty mutineers on Pitcairn Island, who remained completely undiscovered for decades after settling, and in effective cultural isolation for some years after. In this period they were able to evolve a unique culture incorporating both European and Polynesian elements, including a new creole English language (Refshauge & Walsh 1981). If one imagines a thought experiment in which the Bounty mutineers colonised two separate islands, mutually unconnected, the model predicts that they would become memetic isolates for many traits. This rarely applies to many island archipelagos, as the inhabitants are often in contact, albeit infrequently. This leads to the next scenario where such an arrangement is permitted.

### **3.3 Scenario 3: Introducing Limited Cultural Exchange between Societies - the "Paleolithic model"**

Parameter **n** is now raised from zero to 0.01. This permits cultural exchange with a neighbouring society at a probability of  $0.01 \times 0$  per individual per generation. Therefore at **o** of 0.5, the rate of inter-cultural exchange is 0.005. The societies are now no longer totally isolated "desert islands", but have the kind of casual contact perhaps normal for sparsely distributed and occasionally trading nomadic bands. Over 100 runs of the scenario, again for the standard 100 generations, average genetic isolation remains around the background level at 3.12%, but memetic isolation drops to 1.78%. A paired t-test shows this to be a significantly different value for memetic isolation at levels of **p** tending to zero. Thus even a small amount of cultural exchange between our artificial societies may, *ceteris paribus*, create tendencies to cultural homogenisation on a local geographical scale. Cultural blocs (Axelrod 1997) thus begin to emerge out of a patchwork background. Where parameter **n** is increased, memetic isolation falls even further: at **n** = 0.1, memetic isolation falls to 0.71%. Thus in a scenario of 100 societies, most runs terminate without a single memetic isolate at the end of 100 generations. This is called the 'Paleolithic model' as it very roughly recreates a situation hypothesised to have existed in Late Ice Age Europe, where nomadic bands of hunters of large herds of bison, mammoths etc, would have existed in isolation punctuated by brief seasonal periods of contact at assembly points, perhaps associated with exchange of gifts, arrangement of marriages and ritual behaviour (e.g. Rudgley 2000). A more modern example may be provided by the Andamanese, who remained in relative isolation within their archipelago for many thousands of years (e.g. Thangaraj et al. 2003). Again it bears reemphasis that the correspondence between the artificial societies and the anthropological examples given is quite naïve, and really just raises the issue of how better models of such societies could be created, rather than claiming to have created one here.

### **3.4 Scenario 4: Introducing Natural Selection - the "Neolithic model"**

Meme "A" is now submitted to natural selection. This means that those agents with Meme attribute "A" have twice the normal chance of reproduction of those with the other three allomemes. Meme "A" agents thus reproduce above replacement at a probability of 0.02 per generation, instead of the 0.01 for the other three allomemes. This is termed here the "Neolithic Model" as it is reminiscent of one of the main theories for the Neolithic transition in Europe, namely that the cultural activity of farming provided a reproductive and survival advantage, in that early farmers were able to feed more children and live longer to reproduce more, compared

to their Mesolithic hunter-gatherer neighbours (Renfrew 1987). If **m**, **n** and **o** are all set to zero, so that there is no migration or cultural exchange within the population, but natural selection is permitted to increase **r** from 0.01 to 0.02 for Meme "A" individuals, genetic isolation after 100 generations is 3.68% and 2.34%. This memetic isolation is significantly different to the background "pre-cultural model" (See 3.1), although both models have no cultural exchange. The reason for this is that the evolutionary advantage of Meme "A" in terms of natural selection means that this meme will tend to take over the populations in which it is found. If it happens to occur in two adjacent societies, even at initial low frequencies, it will eventually run to fixation in those societies. Of course, this is not really a good "Neolithic model", as Neolithic societies also had cultural exchange as well as natural selection for a meme (assuming for the present that this particular theory of Neolithic transition by demic expansion is true). A second "Neolithic model", allowing levels of cultural exchange equivalent to the "Paleolithic model", i.e. **o** = 0.5 and **n** = 0.01, has levels of genetic and memetic isolation of 3.37% and 1.11%, respectively. A t-test shows that this is a significantly different level to that when there is no natural selection. Natural selection of a meme thus reduces its likelihood of memetic isolation.

### **3.5 Scenario 5: Introducing Cultural Selection - the "Contagion model"**

The second "Neolithic model" above, now has natural selection of Meme "A" replaced by cultural selection. This means that Meme "A" will be transmitted at double the normal frequency of cultural exchange. Thus with **o** = 0.5 and **n** = 0.01, as above, Meme "A" has probabilities of transmission of **oA** = 1 and **nA** = 0.02. Genetic and memetic isolation after 100 generations average to 3.24% and 0.21%, respectively. Therefore, cultural selection almost removes memetic isolation entirely.

### **3.6 Summary of scenarios**

**Table 1 show the summary of the scenarios**

Scenario number, number of section where described and name of model	<b>m</b>	<b>n</b>	<b>o</b>	select.	Genetic isolation (%)	Memetic isolation (%)
1: 3.1 Strictly vertical	0	0	0	None	3.23	3.00
2: 3.2 Desert island	0	0	0.5	None	3.16	4.08
2: 3.2 Desert island - 2	0	0	1	None	3.26	4.04
3: 3.3 "Paleolithic"	0	0.01	1	None	3.12	1.78
3: 3.3 "Paleolithic" - 2	0	0.1	1	None	3.33	0.71
4: 3.4 "Neolithic"	0	0	0	Natural	3.68	2.34

4: 3.4 "Neolithic" - 2	0	0.01	0.5	Natural	3.37	1.11
5: 3.5 Contagion	0	0.01	0.5	Cultural	3.24	0.21

**Table 1:** Summary of the average genetic and memetic isolations resulting from the various scenarios. The 'strictly vertical' model provides the random background level of isolation. The 'desert island model' shows the highest level of memetic isolation. **m**: migration, **n**: cultural interaction within societies, **o**: cultural interaction between societies

Several questions may now be answered by using t-tests to compare the statistical significance of differences between rows in Table 1. In this section, t-tests are two-tailed and non-paired, since comparison is between different scenarios. An F-test is first carried out to establish if the t-test should be heteroscedastic or homoscedastic.

What factors affect genetic isolation?: Method: t-tests between genetic isolation data for all rows against row 3.1 (where there is the random background level of genetic isolation). Answer: no model significantly changes the likely degree of genetic isolation of each society. The differences within the range of 3.12% to 3.68% are all explicable as statistical random variation. This applies equally to cultural traits that are strictly vertically transmitted.

What factors affect memetic isolation?: Method: t-tests between memetic isolation data for all rows against row 3.1 (where there is the random background level of memetic isolation). Answer: by contrast with the above, all models have a level of memetic isolation significantly different to the random background at  $p < 0.01$ . The answer is therefore completely different to the above, in that one might say that several factors may influence memetic isolation, including cultural and natural selection, and the rates of cultural transmission.

Does the "Neolithic - 2" have less memetic isolation than the basic "Neolithic" model? Method: t-tests between memetic isolation data for rows "3.4 Neolithic" and "3.4 Neolithic - 2". Answer: yes, at  $p$  tending to zero. Therefore, it is possible to conclude that an increase in social interaction within a society decreases memetic isolation when the meme is under natural selection, whereas it increases it when natural selection is not operating (compare "Strictly vertical" and "Desert Island").

Does "Contagion" have less memetic isolation than any of the other models? Method: t-tests between memetic isolation data for all rows against row 3.5. Answer: yes, "Contagion" has a significantly lower level of memetic isolation than all other models at  $p$  tending to zero. It is thus possible to say that culturally selected traits are unlikely to exist in isolation.

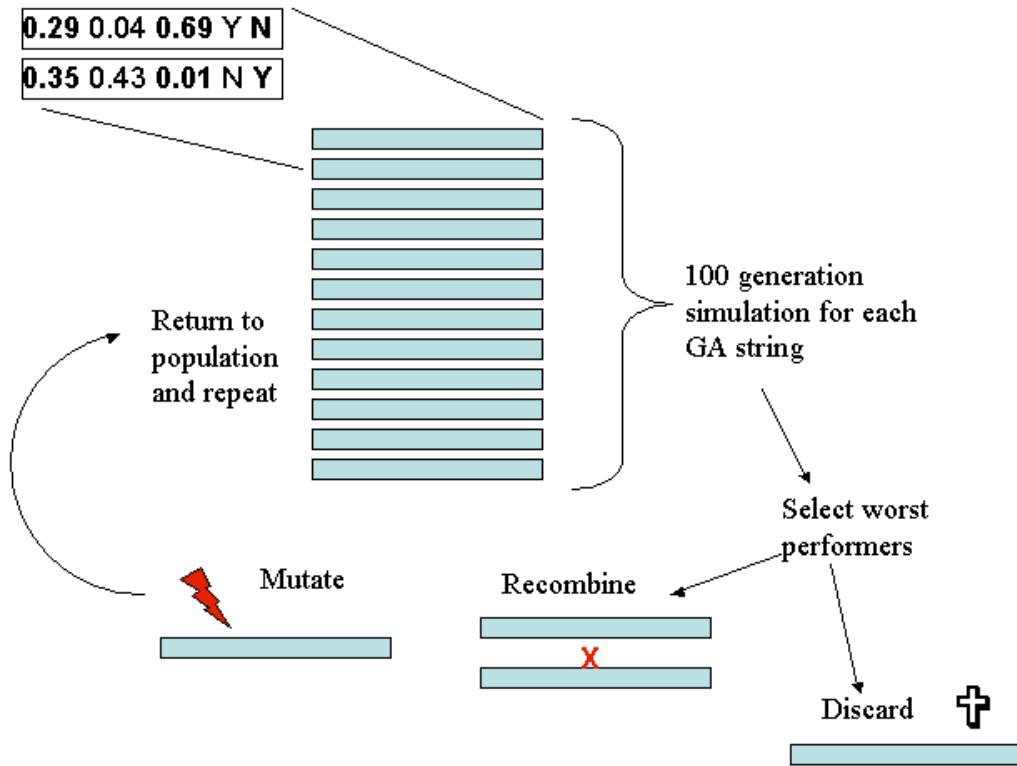
#### 4. The Genetic Algorithm

The above set of simulations only dealt with seven scenarios as compared to a random background. However, since there is a combinatorial explosion of possible scenarios, a heuristic method must be used to select the scenario most likely to give rise to extreme conditions.

A standard method for use in such situations is the Genetic Algorithm (GA - reviewed by Bentley 1999). The general form of the GA used here is as follows:

```
START: Randomise representations of the parameters in a linear string.  
FOR n CYCLES  
{  
    Assess the result of those parameters by 100-generation simulation  
    Rank strings by the result under investigation (e.g. maximal memetic isolation)  
    Discard very worst performers  
    Recombine and/or mutate remainder of unacceptable perfumers  
    Retain and reproduce acceptable performers to maintain constant numbers of strings  
}  
END.
```

In the GA presented here, the threshold below which GA strings are discarded is 0.775 standard deviations below or above the average memetic isolation, depending on whether the GA is attempting to maximise or minimise memetic isolation. The threshold for mutation and/or recombination is 0.5 standard deviations above/below the average. Mutation and recombination occur with probabilities of 0.05 and 0.1 respectively, but only within the sub-standard group. The better performing strings are thus protected from mutation and recombination (this is quite a Lamarckian GA!). These properties were derived by a process of empirical trial and error, and there are many other forms of GA that could be used. The balance between homogenisation, due to selection, and diversity, due to mutation and recombination, is a delicate one, and has to be solved afresh for every situation in which a GA is used. It should be noted that the fitness landscape is extremely noisy. In all the scenarios above in section 3, the standard deviation of the memetic isolation is high. For this reason, it appears to be more profitable to protect the better performing strings. The GA process is illustrated in diagrammatic form in Figure 5, and the continuing high standard deviation may be seen in Figure 6.



**Figure 5:** Diagrammatic representation of the GA process. The starting sets of parameters for the model are represented as a string of length 5. Each string is run as a separate simulation, and the results are assessed according to the desired outcome (e.g. high memetic isolation or low memetic isolation). The 'worst' performers according to the desired criteria are either discarded, mutated or recombined with a randomly chosen other member of the population. The discarded strings are replaced by randomly chosen strings from the better performing part of the population. Selection of GA strings thus tends to homogenise the population while mutation and recombination maintain some variability. The whole process was repeated over 100 cycles.

A set of GA strings was prepared by semi-randomly initialising the variables, **m**, **n** and **o**, and the selective conditions for cultural and natural selection, 50 times. The initialisation is semi-random as the starting set was biased to give fairly low levels of migration. The initial set produced is shown in abbreviated form in Table 2. After 100 iterations of a GA process, selecting for maximum memetic isolation at each stage, the 50 populations corresponding to the 50 GA strings have average memetic isolation of 4.60% and average genetic isolation of 4.72%. The set after 100 GA generations is shown in abbreviated form in Table 3. The most obvious feature of the action of the GA over 100 generations of searching for best parameters for maximum memetic isolation is to eliminate cultural and natural selection. The GA also optimises memetic isolation by reducing the rate of cultural exchange within the culture, variable **o**, from the initial average of 0.462 to less than 0.005. Migration is increased from 0.048 at initialisation to 0.559. The optimal value for teaching to other societies is also maintained fairly high, at 0.432 from

0.501 at initialisation, but this will still represent a small level of cultural exchange since the background probability of cultural exchange is less than 0.005.

Growth rate	Teach own society	Teach other society	Migration	Cultural Selection	Natural Selection
1 0.01	0.24	0.89	0.04	Y	N
2 0.01	0.41	0.46	0.07	Y	Y
3 0.01	0.44	0.58	0.05	N	Y
Several rows.....					
49 0.01	0.12	0.06	0.01	N	N
50 0.01	0.14	0.90	0.07	Y	N
Average	0.462	0.501	0.048	21 Y	18 Y

**Table 2:** Truncated listing of 50 GA strings semi-randomly initialised at 5 parameters

Growth rate	Teach own society	Teach other society	Migration	Cultural Selection	Natural Selection
1 0.01	0	0.81	0.5	N	N
2 0.01	0.01	0.81	0.5	N	N
3 0.01	0	0.81	0.5	N	N
Several rows.....					
49 0.01	0	0.81	0.5	N	N
50 0.01	0	0.81	0.5	N	N
Average	0.0048	0.432	0.5592	0 Y	0 Y

**Table 3:** Truncated listing of 50 GA strings after 100 generations of the GA process, selecting for maximum memetic isolation

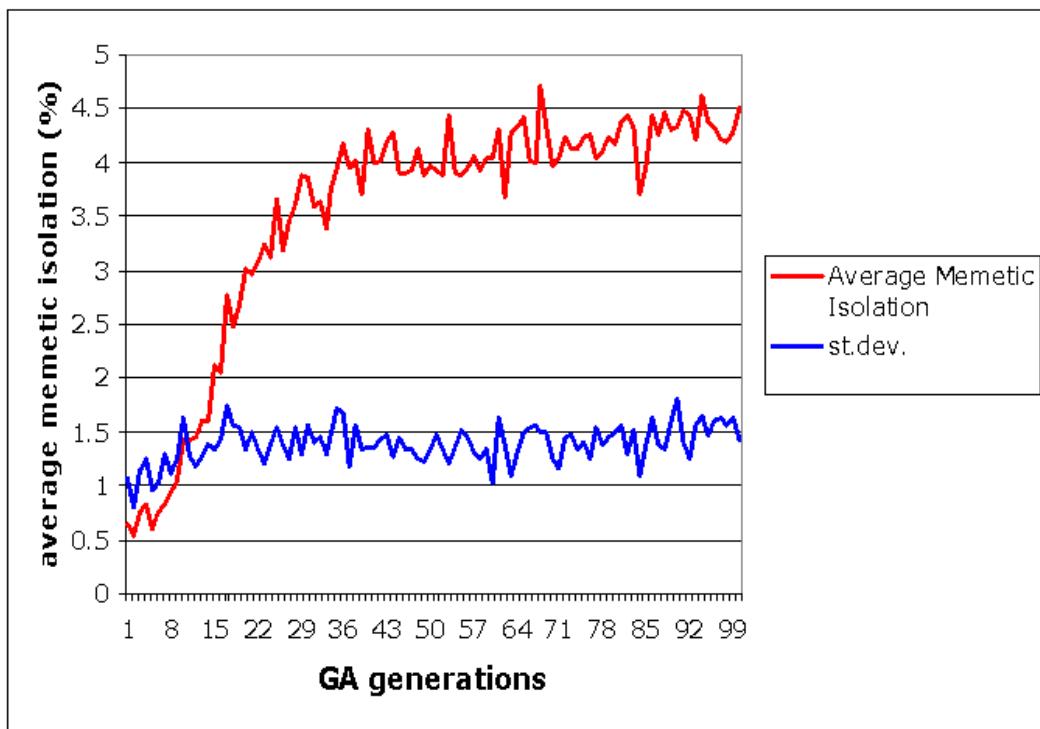
When the same process is performed, but the selective conditions are to minimise memetic isolation, the resulting average parameters are compared in Figure 4. The minimisation GA was

only performed over 7 rounds of selection, as compared to 100 rounds for maximization, as it rapidly achieves a near-zero rate of memetic isolation.

GA seln. (rounds)	Teach own soc.	Teach others	Mig.	Cult. Seln.	Nat. Seln.	Mem. Isol.	Gen. Isol.
Minimise (7)	0.49	0.60	0.49	23/25 Y	23/25 Y	0.02%	3.43%
Maximise (100)	0.0048	0.432	0.5592	0/25 Y	0/25 Y	4.60%	4.72%

**Table 4:** Comparison of 25 rounds of GA selection for the optimal parameters to maximise memetic isolation, with 7 rounds of selection for the optimal parameters to minimise memetic isolation.

The progress of the memetic isolation-maximising GA is shown in Figure 6.



**Figure 6:** Progress of the genetic algorithm in selecting parameters to maximise memetic isolation. Red: the average memetic isolation in the 50 populations corresponding to the 50 GA strings. Blue: the standard deviation of that average throughout the 50 populations. The persistence of a relatively high standard deviation illustrates the 'noisiness' of the system.

## 5. Discussion

The contemporary phenomenon of globalisation involves cultural as well as economic aspects. It has been estimated that 5.6% of the world's languages are in danger of extinction, and a further 4.4% have recently become extinct (Sutherland 2003). In the Social Interaction Model (Axelrod 1997), this phenomenon was seen as an emergent property. SIM is another grid-based agent model in which neighbouring agents are permitted to exchange information with other societies providing they shared at least one cultural trait. SIM has a single agent per grid square, which can either be taken as an individual or as a whole, and necessarily homogeneous, society (Axelrod suggests 'homogeneous villages'). There is no migration. SIM leads to the increasing homogenisation of agents in close proximity, while at the same time agents that had by chance no common traits with their neighbours, remain completely differentiated. If the SIM agents are treated as individuals, SIM demonstrates how neighbouring individuals can form homogeneous cultural areas. Similarly, if the SIM agents are regarded as homogeneous societies, SIM demonstrates the formation of cultural 'blocs' on a local geographical basis. The model presented here differs from SIM, in that it is necessarily based on individuals. Societies are simply geographical areas, whose cultural profiles depend on those of the agents that inhabit them.

Another important difference is that in the present model, the level of social interaction between societies is constant for each run of the scenario, and is defined by parameter **n**. However, in SIM, the analogous parameter is proportional to the number of cultural traits that neighbouring agents have in common. Thus, if there are no traits in common, neighbouring agents will not interact, thus creating a local 'desert island' model. If the SIM agents are seen as homogeneous societies, SIM thus simultaneously operates a 'desert island' model and a 'Palaeolithic model' within the same scenario, with the two models operating in different parts of the grid, with shifting boundaries as the scenario progresses. A model more similar to the present one in its structure, is given by Ray et al. (2003), but those authors focus exclusively on the genetic phenomena and do not deal with horizontally transmitted cultural entities.

The genetic algorithm process presented here seeks to derive parameters for the prevention of cultural blocs. It delivers maximal memetic isolation (an average of 4.6% of societies are memetic isolates after 100 generations of the scenario) at levels of social interaction tending to zero (less than 0.5% per person per lifetime), and with high migration (over 55% per person per lifetime), where no selective conditions operate. This describes a rather nightmarish scenario where there is little or no social communication between continually wandering individuals. There is some superficial similarity here to theories of 'alienation' (Geyer 1994). Although technically, this might be the optimal way of maintaining memetic diversity in a grid of artificial societies, it is clearly not a solution that any would prescribe in the real world. It is also completely vulnerable to selection pressures where they arise. Even in situations of very low social interaction, a culturally selected meme will tend to spread through a population (see Gatherer 2002a, Fig. 4). When there is no social interaction at all, and cultural selection ceases to operate, natural selection can still push memes to fixation through growth of the lineages that exhibit them. In effect, extreme migration rates cause the cellular nature of the scenario to break down, and effectively become a single society. Parameters **o** and **n** then cease to have any significance, as cultural transfer occurs principally by transfer of people.

How would one make the model more comparable to the real world? Empirical studies of Australian Aborigines suggest a realistic figure of 0.12 to 0.13 for  $\mathbf{m}$ , where the pattern of migration is mostly to neighbouring tribal groups and rarely to more distant tribes, reflected in the present model (Lasker & Crews 1996). Migration rates of  $\mathbf{m} \sim 0.55$  are rare in the real world. Where they do occur they are often in specific directions, e.g. invasions, flight of refugees, economic migration, and are not reproducible in the present model. Even when they do, they are often smaller than  $\mathbf{m} \sim 0.55$ . For instance even at the height of the potato famine in Ireland, the total decline in population over the decade 1841–1851 was about 20%, including migration and death (although of course a decade does not constitute a lifetime, so the figures are not strictly comparable). The largest migration occurring in a single year, the 17 million people that migrated from one part of the Indian subcontinent to another after partition, still only accounted for some 7% of the Indian population at the time. Modern radiochemistry is beginning to show that Neolithic populations may have been more mobile than previously thought (e.g. Montgomery et al. 2000). There are some interesting questions that could be addressed in this area using scenarios of the sort presented here, once accurate parameters can be estimated from the empirical data, such as why late Neolithic Malta became a cultural isolate after two millennia in contact with the rest of the Mediterranean (e.g. Robb 2001; Schulting & Richards 2002). Empirical studies to determine realistic values of parameter  $\mathbf{o}$  may be more problematic, but there may be some clues from anthropological field studies of cultural transmission (e.g. Hewlett & Cavalli-Sforza 1986).

When set to minimise memetic isolation, the GA simply identifies cultural selection as the overriding factor in this. As modelled here, contagious memes run to fixation within societies, and when transferred to neighbouring societies, do the same there. However, it is also clear that this is a naïve representation of the likely course of contagion within a society. Gatherer (2002a) shows that there are several factors that can derail 'the contagionist paradigm', and that random, or pseudo-random, walks are more likely than sigmoid curves when considering meme frequencies within any society.

The effect of modern mass communication media is to amplify the horizontal mode of transmission so that one individual can communicate with many more individuals simultaneously than would be possible in a lifetime of individual contacts (Cavalli-Sforza & Feldman 1981). This situation is not explicit modelled either here, or in SIM, but could easily be incorporated. Axelrod (1997) suggests a model where long-range interactions are permitted and probabilities of interaction are based on such long-range communication networks, rather than geography. Gould & Tornqvist (1971) describe a transition from a spatially contagious pattern of cultural spread to a hierarchical one dominated by rapid transfers between urban centres. These authors cite Pyle's (1969) work on cholera by way of example: the 1832 epidemic in Montreal and New York spread slowly down the Ohio/Great Lakes river system. In 1865 by contrast, several major cities, by that time connected by rail, were struck simultaneously, and the second wave of the epidemic struck towns connected to those major cities. Thus, for societies in the industrial phase of development, it is clear that such a spatial grid model (whether using SIM or the model presented here) is inadequate.

However, although the initial model was conceived as a representation of a geographical grid of societies, in order to compare the results of simulations with a geographically ordered

ethnographic data set (Gatherer 2002b), it is also possible to consider it more abstractly, as a single large society partitioned into smaller social networks. This effect is in any case forced on the scenario at high values of **m**, such as that produced by the GA. The arrangement of the cells on the grid does not then refer to any geographical separation of sub-societies, but to their separation in terms of the likelihood of weak ties. Sub-societies that are not adjacent on the grid will have no weak ties at all between them, and those that are adjacent will have weak ties with probability **n** per individual per generation.

In summary, it has been known for over a quarter of a century that memes have a greater tendency to homogeneity within populations than genes, and also that memes are more likely to be spatially clustered (reviewed by Cavalli-Sforza 1979). In the context of the present scenario, 'spatial clustering' can be taken to mean an absence of memetic isolation, or in Axelrod's terms the formation of 'cultural blocs'. The scenarios presented here demonstrate that, where societies are totally isolated (the "desert island" model, section 3.2) the tendency to memetic isolation is actually increased. However, where cultural exchange, even at a tiny level, is permitted between neighbouring societies, memetic isolation declines precipitately (the "Paleolithic" model, section 3.3). Where the meme is also under selection, especially cultural selection, memetic isolation falls towards zero (the "Neolithic" and "Contagion" models, sections 3.4 and 3.5).

Therefore the inherent tendency of memes to be more spatially clustered than genes depends on the existence of cultural bridges, or 'weak ties' (Granovetter 1973), between societies. The GA results confirms the primacy of selective pressures in creating spatial homogeneity, but also suggests that such homogeneity can be rapidly fractured by high migration rates. Memetic systems can be seen to be highly unstable, especially when the selective forces at work within them are also cultural. Culture is always close to chaos.

## Appendix: Code Listings

All listings require Perl to be installed, and should run under both windows and Unix. In the event of any problems, please contact the author.

The code for the graphical simulator:

As a perl file: sim.pl

As a text file: sim.pl.txt

The code for the GA:

As a perl file: ga.pl

As a text file: ga.pl.txt

## References

Ambrose SH (1998) Late Pleistocene human population bottlenecks, volcanic winter, and differentiation of modern humans. *Journal of Human Evolution* **34**, 623-651.

Axelrod R (1997) The dissemination of culture: a model with local convergence and global polarization. *Journal of Conflict Resolution* **41**, 203-226.

Barbujani G & Sokal RR (1990) Zones of sharp genetic change in Europe are also linguistic boundaries. *Proceedings of the National Academy of Sciences, USA* **87**, 1816-1819.

Bentley P (1999) (ed.). *Evolutionary Design by Computers*. Morgan Kaufmann, San Francisco.

Burnell K (1998) Cultural variation in savannah sparrow, *Passerculus sandwichensis*, songs: an analysis using the meme concept. *Animal Behaviour* **56**, 995-1003.

Cavalli-Sforza LL & Feldman MW (1973) Models for cultural inheritance. I. Group mean and within group variation.. *Theoretical Population Biology* **4**, 42-55.

Cavalli-Sforza LL & Feldman MW (1981) *Cultural Transmission and Evolution*. Princeton University Press, Princeton.

Cavalli-Sforza LL (1979) Cultural change and its relevance for human genetics. *Ciba Foundation Symposium* **66**, 5-23.

De Silvestri A & Guglielmino CR (2000) Ethnicity and malaria affect surname distribution in Consenza Province (Italy). *Human Biology* **72**, 573-583.

Durham WH (1991) *Coevolution: Genes, Culture and Human Diversity*. Stanford University Press, Stanford.

Gatherer D (2002a) The Spread of Irrational Behaviours by Contagion: An Agent Micro-Simulation. *Journal of Memetics - Evolutionary Models of Information Transmission* **6**.  
[http://jom-emit.cfpm.org/2002/vol6/gatherer\\_d.html](http://jom-emit.cfpm.org/2002/vol6/gatherer_d.html)

Gatherer D (2002b) Identifying Cases of Social Contagion Using Memetic Isolation: Comparison of the Dynamics of a Multisociety Simulation with an Ethnographic Data Set. *Journal of Artificial Societies and Social Simulation* **5**(4). <http://jasss.soc.surrey.ac.uk/5/4/5.html>

Geyer F (1994) Alienation, Participation and Increasing Societal Complexity. *Kybernetes* **23**, 1-3

Gonzalez Jose R, Garcia-Moro C, Dahinten S & Hernandez M (2002) Origin of Fueguian-Patagonians: an approach to population history and structure using R matrix and matrix permutation methods. *American Journal of Human Biology* **14**, 308-320.

Gould P & Tornqvist G (1971) Information, Innovation and acceptance. In: Hagerstrand T & Kuklinski A (eds.) *Information Systems for Regional Development*, Lund Series in Geography, Ser. B, *Human Geography* **37**, 149-168.

Granovetter M (1973) The strength of weak ties. *American Journal of Sociology* **78**, 1360-1380.  
Harpending HC, & al. (1998) Genetic traces of ancient demography. *Proceedings of the National Academy of Sciences, USA*, **95**, 1961-1967.

Hewlett BS & Cavalli-Sforza LL (1986) Cultural transmission among Aka pygmies. *American Anthropologist* **88**, 922-934.

Kittles RA, & al. (1999) Autosomal, mitochondrial, and Y chromosome DNA variation in Finland: evidence for a male-specific bottleneck. *American Journal of Physical Anthropology* **108**, 381-399.

Lasker GW & Crews DE. (1996) Behavioral influences on the evolution of human genetic diversity. *Molecular Phylogenetics and Evolution* **5**, 232-240.

Lynch A, Plunkett GM, Raker A & Jenkins P (1989) A model of cultural evolution of chaffinch song derived with the meme concept. *The American Naturalist* **133**, 634-653.

Montgomery J, Budd P & Evans J (2000) Reconstructing the lifetime movements of ancient people: a Neolithic case study from southern England. *European Journal of Archaeology* **3**, 370-386.

Pyle G (1969) Diffusion of cholera in the United States. *Geographical Analysis* **1**, 59-74.  
Ramirez-Duque P, Arends T & Merino F (1982) Chediak-Higashi syndrome: description of a cluster in a Venezuelan-Andean isolated region. *Journal of Medicine* **13**, 431-451.

Ray N, Currat M & Excoffier L (2003) Intra-Deme Molecular Diversity in Spatially Expanding Populations. *Molecular Biology and Evolution* **20**, 76-86.

Reader SM & Laland KN (2000) Diffusion of foraging innovations in the guppy. *Animal Behaviour* **60**, 175-180.

Refshauge WF & Walsh RJ (1981) Pitcairn Island: fertility and population growth, 1790-1856. *Annals of Human Biology* **8**, 303-312.

Renfrew C (1987) *Archaeology & Language: The Puzzle of the Indo-European Origins*. Jonathan Cape, London.

Robb J (2001) Island identities: ritual, travel and the creation of difference in Neolithic Malta. *European Journal of Archaeology* **4**, 175-202.

Rudgley R (2000) *Secrets of the Stone Age: A Prehistoric Journey*. Century, London.

Rychkov YG & Sheremeteva VA (1977) The genetic process in the system of ancient human isolates in North Asia. In: Harrison GA (ed.), *Population Structure and Human Variation*, Cambridge University Press, Cambridge, 47-108.

Schulting RJ & Richards MP (2002) The wet, the wild and the domesticated: the Mesolithic-Neolithic transition on the west coast of Scotland. *European Journal of Archaeology* **5**, 147-189.

- Scriver CR (2001) Human genetics: lessons from Quebec populations. Annual Review of Genomics and Human Genetics **2**, 69-101.
- Sokal RR, Oden NL, Legendre P, Fortin MJ, Kim JY & Vaudor A (1989) Genetic differences among language families in Europe. American Journal of Physical Anthropology **79**, 489-502.
- Sutherland WJ (2003) Parallel extinction risk and global distribution of languages and species. Nature **423**, 276-279.
- Thangaraj K, & al. (2003) Genetic Affinities of the Andaman Islanders, a Vanishing Human Population. Current Biology **13**, 86-93.
- Torrington M & Viljoen DL (1991) Founder effect in 20 Afrikaner kindreds with pseudoxanthoma elasticum. South African Medical Journal **79**, 7-11.
- Uyenoyama M, Feldman MW & Cavalli-Sforza LL (1979) Evolutionary effects of contagious and familial transmission. Proceedings of the National Academy of Sciences, USA **76**, 420-424.
- Van Schaik CP, & al. (2003) Orangutan cultures and the evolution of material culture. Science **299**, 102-105.
- Vivenes De Lugo M, Rodriguez-Larralde A & Castro De Guerra D. (2003) Beta-globin gene cluster haplotypes as evidence of African gene flow to the northeastern coast of Venezuela. American Journal of Human Biology **15**, 29-37.
- Vona G, & al. (1996) Genetics, geography, and culture: the population of S. Pietro Island (Sardinia, Italy). American Journal of Physical Anthropology **100**, 461-471.

# **MEMES AND THE PERSISTENCE OF ORGANIZATIONAL STRUCTURES**

**By**

**Hélène Giroux, Lecturer, Hautes Études Commerciales, Montréal QC Canada**

**James R. Taylor, Professor of Communication, Université de Montréal**

**François Cooren, Research Professor of Communication, University of Cincinnati**

Paper prepared for the *Symposium on Memetics:  
Evolutionary Models of Information Transmission*

15th International Congress on Cybernetics

Namur (Belgium), August 24-28, 1998

June 1, 1998

All return mail should be addressed to Professor Taylor at  
3051 Cedar Avenue, Montreal, Quebec, H3Y 1Y8, Canada

Telephone: 1-514-937-5069, Fax: 1-514-343-2298

email: either [taylor@com.umontreal.ca](mailto:taylor@com.umontreal.ca) or [ejtaylor@sympatico.ca](mailto:ejtaylor@sympatico.ca).

## **Abstract**

An impressive body of discourse analysis of organizational communication has added credence to the view that organization emerges in communication, but, as Nardi (1996) observes, the problem of explaining organizational persistence remains unresolved. In this paper the authors propose a model of organizational persistence as a memetic phenomenon. We argue that structures of meaning that describe action are composed of a limited set of bases (the equivalent of nucleotides): agents, events, objects, recipients, intentions, etc. These bases are then strung together to produce structured units describing narratively-based archetypes of action by means of which organization is both generated and understood. However, unlike DNA, such structures are single-stranded. We hypothesize that for them to replicate, the participation of other agents capable of playing a complementary role is necessary. We define such interaction as *conversation*. The equivalent of mRNA is thus the speaker of language, by means of which the patterns inscribed in the 'memome' of the speaker are 'read off' and produced as speech or written text. In linguistics this corresponds to the transposition of semantics into syntaxes. Conversation is thus the process of conversion of narrative into action, mediated by language, and constrained by the forms of conversation (as studied in Conversation Analysis). Similarly, textualization is defined as the process of conversion of action into archetypal narrative or "making sense" (Taylor et al, 1996). The function of interaction is to give structure to the context of action in which communication is embedded (including material objects that enter into action). The social equivalent of the cell is thus the organization which is composed of the set of self-reproducing systems. Communication thus both organizes context, and explains it. It is the sum of material, symbolic and social elements that compose the 'molecular soup' of the work environment that makes up organization, and not just communication. Nevertheless, communication is the carrier of structure, and explains organizational persistence.

## **Introduction**

An impressive body of discourse analysis of organizational communication has added credence to the view that organization emerges in communication (Taylor & Van Every, *in press*), but, as Nardi (1996) observes, the problem of explaining organizational persistence remains unresolved. In our attack on the problem, we take the structuration theory of Giddens (1976, 1979, 1984) as our starting point.

### **Structuration as an explanation of organizational persistence**

Giddens breaks with the established dualistic tradition of Western thought - the life forms of society and of the individual conceived as ontologically distinct levels of reality - by proposing a theory of *duality*, where both social and individual existences are co-constructed in the ordinary sphere of daily interaction. To accomplish this shift of perspective, he proposes to focus on the stream of activity which constitutes, without exception, the immediate experience of humanity (a version, perhaps, of Heidegger's concept of 'thrownness'). That which must then be explained, sociologically, is the "continuity of practice" transcending the strictly local and time/space situated (what we described above as "persistence") which clearly characterizes social forms of organization whose continuities extend over time, and across space. His central hypothesis is that this is achieved through reflexivity, i.e., "the monitored character of the ongoing flow of social life" (Giddens, 1984: 3). Human action occurs as an uninterrupted flow (that he calls a '*durée*'). It is experienced, however, as a structure of actions and reactions, each informed by human purpose. That which makes the continuous flow of conduct and cognition appear to be structured is what he calls "rationalization" (Giddens, 1984: 3). Such rationalization is "a routine characteristic of human conduct, carried on in a taken-for-granted fashion" (Giddens, 1984: 4) to produce the "encounters and episodes" that allow us to interpret life as motivated by intention and socially meaningful. That which explains persistence of organizational structure must thus be found within the rationalization process, as a central characteristic of it.

It seems evident to us (although this is not Giddens' primary emphasis) that the rationalization of which he writes is accomplished as an effect of the mediation of language in a context of communication. If the monitoring of activity and its rationalization to generate a motivated universe of action is achieved by its representation in forms of language, as cognition or conversation, then it is the monitoring-->rationalization process, as an effect of languaging, which explains persistence: That which persists is not the transient instantaneity of the flow but its *interpretation* or symbolic representation. Since rationalized experience becomes in turn how the situations that frame subsequent activity are experienced (Weick's 1979 concept of *enactment*) it is in this way that people are enabled to discriminate the structural continuities that enable social life, as they migrate from one time/space configuration to another. It is this that makes social (as well as cellular) life not just organized, but self-organizing (Taylor, 1995). Our hypothesis is that the rationalization process can be explained as a memetic phenomenon. There is in other words a homomorphism linking social organizational and organismic cellular processes of self-organization. We take this to be not an exact model, but a loose analogy, and it is this analogy we are exploring.

## The genetic analogy

The central unit of biological self-organization is the cell. The cell survives as an entity because it is self-regenerating (autopoietic). We propose that, by analogy, human social organization is similarly self-regenerating, and that the role played by gene structures in biological reproduction has a memetic analog in organizational persistence, although the analogy is far from one-to-one, in that while language presumably has structuring properties analogous to those of DNA its role is not identical, and the process by which its patterns are reproduced manifest significant differences from those studied by geneticists. The purpose of this paper is to outline, in a very preliminary way, some of the commonalities and divergences.

In exploiting the genetic model of cellular reproduction we inherit three concepts: replication, transcription and translation. In cellular reproduction, replication is accomplished by the splitting of the double strand of the gene to produce two identical copies. Transcription of the information encoded in the DNA requires the agency of mRNA which "reads" off the genetically encoded information and transfers it to the "manufacturing plant" of the cell where, assisted by enzymes, it is then "translated" into proteins. The proteins then become the active agents in sustaining the ongoing process. Each of these three stages, we propose, are also characteristic of the memetic cycle, even though the modes of accomplishment are by no means identical. We consider each of the phases in turn, beginning with transcription.

## Memetic transcription

There is, as is well known, a genetic syntax: Strings of DNA are composed of one of four bases (nucleotides), grouped into triplets (codons), organized into larger structures (genes), then into chromosomes and eventually a genome. But there is also a semantics of genes: Genotypes map (in complex ways) characteristics of the phenotype. The organism is constructed out of materials whose patterns are genetically encoded: A gene can for example be conceptualized as the recipe for a specific protein. In a general way, this correspondence also characterizes the role of language with respect to social organization. Language is also distinguished by a hierarchy of levels, from phonemic to grammatical to narratively informed units to text. At the level of strings of elementary symbols (beginning with the sentence) language encodes social structures: "I just bought myself a new winter coat" is not just a statement in language, it is an encoded organizational event that supposes the existence of buyers and sellers, markets and stores, and institutional arrangements that govern them. Language, at the level of semantics, distinguishes actors, agents, events, actions, instruments, objects, exchange dynamics, process, place and circumstance. And it does so in a way that constructs time as a structured unfolding: a pattern of interconnected events. It contains, in other words, all those elements that allow organization to be constructed, cognitively, as a lived reality - "rationalized." Such semantic patterns, we further propose, correspond to a semionarrative patterning of meaning: mirrored in the single utterances and sentences of speech, but realized in suprasential configurations having their basis in story syntaxes (Taylor, 1998; Cooren & Taylor, in press; Groleau & Cooren, in press; Robichaud, in press).

There is not space in this brief paper to develop this idea in full. Instead, we state as a working hypothesis that the semantics of language have precisely those properties of representation that

make Giddens' theory of reflexive monitoring and rationalization feasible. Experience, once textualized, is made meaningful. The "meme" is (and here the parallel with the gene is evident) that component of language which is capable of mapping a single feature of organization. Our hypothesis is that when it is read off, in a context of conversational exchange, it also provides the template for the *building* of the elements of organization, as a discursively realized phenomenon. This transcription is what we term 'actualization' (Robichaud, 1998). For such a hypothesis to be tenable, we would have to suppose that, in spite of the great variety of human spoken languages, the underlying semantic properties of all are the same. There is, fortunately, encouraging evidence to support this view that the semantics of language are characterized by patterns that are found universally in all the diverse tongues spoken by humankind (Bickerton, 1990).

It follows that the memetic equivalent of mRNA is speech. We assume that people, in speaking, draw on a repertoire of symbolic material - words, phrases, imagery, sayings, stories, analogies, lines of argument, etc. - that reflects stable properties of the speaker's semantic representation of the organizational world which formed him or her: He or she is transcribing memetically encoded patterns in the process of speaking. Again, given the constraints of this paper, we are obliged to enunciate this as a working hypothesis, without further defense.

## **Memetic translation**

Proteins are constructed by adjoining amino acids to form the complex patterns which distinguish one protein from another. We hypothesize that the "proteins" of social organization are patterned sequences of consecutive, inter-related acts, equivalent to what Schank and Abelson (1977) described as "scripts." These authors draw, as a concrete example, on what they call a "restaurant script" to show that the actions of both restaurateurs and clients fit together interactively to construct an event or episode, not because the spoken interventions of those involved are self-containedly meaningful, but because the indexical expressions used by participants ("Table for two, sir?", "Smoking, non-smoking?", "Bill?") is a recognizable transcription of components of the script. That which is being produced is not just speech, but a meaningful interaction that contributes to the persistence of a social institution, eating at a restaurant. The role of language is double: to encode the pattern of the event (in what we call its "text") and to mobilize the practical elements going into a meal - food, kitchen, building, check-out counter, etc. - by supplying the necessary procedural guidance by means of which the material substructure of dining is effectively organized. In speaking the actors are actually *acting out* the organizational structure (Giddens' point). By analogy, this is how all organizations are constructed.

There is, however, an obvious difference in the transcription processes of the cell and the human organization. In the cell, there is a direct transfer of the information encoded on one of the strands of the double helix, via the agency of mRNA, in the aligning of amino acids to form a given protein. In the organizational conversation, the production of a unit of organization necessarily supposes a mutuality of patterning: The restaurant customer and the waiter, for example, have complementary roles to play. The "double helix" of organization is thus constituted, not as a permanent binding of complementary strands, but as an occasioned event where the "code" of the organization will have been successfully realized if the occasioned 'eventing' is regularly accomplished.

## **Memetic replication**

The replication of the memetic pattern is thus intrinsically dependent on the dynamic of interaction. The persistence of cultural patterns, historically and geographically, gives empirical support to the idea that, in general, replication occurs, organizationally as well as biologically, with relatively constant fidelity. But it should be obvious that the process is contingent on situational variables, almost certainly to a greater degree than is the case for genetic replication. There is no reason to think, though, that the *principle* of replication is any less applicable to organizational as to organismic reproduction. All that is required is that the participants in the exchange find their model confirmed by their experience: The satisfied diner leaves with his or her 'restaurant meme' once more confirmed, and thus replicated; the same goes for the restaurateur. On the other hand, the model accounts easily for innovation. A meme that is never realized in interaction will presumably decay, while, on the other hand, the actor exposed to a new situation, in interaction with people who have an encoded map of its script, is likely to quickly learn the complementary exigencies, and thus the memetic representation of the organization is adapted to prepare the actor for subsequent encounters of a similar kind. (Note that the learner acquires, not the same code as the first, but the complement.)

## **The self-organizing properties of the system**

The internal dynamics of the cell include, of course, much more than the reproductive activities of the genetic system. That which explains the importance of the latter is that it accounts for the production of proteins, among which are to be found the enzymes that then become active agents in organizing the life of the cell and assuring the continuation of the reproductive cycle. We would argue that the function of communication, with respect to organization, is similar, in that it is by communication that the material preoccupations of organizational members are shaped, while at the same time the communication folds back onto itself, as Giddens proposes, to assure that the organizational pattern is replicated, and organization thus persists over time, and across space - the structuration hypothesis.

## **The organization is the cell**

Under this hypothesis the organization *is* the cell of social life. Society is thus multi-cellular. An issue that the limits of this paper do not allow us to consider is that of how to establish the boundaries of the organization (for elaboration see Taylor & Van Every, forthcoming).

## **Conclusion**

In Dawkins' initial treatment of memes he referred to them as "unit[s] of information residing in the brain" (p. 109) and described their propagation as "leaping from brain to brain via a process which, in the broad sense, can be called imitation" (p. 192). In a general way this is still how memes are principally conceived. Calvin (1996), for example, describes memes "as those things that are copied from mind to mind...from words to dances" and gives as examples advertising jingles and rumors (p. 18). Deacon (1997) similarly describes them as "bits of copied cultural information" (p. 115). In general, memes are seen to correspond to a cognitive pattern (or

"cerebral code" in Calvin's terms) and the spread of memes as a phenomenon of pattern cloning in the brain.

There are, as we see it, two principal problems with this way of characterizing the replication of memes. First, compared to genetic theory, it leaves the mechanism of imitation largely unspecified as a precise operation. It fails to explain why some patterns spread and some do not (an issue that has formed the matter of diffusion theory for more than a half century without producing any definite conclusion other than communicational). Second, it takes no account of the complementarity of patterns which, as the restaurant script illustrates, is an essential feature of cultural persistence. This latter omission may be traced to the dominance in memetic theory of considerations of cognitive dynamics and the relative neglect of communication. What we believe is that a reconsideration of the memetic *problématique* taking account of the role of language in communication provides both a more satisfactory explanation of the reproductive cycle, and a better account of the relation of memes to cultural persistence, as well as change.

It is obvious that such a theory will have to take account of non-linguistic as well as linguistic patterning: cooking and playing the violin as well as conversing. This does not seem to us an insuperable extension of the theory, but it is not one we have the space to consider here. Stated very generally, the issue is not of how memes "leap from brain to brain" but how they connect up with practice and thus become reinforced (or decay) by either having, or not having, an application. Any memetic pattern that is not, to use Weick's term, enacted is unlikely to survive. The spreading of a meme is thus not a "leap" but is inevitably mediated by practice.

## References

- Bicherton, D. (1990). Language and species. Chicago: University of Chicago Press.
- Calvin, W.H. (1996). The cerebral code: Thinking a thought in the mosaics of the mind. Cambridge, MA: MIT Press.
- Cooren, F. & Taylor, J.R. (In press). The institutional and rhetorical modes of the organizing dimension of communication: Discursive analysis of a Parliamentary commission. *Communication Review*.
- Dawkins, R. (1982). The selfish gene. Oxford: Oxford University Press.
- Deacon, T.W. (1997). The symbolic species: The co-evolution of language and the brain. New York: W.W. Norton.
- Giddens, A. (1976). New rules of sociological method. London: Hutchinson.
- Giddens, A. (1979). Central problems in social theory: Action, structure and contradiction in social theory. Berkeley and Los Angeles: University of California Press.
- Giddens, A. (1984). The constitution of society: Outline of the theory of structuration. Oxford: Polity Press.

Groleau, C. & Cooren, F. (In press). A socio-semiotic approach to computerization: Bridging the gap between ethnographers and systems analysts. *Communication Review*.

Nardi, Bonnie A. (1996). Studying context: A comparison of activity theory, situated action models, and distributed cognition. In B.A. Nardi, ed., *Context and consciousness: Activity theory and human-computer interaction*. Cambridge, MA: MIT press.

Robichaud, D. (1998). Au delà de l'action et de la structure: Traduction, réseaux d'actants et narrativité dans un processus de discussion publique. Doctoral dissertation, Department of Communication, University of Montreal.

Robichaud, D. (In press). Textualization and organizing: Illustrations from a public discussion process. *Communication Review*

Schank, R. & Abelson, R. (1977). *Scripts, plans, goals and understanding*. Hillsdale, NJ: Lawrence Erlbaum Associates.

Taylor, J.R. (1995). Shifting from a heteronomous to an autonomous worldview of organizational communication: Communication theory on the cusp. *Communication Theory*, 5 (1), 1-35.

Taylor, J.R. (1998). What is "organizational communication"? Communication as a dialogic of text and conversation. *Communication Review*.

Taylor, J.R., Cooren, F., Giroux, H. & Robichaud, D. (1996). The communicational basis of organization: Between the conversation and the text. *Communication Theory*, 6 (1), 1-39.

Taylor, J.R. & Van Every, E. (In press). The emergent organization: Communication as its site and surface. Mahwah, NJ: Lawrence Erlbaum Associates.

Weick, K.E. (1979). *The social psychology of organizing*. New York: Random House.

# MUTATION, SELECTION, AND VERTICAL TRANSMISSION OF THEISTIC MEMES IN RELIGIOUS CANONS

By

**John D. Gottsch**

The Johns Hopkins University School of Medicine  
Johns Hopkins Hospital, 600 N. Wolfe Street, Baltimore, MD 21205  
jgottsch@jhmi.edu

## Abstract

A study of ancient and modern Near Eastern religious canons reveals the mutation, selection, and vertical transmission of fitness-enhancing textual units, defined as theistic memes. The earliest recorded theistic memes dealt with human fear of death and defined man's earliest relationship to god. Theistic memes that could theoretically affect fitness through selection and incorporation into religious canons included those dictating beliefs about (a) self-awareness in an unknown world, (b) strategies and behaviors toward others and within the nuclear family, and (c) appropriate sexual behaviors within marriage. Prohibition of aberrant sexual practices such as incest, adultery, homosexuality, bestiality, castration, and religious prostitution would have further maximized fitness. A remarkable mutation of the ancient Near Eastern theistic meme of child sacrifice is documented in the Old Testament in the story of Abraham and Isaac. Vertically transmitted theistic memes in the Hebrew canon were largely incorporated into Christian and Muslim religious canons (New Testament and Qur'an). Mutations of theistic memes during vertical transmission into these other canons allowed the same fitness-enhancing stability for the gentile and Arabic populations and are notable for the different strategies used to produce homogenized, orthodox canons.

**Keywords:** Religious canon, fitness, meme, mutation

## 1 Introduction

Dawkins (1976) coined the term *meme* (from the Greek *mimeme*, meaning "imitation") to describe a unit of cultural transmission. A meme serves as a memory device that can be propagated from person to person and from generation to generation in the form of spoken or written words or by example. Memes include ideas, knowledge, and concepts. Blackmore (1999) has further defined the meme as having properties that convey memory, can randomly vary, and are subject to selection. Heylighen (1998) has outlined the selection criteria necessary for memes to be successfully propagated from generation to generation. In the first step of meme propagation, *assimilation*, the meme must be assimilated and accepted by a host through infection (exposure) and must be somewhat plausible. In the next step, *retention*, the meme must be retained by the host through memory. With this step meme selection occurs. Without memory, a meme is extinguished and cannot proceed to the next step, *expression*, when infection of another host may occur. Retention of a meme depends on the perceived usefulness, interest, and relevance to the host. The last step is *transmission*, the physical means by which the meme

can be conveyed to the next potential host. Transmission may occur through physical gestures (hand or facial movements), speech, the spoken or written word, or pictures.

Memes that convey dictated social behavior can be categorized as either secular or theistic. According to Dawkins (1976), the *god meme* concept "provides a superficially plausible answer to deep and troubling questions about existence." Blackmore (1999) recognizes that religions contain memes that regulate certain behaviors, including sexual practices and some cooperative behaviors, and can modify aggression. Operationally, in most religious canons the god meme consists of a number of explicit pronouncements and commandments purportedly attributed to a god. In this study, these specific attributes and dictated behaviors are called *theistic memes*.

The word canon is derived from the Greek *kanon*, meaning "reed" or "cane," a device used by the ancients to ensure correct measurements in masonry. The term came to signify an act of setting standards, and by the fourth century implied a list of orthodox texts (Mellor 1972). Although the term was first used to describe the Christian literature, canonization can be documented in the earliest traditions of religious texts. A process in which those exercising authority in a population select and exclude material in written texts, thus developing a non-contradictory religious canon with coherent theistic memes, has occurred since the time of the ancient Babylonians. In general, theistic meme selection is attenuated over time as the text hardens into inviolable canon. Here, *religious canon* is defined as consisting of theistic memes faithfully transcribed from generation to generation. In the genesis of a canon or orthodox literature, after an initial period of competing theistic memes one group eventually prevails over others. Theistic memes may *mutate* as accidental or intentional alterations. Accidental alterations include misspellings, deletions, or repetitions. Ehrman (1993) has categorized intentional alterations as harmonization of the text with parallel passages, elimination of grammatical errors, or the smoothing out of ambiguities. However, in some instances the author of the alterations has the sole intention of creating text to render a canon more orthodox and does this by importing biased memes or by modifying or eliminating memes with contrary or competing views.

## 2 First Evidence of Memes

Schmandt-Besserat (1986) described the first definitive evidence of the creation of tangible objects, or *tokens*, as memetic devices designed to convey ideas among fellow humans. Tokens were used widely throughout the Near East from 8000 BCE onward. During this era of nascent agrarian cultures, societies required means of keeping records and indicating weights and measures to deal with the realities of increased food stores. Tokens were made from fired clay in simple geometric shapes -spheres, discs, cylinders -and used as mnemonic devices representing units of tangible goods. These tokens are the first evidence of tangible memes created to convey messages over distance and time.

Around 4000 BCE, more sophisticated memes were developed as *complex tokens*, clay figures with inscriptions indicating some abstract idea. Complex tokens originated at about the same time and location as Sumerian writings in the form of *cuneiform text*, some of which have survived to this day. The oldest writing identified to date (around 3500 BCE) is on a small limestone tablet excavated from the ancient city of Kish (Finegan 1979). Pictographs of a human head, hand, and foot and a threshing sledge appear on both sides of the tablets. The message

inscribed is as yet undecipherable. One of the first deciphered texts on a Sumerian writing tablet describes a hero named Enmerkar who ruled the city Uruk and lamented for a world of the past in which there was no fear (Kramer 1972a, 1972b).

*In those days there was no snake, there was no scorpion, there was no hyena,  
There was no lion, there was no wild dog, no wolf,  
There was no fear, no terror,  
Man had no rival.*

This, one of the oldest recorded memes discovered to date, expresses the human belief in an imagined past when man was without anxiety, without fear, and without enemies who would do him harm. These Sumerian memes, preserved in clay, were meant to be read by others and were likely understood by many succeeding generations.

### 3 First Theistic Memes in Religious Canons

#### 3.1 *Epic of Gilgamesh*

The earliest theistic memes to pass from generation to generation over a large territorial area are those found in the *Epic of Gilgamesh*. Tigay (1982) found the origins of the story in third millennium Sumeria; it was propagated to the old Babylonian empire (2000 BCE), the Hittite and Hurrian empires in the Middle Babylonian period (1400 BCE), and the Assyrian empire (700 BCE). This epic in its basic form, passed as a written document on clay tablets, was most probably read to the illiterates of the population, who then passed on the story in some oral tradition. Because the story was committed to a written form and was copied faithfully by learned scribes, the epic has been preserved for thousands of years. The *Epic of Gilgamesh* is a remarkable collection of theistic memes vertically transmitted over the centuries as a story of human fear of death and how humans should cope with death anxiety.

Gilgamesh, after the death of his best friend, Enkidu, finds himself crying "bitterly like unto a wailing woman." For seven days and nights he weeps over his friend, not permitting his burial. Grief stricken, he becomes obsessed with the fear of death: "When I die, shall I not be like unto Enkidu?" He eventually decides that he wants to find immortality and sets out to discover the dwelling place of Utnapishtim, the survivor of the Great Flood, who will be able to tell him how to achieve eternal life.

At the edge of the sea, Gilgamesh receives advice about his quest from Siduri, the divine barmaid.

*The life you pursue you shall not find,  
When the gods created mankind.  
Death for mankind they set aside.  
Life in their own hands remaining,  
As for you, Gilgamesh, let your belly be full,  
Make merry day and night.  
Of each day make a feast of rejoicing,*

*Day and night dance and play! Let your garments be sparkling fresh,  
 Your head be washed; bathe in water,  
 Pay heed to a little one that holds your hand.  
 Let a spouse delight in your bosom,  
 For this is the task of a woman.*

Gilgamesh eventually finds Utnapishtim, who recounts in detail the Great Flood, his survival, and the subsequent gift of immortality. Utnapishtim reveals the secret of the gods: at the bottom of the sea is a plant that gives eternal life. Gilgamesh departs in a boat, dives to the bottom of the sea, and retrieves the plant. But a serpent rises up from the water, smells the fragrance of the plant, and eats it. Gilgamesh weeps bitterly, but he realizes that nothing can be done to obtain immortality.

The Gilgamesh epic is a collection of ancient Near Eastern written theistic memes that expressly deal with questions about the fear of death and the afterlife. The instructions of the barmaid are a set of memes defining social strategies by which man is supposed to live his life. These memes are explicit: "The life you pursue you shall not find, / When the gods created mankind. / Death for mankind they set aside . . . / Pay heed to a little one that holds your hand. / Let a spouse delight in your bosom."

Oral traditions are believed to have preceded the written memes of the *Epic of Gilgamesh*. Tigay (1982) has documented the canonization of the Epic by comparing surviving clay tablets from different locales and chronological eras (Table 1). Gilgamesh, the Sumerian king, is believed to have lived around 2700 bce. Historical evidence suggests that Gilgamesh rebuilt a shrine to the god Enil at Nippur and constructed a wall at Uruk. The earliest texts concerning Gilgamesh date to 2100-2000 bce, and these were likely based on oral memes transmitted from the time of his reign. The original Sumerian tales included six non-interrelated compositions describing the true Sumerian hero king, Gilgamesh (Kramer 1972a, 1972b). The composition and meter of the tales suggest that they may have served as hymnic doxologies used in temples as part of hero worship of an ancient king. At this time Gilgamesh was considered part god. Although each tale had separate plots, three of the tales concerned Gilgamesh's fear of death and his wish to find eternal life.

Passage	Change from	Change to	Meme mutation type
Gilg. P. i, 7/GE 1,v,28	Descended upon me	Keeps descending	Synonyms or words functioning similarly
/GE II, ii, 40	The populace gathered around him	The populace <i>jostles toward</i> him	
Gilg. P. vi, 14/GE II,iii,47	He did not allow Gilgamesh to enter	He <i>does not</i> allow Gilgamesh <i>to be brought in</i>	Different grammatical and lexical forms of the same word
Gilg. P. ii, 12/GE I, iv, 35	Wild creatures (fem)	Wild creatures (masc)	
Gilg. P. ii, 16/ GE I, iv, 37	To the pure temple, abode of Anu	To the pure temple, abod of Anu <i>and Ishtar</i>	Added words or phrases
Gilg. P. i, 15,	The mother of	The mother of Gilgamesh <i>is</i>	

37/GE I, v, 39; vi, 16	Gilgamesh, knowing all	<i>wise, knowing, she knows all</i>	
Gilg. P. I, 10/GE I, v, 31-34	Uruk-land was gathered about it	Uruk-land stands <i>about [it]</i> <i>[The land is gathered] around it, [The popula] ce [jostles] t[oward it], [The men ma]ss against it</i>	Expansion by parallelism
(Gilg. P. ii, 14-18)/(GEI,iv,36-37)	Come let me lead you to Broad-martred Uruk, to the pure temple, the abode of Anu, Enkidu, arise, let me direct you to Eanna, the abode of Anu	Come, let me <i>direct</i> you to Uruk (of?) the <i>sheepfold</i> , to the pure temple, the abode of Anu and Ishtar	Telescoping of parallel lines
Gilg. Mi, iii, 9/GE X, ii, 16-17	Show (me) the path [...]	<i>[Now], barmaid, what is the road to Utnapish[tim]?[What is] its [/]andmark!</i>	Reformulation with new idea added
Gilg. P. i, II/GE I, v, 35	The men Kiss his feet	<i>[Like a ba]by, an [in]fant, they kiss his feet</i>	
Gilg. Y.iv, 2/GE II, v, 2, 5; cf. GEH rev. vi, 12	Seven terrifying haloes	a terror <i>to people...</i>	Reformulation with meaning changed completely
Gil. P.Y.vi, 15-16/GE II, ii, 48	They grappled with each other, like champions (lit. victors)they bent the knee	They grappled with each other <i>in the gate of the marital chamber</i>	
C, I, from Gilg. P. ii, 16, 18 and GE 1, iv 37, 44	“Abode of Anu and Ishtar”	“Abode of Anu”	Change due to Religious Ideology: <i>The deliberate omission of Ishtar in the Old Babylonian version is related to an antipathy toward Ishtar/Inanna’s unfavorable portrayal reflected in the Bull and Heaven episode.</i>

**Table 1:** Theistic Memetic Mutations in the Gilgamesh Epic from Old Babylonian to Late Version (Adapted from Tigay 1982)

The transformation of the Sumerian tales into a religious canon began in the Old Babylonian period (2000-1600 BCE). The broad memetic story line of the Sumerian tales was creatively adapted into an integrated and well-meshed text with a definitive message. The central theistic memetic message was Gilgamesh's anguish over the death of a friend, his own fear of death, and

his search for immortality. The proto-theistic memes of the Sumerian tales were mutated. Enkidu changes from a servant to a friend, thereby increasing the emotional impact of his death on Gilgamesh. Added to the story were the all-important theistic memes of ancient Sumerian wisdom sayings, admonishing Gilgamesh to concentrate on his present earthly life, with a woman and children, and not to worry about immortality.

The Old Babylonian *Epic of Gilgamesh* spread throughout the ancient Near East in the Late and Middle Babylonian period. The canon was found in the original Akkadian and translated into Hittite and Hurrian, at which point the text was subjected to contemporary editing. These textual mutations were mostly in the form of unintentional scribal errors. However, some of the redaction included grammatical and lexical changes, addition of new words or phrases, or expansion of some ideas by parallelism. Some changes in the text included the addition of new ideas and some reflected changing religious ideology (Table 1). These mutations were incorporated into the text, which thereafter became standardized canon. The theistic memetic message of the *Epic* was probably relevant to populations throughout the ancient Near East, as it was widely distributed and considerable effort was put into preserving the text by meticulously copying it onto tablets.

The story of the Great Flood has ancient roots. Many early texts referred to the event(s) (Table 2), and the actual flood most likely occurred in the early third millennium. Two prominent early accounts were the Sumerian *Deluge* and the Akkadian *Atrahasis Epic*. Neither was incorporated into the Old Babylonian version. However, The *Atrahasis Epic* continued to be transcribed as independent canon until late in the first millennium (around 1250 BCE), when it was incorporated nearly verbatim into the *Epic of Gilgamesh*, greatly expanding the Utnapishtim flood story.

	<i>Gilgamesh Epic</i>	<i>Old Testament</i>
Author	Enlil, the Sumerian god of Babylonia.	God
Reason	In (Tablet XI:14) the heart of the great gods prompted them to bring a deluge	In (Gen. 6:1-13) The earth was corrupt before God and was filled with violence because of man, for all flesh had corrupted its way upon the earth.
Hero	The Sumerian account calls him Ziusudra meaning "he who laid hold on life of distant days." Other versions call him Atrahasis meaning "the exceedingly wise."	Noah meaning "rest.". (Gen. 6:9).
Announcement	Ea appears to the Utnapishtim telling him to abandon his possessions and to build a ship with certain specifications and take into it every seed of life. (Tablet XI:19-31). This revelation was made contrary to Enlil's plan and without his knowledge. (Tablet XI:173).	In (Gen 7:1), God communicates with Noah. "Go into the ark, you and your whole family, because I have found you righteous in this generation."

Period of Grace.	There was no thought of granting mankind an opportunity to repent. According to Atrahasis Epic man was granted several periods of grace before Enlil determined to destroy humankind by means of a flood.	Gen. 6:3, Man was granted period of grace one hundred and twenty years.
Ark and Dimensions	Elippu, meaning "a great vessel," "boat," "ship." Utnapishtim's boat , length height, and width were 120 cubits, seven stories.	Ark is "teba" related to the Egyptian db't meaning, "box," "chest," "coffin." Noah's ark was "three hundred cubits, its breath fifty cubits, and it's height thirty cubits" (Gen. 6:15). Noah's ark had 3 stories.
Occupants	Utnapishtim loaded aboard all he had of "the seed of all living creatures," " the game of the field, the beasts of the field, all the craftsmen," (Tablet XI: 80-85 and 94-95).	Gen. 6:19-21: " of every living thing, of all flesh, thou shalt bring two of every into the ark to keep them alive with thee; they shall bea male and a female. Of the birds....." " And Noah, with his sons, his wife, and the wives of his sons, went into the ark, because of the waters of the flood."
Flood Begins	Tablets silent on this point.	"In the sixth hundredth year of Noah's life , in the second month, on the seventeenth day of the month, on that very day all the fountains of the great deep (tehom) were broken open, and the windows of the heavens were opened" (Gen. 7:11).
Cause	The destructive forces listed in the Summerian tablet are amuru, meaning "rainstorm," rain flood," or "cloudburst," and mighty winds. These two elements accompanied by thunder and lightning are mentioned in the Gilgamesh Epic.	Torrential rains from the heaven and the eruption of the subterranean waters. (7:11-12).
Duration	Rained in the evening followed by a storm that lasted for six days and six nights.	For forty days and forty nights it rained upon the earth. (Gen 7:11)
Storm magnitude	"As soon as the first shimmer of the morning beamed forth, a black cloud came up from out the horizon."	" The waters increased and lifted the ark, so that it rose above the earth.....the waters grew exceedingly strong upon the earth; and all the mountains under the whole heaven were covered."
Ark Landing	Mount Nisir, " mount of Salvation."	" on one of the mountains of Ararat."

Bird Scene	Utnapishtim released a dove for testing the subsidence of the water on the seventh day after the landing at Mount Nisir.	Forty days after the tops of the mountains had become visible, Noah opened the window of the ark and sent forth a raven. (Genesis 8:5-7). Having waited seven more days, Noah released another dove, which did not come back until toward evening. (8:10-11) After another seven days, Noah sent out a third dove, but she never returned (8:12) Also this was a good sign; It showed that the lowlands as well as the mountains were free of water
Exit from the Ark	"sent forth to the four winds"	God said to Noah, "Go forth of the ark, thou and thy wife, thy sons and the wives of thy sons with thee. Bring forth with thee every living thing that is with thee.....and be fruitful and multiply on the earth" (8:16-17)
Sacrifice	Utnapishtim prostrated himself before the sun-god and offered up an Ox and an abundant sacrifice of sheep.	Noah built an alter unto the Lord and offered burnt-offerings "of every clean beast and of every clean fowl" (8:20).
Divine Blessings	Removal from the ken of mortal man	Man may not be slain in impunity either by man or by beast (Gen. 9:1-7)

Table 2: Theistic Memes of Flood Story in the Gilgamesh Epic and Old Testament (Adapted from Heidel 1963)

### 3.2 Enuma elish

The Gilgamesh memes did not deal with questions about how gods and humans were created. These were recorded in a second set of written theistic memes found throughout the ancient Near East, called the *Enuma elish* (When above) myth. The memes incorporated in this myth likely date to the rise of the first Babylonian dynasty (1894-1595 BCE) when the god Marduk was legitimized, especially under the reign of Hammurabi (1792-1750 BCE), the most powerful of the Babylonian kings.

The creation story opens with a description of a time when there was nothing but the divine spirits, the parental gods, Apsu and Tiamat, and their son, Mummu. Each represents cosmic matter, different forms of water (fresh water, salt water, and clouds) that mingled together and from which light emanated and the universe was made. After several generations of gods, the great wise god, Marduk, is born; he becomes supreme by slaying the evil gods and creating man. This complicated story raised as many questions as it answered. However, as a memetic device it provided "an answer" to questions about creation that endured for centuries. Ultimately, the myth became logically untenable and was abandoned, but the basic memetic structure survives to this day in the Old Testament.

	<i>Enuma Elish</i>	<i>Passage</i>	<i>Genesis</i>
Tablet VII: 86	Divine Spirit and cosmic matter are coexistent and coeternal	Gen. 1:1-2	Divine spirit creates cosmic matter and exists independently of it
Tablet VI: 23-30	Primeval chaos; Tiamat enveloped in darkness	Gen. 1:2	The earth a desolate waste, with darkness covering the deep(tehom)
Tablet I: 38	Light emanating from the gods	Gen. 1:3	Light created
Tablet I: 59-72	The creation of the firmament	Gen. 1:6	The creation of the firmament
Tablets IV: 143-45	The creation of the dry land	Gen. 1:9	The creation of dry land
Tablet V: 1	The creation of the luminaries	Gen. 1:14-18	The creation of the luminaries
Tablet VI: 1-38	The creation of man	Gen. 1:26	The creation of man
Tablet VII	The gods rest and celebrate	Gen. 2:2	God rests and sanctifies the seventh day

**Table 3:** Theistic Memetic Structures in Enuma Elish Creation Myth and Old Testament Genesis (Adapted from: Heidel 1963)

A superficial comparison of the Hebrew Genesis and the Babylonian *Enuma elish* might suggest the two stories are different in style, complexity, logic, and maturity. However, Heidel (1942) demonstrated that when the internal memetic structures (story lines) are outlined, they prove to be identical (Table 3). The memetic influences of the *Enuma elish* were also found to have influenced the Canaanite population, with the creation of Baal paralleling the genesis of Marduk. This parallelism is also evident in the stories that deal with the struggles of Tiamat and the struggles of Yahweh in slaying the primordial evil monsters. These two texts are likely memetic homologues.

*When thou shalt smite Lotan, the fleeing serpent, And shalt put an end to the tortuous serpent, Shalyat of the seven heads.*  
(Ras Sharma Tablets)

*On that day the Lord will punish with His sword, which is hard and great and strong, Leviathan, the fleeing serpent, and Leviathan, the tortuous serpent, And He will slay the crocodile that is in the sea.*

(Isa. 27:1)

And theistic memes as hymns of praise for Baal were most likely memetically mutated in form and context to become Psalms for the Hebrew God, Yahweh.

*Behold, thine enemies, O Baal;  
Behold, thine enemies thou shalt smite.  
Behold, thou shalt destroy thine adversaries.  
(Ras Sharma Tablets)  
For, behold, thine enemies, O Lord,  
For, behold, thine enemies shall perish!  
All the workers of iniquity shall be scattered!*  
(Ps. 92:10)

### 3.3 Law Codes

The third widely distributed written tradition of memes passed from generation to generation in ancient Near Eastern cultures consisted of sets of law codes. Some codes date to the Sumerian period, the laws of Ur-Nammu (2112-2095 BCE). Not all the Ur-Nammu law codes were preserved, and the number and subject and the punishments decreed by the king cannot be known.

The law code memes were presented as divinely inspired and were preceded by a prologue establishing the authority of the king and his relationship to the gods. Thus the laws had a basis and a credibility in the population as enforced by the gods. These are therefore theistic memes. An epilogue details punishments for those who chose to disobey or deface the inscribed codes. In the code of Hammurabi, the chief god, Enlil, proclaims

*When lofty Anum, king of the Anunnaki, and Enlil, lord of heaven and earth, the determiner of the destinies of the land, determined for Marduk, the first-born of Enki, the Enlil functions over all mankind, made him great among the Igigi, called Babylon by its exalted name, made it supreme in the world, established for him in its midst an enduring kingship, whose foundations are as firm as heaven and earth--at that time Anum and Enlil named me to promote the welfare of the people, me Hammurabi, the devout, god-fearing prince, to cause justice to prevail in the land, to destroy the wicked and the evil, that the strong might not oppress the weak, to rise like the sun over the black-headed people and to light up the land.*

The memetic law codes of Eshnunna and Hammurabi were directly incorporated into the *Covenant codes* of the Torah (Mellor 1972, Patrick 1985). Mellor (1972) has noted some thirty-five to fifty Covenant memetic codes that are directly related to the Hammurabi memetic codes. Finkelstein (1981) has provided convincing evidence of direct meme transfer from Near Eastern codes to the Hebrew laws dealing with the issue of the goring ox. Early agricultural populations depended on oxen for heavy labor. The ox, although useful for many tasks, was also unpredictable and capable of inflicting injury by goring its owner or others. The people of the ancient Near East understood that some oxen were prone to goring innocent bystanders and that the owners of these oxen should be aware of this and should take appropriate protective measures. Owners who did not do so were guilty of "culpable negligence," a concept that ancient Near Eastern populations thought worthy of setting into theistic memetic laws to be passed from

generation to generation. These laws are notably found in three codes, the Eshnunna, the Hammurabi, and the Covenant.

Oxen memetic laws of Eshnunna, 1850 BCE:

*If an ox has gored another ox and caused its death, the owners of the oxen shall divide between them the sale value of the living ox and the carcass of the dead ox.*

*If an ox was a habitual gorer, the local authorities having so duly notified its owner, yet he did not keep his ox in check and it then gored a man and caused his death, the owner of the ox shall pay two thirds of a mina of silver to the survivors of the victim.*

*If it gored a slave and caused his death, he shall pay fifteen shekels of silver.*

Oxen memetic laws of Hammurabi, 1792-1750 BCE:

*If an ox, while walking along the street, gored a person and caused his death, no claims will be allowed in that case.*

*But if someone's ox was a habitual gorer, the local authority having notified him that it was a habitual gorer, yet he did not have its horns screened nor kept his under control, and that ox then gored a free-born man to death, he must pay one-half mina of silver.*

*If the victim was someone's slave, he shall pay one-third mina of silver to the slave's owner.*

Oxen memetic laws in Covenant code (Exod. 21:28-32), 1200-1000 BCE:

*If an ox gores a man or a woman to death, the ox shall be stoned to death, its flesh may not be eaten, but the owner of the ox is innocent.*

*But if the ox was previously reputed to have had the propensity to gore, its owner having been so warned, yet he did not keep it under control, so that it then killed a man or a woman, the ox shall be stoned to death, and its owner shall be put to death as well. Should a ransom be imposed upon him, however, he shall pay as the redemption for his life as much as is assessed upon him. If the ox gore a slave or slavewoman, he must pay thirty shekels of silver to his owner, but the ox shall be stoned to death.*

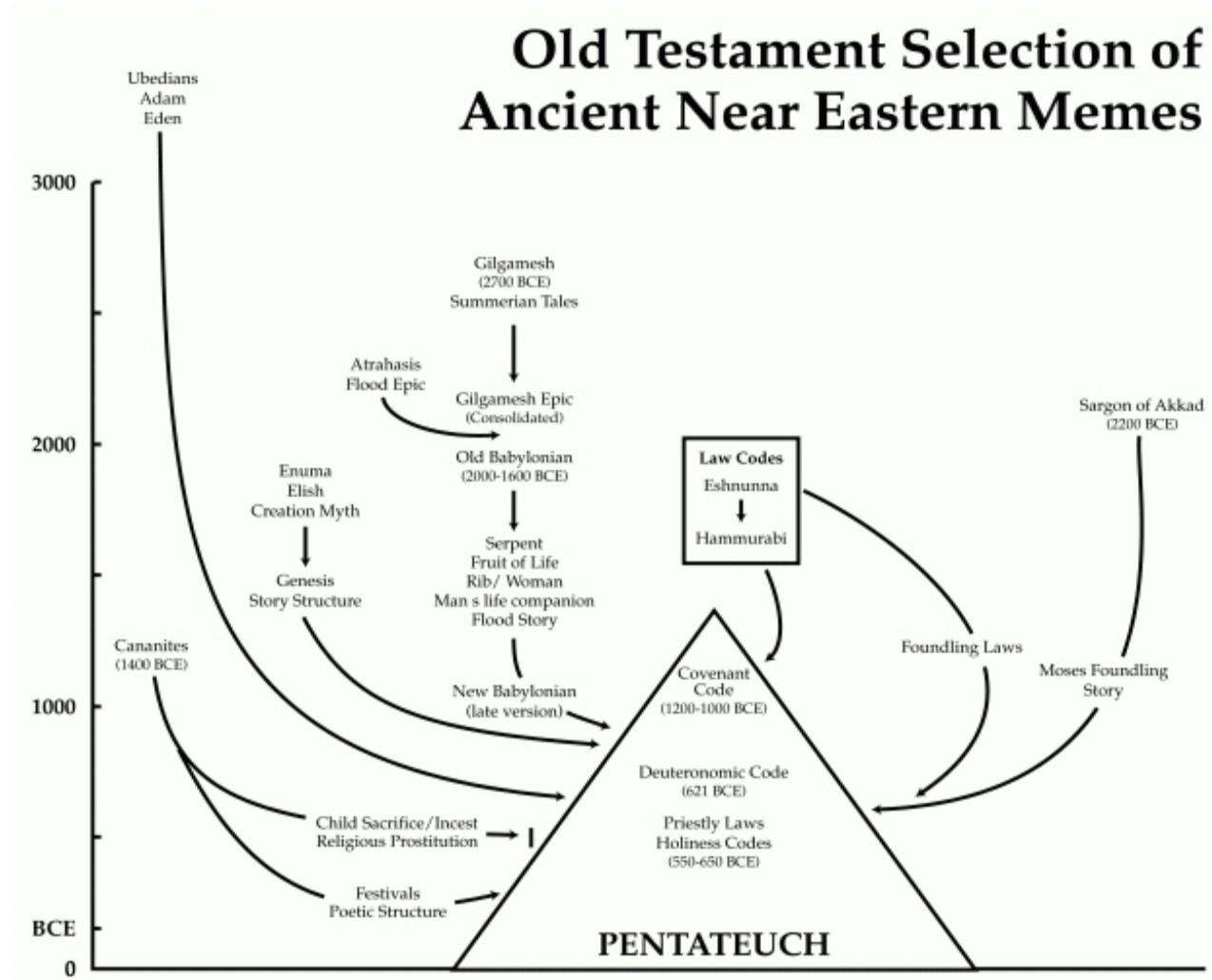
*If an ox belonging to one man gores to death the ox of his fellow, they shall sell the live ox and divide the proceeds, and they shall divide the dead one as well.*

*But if the ox was previously reputed to have had the propensity to gore, and its owner had not kept it under control, he shall make good ox for ox, but will keep the dead one for himself.*

The interrelationships among these texts are unmistakable. Given the identical memetic expressions and situations, it is highly likely that the laws of Eshnunna and Hammurabi were directly related to or drew upon the same prototypical material. The Hebrew writers likewise

undoubtedly knew of the culpable negligence meme and the written Mesopotamian memetic codes, those concerning oxen in particular, when mutating these laws to craft the Covenant code.

#### 4 Theistic Memetic Creation of the Hebrew Canon



**Table 4:** Old Testament Selection of Ancient Near Eastern Memes

A number of theistic memes taken from Mesopotamian canon and other traditions were selected and incorporated into the Hebrew canon, the Pentateuch (Table 4). The advice of the Siduri to Gilgamesh found in Tablet X of the Old Babylonian version is incorporated into Ecclesiastes 9:7-11 as:

*Go then eat your bread in happiness and drink your wine with a cheerful heart; for God has already approved your works.*

*Let your clothes be white all the time, and let not oil be lacking on your head.*

*Enjoy life with the woman whom you love all the days of your fleeting life which He has given to you under the sun; for this is your reward in life and in your toil in which you have labored under the sun.*

*Whatever your hand finds to do, do it with all your might; for there is no activity or planning or knowledge or wisdom in Sheol where you are going.*

*I again saw under the sun that the race is not to the swift and the battle is not to the warriors, and neither is bread to the wise nor wealth to the discerning nor favor to men of ability; for time and chance overtake them all.*

(Segal in press)

Hamblin (1987) has described what may be the oldest memetic story, Adam in the Garden of Eden (probably of pre-Sumerian in origin), as representing man's transition from hunter-gatherer to agriculturist some 7000 years ago. The terms *Adam* and *Eden* have been found in cuneiform texts from early Sumerian times and even the earlier Ubadian period. *Eden* means fertile plain; *Adam* means settlement on the plain. God's banishment of Adam and Eve from the Garden of Eden has been proposed to represent humans' choice of a life of husbandry and agriculture over their earlier, God-ordained life as hunter-gatherers, thus going against "God's will." The "original sin" may in fact have been the choice to grow our own food. The serpent that presents the gift of knowledge is no doubt a memetic mutation of the sea serpent that steals the immortality-giving plant in the *Epic of Gilgamesh*. Also taken from the *Epic of Gilgamesh* (previously the Atrahasis epic) is the theistic meme of the flood story; importantly, Noah becomes the memetic figure Utnapishtim. Heidel (1970) has shown the remarkable and unmistakable parallels between these two stories, including the characters, the ark, and the sequence of events (Table 2). Another likely theistic meme taken from the *Epic of Gilgamesh* and incorporated into the Genesis canon is the concept of a man's life companion. The author of the *Epic* chose a man, Enkidu, for Gilgamesh's companion. The author of the Hebrew Genesis mutated this memetic concept such that a woman becomes a man's life companion.

Of great importance to the Torah is the story of Moses. His early life in the Egyptian court is recounted in detail in Exodus. However, the story of a foundling who is rejected, exposed to danger, and found and nurtured and then grows to adulthood to achieve greatness was a literary meme first used to describe the early life of Sargon of Akkad (2371-2316 bce) (Table 4). Sargon was one of the first great conquerors of history. As Finegan (1979) recounts the legend, Sargon's mother puts him in a basket made of rushes and floats it down a river. The boy is pulled out of the water by a woman named Akki, who raises him in the court of the king Ur-Zababa. Sargon breaks away and establishes his own kingship, eventually capturing and uniting the Sumerian and Akkadian kingdoms. Childs (1965) has pointed out that Sargon of Akkad initiated the foundling child legal principles, articulated in the earliest of the ancient Near Eastern law codes and later incorporated into the Old Testament. These principles included statements of (a) the individual involved, (b) length of time designated by the contract, (c) condition of work, (d) specifications about nourishment, (e) fines for breach of contract, (f) amount of wages, and (g) witnesses. The narrative of the Moses foundling story incorporates these ancient Near Eastern legal memes (except for the breach of contract clause) in this order.

*Then said his sister to Pharaoh's daughter, Shall I go and call to thee a nurse of the Hebrew women, that she may nurse the child for thee? And Pharaoh's daughter said unto her, take this child away and nurse it for me, and I will give thee wages. And the woman took the child and nursed it. And the child grew, and she brought him unto Pharaoh's daughter, and he became her son.*

(Exod. 2:7-10)

The Hebrew writers were not unlike the ancient Near Eastern authors of the *Epic of Gilgamesh*, who understood that a text could be revised and reinterpreted to edify and reconcile new historical and theological events. The Hebrews' examination and probing of these new events to find relevant precedence was called *Midrash*. Midrash was employed in the development of new text as a creative inspiration based on a knowledge of previous texts and traditions (Goulder 1974). Midrashic activity was rationalized by the Hebrew belief that one passage could be used to illuminate another, whether a simple or a complex passage. Deuteronomy is essentially a Midrash of the Covenant code; Chronicles is a Midrash of Kings. Thus in essence, Midrash was the process of theistic memetic mutation that allowed a successful meme to be varied and incorporated into new text. If a new text was accepted as canon, the incorporated, mutated theistic memes could (and frequently did) affect reproductive fitness. Thus the Hebrew literature is an interwoven fabric of theistic memetic historical facts, traditions, beliefs, and rituals. These texts were often related to others by Midrash, or the incorporation of theistic memetic mutations.

The overriding Hebrew theistic meme was that all humans are created in God's image and thus all are created equal (Dimont 1962, Johnson 1988). This theistic meme was distinct and unique in ancient Near Eastern traditions because of its emphasis on human life and the rights of individual believers -a concept that contrasted with other contemporary cultures in which property rights and the rights of royalty were paramount. In the goring oxen memes of the Hebrews, the emphasis on punishment conveyed the importance of human life, whereas in the analogous Babylonian culpable negligence memes the emphasis was on property rights.

The fully developed Jewish text of laws, the Torah or Pentateuch, in its complete form consists of 613 theistic memes, faithfully preserved to this day (*Encyclopaedia Judaica* 1971). These theistic memes accrued over a number of centuries to cover a vast array of codes of conduct defining behavior and beliefs. The codes were interpreted literally, without a well-defined belief in the afterlife, until the first century bce, when several sects developed alternative memes about how the law could be reinterpreted.

Jewish religious canon evolved in three distinct phases, the Covenant code (1200-1000 bce), as discussed above, the Deuteronomic code (700-621 bce), and the Holiness code (587-450 bce). The Deuteronomic code was likely developed in the century before its fabled discovery in the Jerusalem temple during Josiah's reign (621 bce). The Deuteronomic laws were intended to further restate and redefine Hebrew monotheism and to rid the religion of foreign influences by actively repudiating the customs and practices of the Canaanites and the Egyptians. The Canaanite civilization was eventually subjugated, with the land and people incorporated into the Hebrew state. Some of the religious practices of the Canaanites were incorporated into the religious beliefs of the Hebrews, such as worship of the Canaanite gods and the practice of child

sacrifice. The Deuteronomic code amply documents later Hebrew writers' attempts to repudiate these Canaanite cultic practices related to child sacrifice, fertility, and religious prostitution.

*You shall not behave thus toward the Lord your God, for every abominable act which the Lord hates they [the Canaanites] have done for their gods; for they even burn their sons and daughters in the fire to their gods.*

(Deut. 12:31)

*When you enter the land which the Lord your God gives you [Canaan], you shall not learn to imitate the detestable things of those nations. There shall not be found among you anyone who makes his son or his daughter pass through the fire, one who uses divination, one who practices witchcraft or one who interprets omens, or a sorcerer.*

(Deut. 18:9-10)

Child sacrifice must have markedly reduced the fitness of the Canaanites. This behavior was clearly "imitated" by the Hebrews and would have markedly reduced their fitness as well.

*They [the Israelites] did not destroy the peoples, as the Lord commanded them, but they mingled with the nations and learned their practices, and served their idols, which became a snare to them. They even sacrificed their sons and their daughters to the demons, and shed innocent blood, the blood of their sons and their daughters, Whom they sacrificed to the idols of Canaan.*

(Ps. 106:34-37)

The Hebrew practice was later seen as an "abomination" in the eyes of God: "They [Israelites] built the high places of Baal that are in the valley of Ben-hinnom to cause their sons and their daughters to pass through the fire to Molech"(Jer. 32:35). King Josiah destroyed this child sacrificial alter, called a Topheth: "[King Josiah] also defiled Topheth, which is in the valley of the son of Hinnom, that no man might pass his son or his daughter through the fire for Molech"(2 Kings 23:10). And Jeremiah (7:30-32) records,

*For the sons of Judah have done that which is evil in my sight declares the lord, they have set their detestable things in the house which is called by My name to defile it. They have built the high places of Topeth, which is in the valley of the son of Hinnom, to burn their sons and their daughters in the fire, which I did not command, and it did not come into My mind. Therefore, behold, days are coming, declares the Lord, when it will no more be called Topeth, or the valley of the son of Hinnom, but the valley of the Slaughter; for they will bury in Topeth because there is no other place.*

Ultimately, Jewish fitness was enhanced by the absolute prohibition against child sacrifice. No religious practice, except suicide and prohibition against heterosexual sex, could limit fitness more. Human child sacrifice must be a non-genetically driven behavior, a practice dictated by religious mandate. Development of a genetically driven human behavior that would waste such an extraordinary amount of parental investment is improbable. The practice probably evolved as a phenomenon of self-awareness that imagined a vengeful god who would unleash the most horrible of unknowns if not appeased with the ultimate sacrifice -one's own child. The meaning of the story of Abraham and Isaac has baffled many. How could any father be so willing to

sacrifice his own son, even if asked by God? However, in the times of the ancient Near Eastern Canaanites and Israelites, sacrifice of a child was the only way to truly gain favor and to spare the wrath of God. As Maccoby (1982) concludes,

*Perhaps the most interesting story [about human sacrifice], however, is the story of Abraham and Isaac, for here we find not complete transformation, made from an anti human-sacrificial stand-point, but a phase transition, in which the yearning for human sacrifice is still struggling with the desire to abolish it. The purpose of the story is to show that God Himself ordained that animal sacrifice should be substituted for human sacrifice. At the same time, the story contains no moral revulsion from the very idea of human sacrifice. On the contrary, it is imputed to Abraham as extraordinary merit that he was willing to sacrifice his favourite son, Isaac, at the behest of God. We see here the dynamics of the historic move from human to animal sacrifice: on the one hand, this is a revolutionary step, by which a higher morality is brought into effect; on the other, the benefits of human sacrifice cannot be lightly relinquished, and the transition from human to animal sacrifice must appear plausible in the sense that animal sacrifice must acquire the same aura of reverence and holiness that previously belonged to human sacrifice.*

In effect, the story of Abraham and Isaac represents a transition of the Hebrews to a more fit religious population through the memetic mutation of their religious canon to exclude the fitness-reducing practice of child sacrifice.

After the Babylonian exile, the Hebrew canon was further infused with fitness-enhancing memes that came to be known as the *Holiness code*. The overall effect of the Holiness code was to make the Hebrew population more a religious community than a nation-state. Although many of the theistic memes were related to ritual and sacrifice, a substantial number were direct and explicit sexual prohibitions. Many Canaanite sexual practices (incest, bestiality, and prostitution) were detrimental to maintaining the nuclear family and most likely contributed to the population's decline in fitness. In contrast, the Holiness code prohibited incest, with the most inclusive set of theistic memes to counter consanguinity. Laws also prohibited adultery, homosexual sex, and bestiality. These Jewish theistic memetic laws greatly enhanced fitness and were enforced by the most explicit (attributed to God) theistic memetic punishments.

*Defile not ye yourselves in any of these things: for in all these the nations are defiled which I cast out before you: And the land is defiled: therefore I do visit the iniquity thereof upon it, and the land itself vomiteth out her inhabitants. That the land spew not you out also, when ye defile it, as it spewed out the nations that were before you. For whosoever shall commit any of these abominations, even the souls that commit them shall be cut off from among their people. Therefore shall ye keep mine ordinance, that ye commit not any one of these abominable customs, which were committed before you, and that ye defile not yourselves therein: I am the Lord your God.*

(Lev. 18:24, 25, 28-30)

Here, human behaviors are dictated not by genes but by mental constructs backed up by another mental construct, the fear of God.

At the core of the fitness-enhancing Jewish canon are other unique theistic memes that allowed all individuals of the religious population to exercise and to be governed by the same behavioral memes. Collectively, the newly constituted theistic memes, as the *Laws of Israel*, facilitated the reproductive success of the entire population. Jewish theistic memes provided more credible answers than did other Near Eastern genesis stories to self-awareness questions such as: Where did I come from? Who is God? What happens when I die? When my family dies? Is there life after death? On this platform of increased credibility were established other theistic memes that facilitated reproductive success. Decreasing conflict among individuals was chief among the effects of these memes. The Hebrew God commanded, "Thou shalt not hate thy brother in thine heart: thou shalt not in any wise rebuke thy neighbor, and not suffer sin upon him. Thou shalt not avenge, nor bear any grudge against the children of thy people, but thou shalt love thy neighbor as thyself. I am the Lord" (Lev. 17, 18). The overall strategy had clearly shifted to cooperation for the whole population. Other theistic memes stabilized the nuclear family through strict laws on obedience to and respect for one's parents and the maintenance of the family through strict observance of religious ritual. Other remarkable theistic memes in the Jewish canon promoted the reproductive success and fitness of the population. The Law states that Jewish men and woman should marry and that as husband and wife they should have children.

## 5 Theistic Memetic Mutation in Creation of the Christian Canon

The *Sadducees* were an authoritarian group of Jewish priests and the well-to-do who proposed that all Law was to be written and that a strict and literal adherence to the Law was necessary to be holy and to be a Jew. The Law was not to be changed; it was fixed for all time. Because the Torah contained few specific memes regarding an afterlife, the Sadducees ignored the subject. In contrast, the *Pharisees* embraced the idea of an apocalypse and an afterlife, as recounted by Isaiah, and the concept that those who had fulfilled the Law and were denied compensation in this earthly life would be compensated in the next. The Pharisees proposed that if a person could not fulfill the Law literally because of penury or physical circumstances, then executing the spirit of the Law was sufficient. The *Essenes* were an ascetic sect on the fringes of Jewish society. They believed in resurrection, immortality of the soul, and a messiah who would lead the way to immortal salvation. The wicked would be punished in hell and the good would live with God in heaven. The Essenes developed elaborate purification rites, including baptism, in which new members were admitted into a "New Covenant" with the remission of sins. The messiah, the "Teacher of Righteousness," preached penitence, poverty, and love of one's neighbors. In this Jewish Essene church, the essential rite was the sacred meal. The Essenes believed in celibacy, yet permitted marriage. Most new members were adopted from other sects and indoctrinated in the faith.

A fourth Jewish sect was the *Christians* (Johnson 1988). Jesus of Nazareth, baptized by John the Baptist, an Essene, radically changed the interpretation of the Law: no longer were the historical context and roots of the Law important. Fulfillment of the Law, in fact, was comparable only to good works -just a precondition for eternal salvation. Salvation itself depended on faith and belief in Jesus as the Son of God, the fulfillment of the prophecy of the Jewish messiah.

The theistic memes concerning Jesus Christ were that (a) he was the Jewish messiah, (b) he was crucified, and (c) he rose from the dead. The initial community that embraced Jesus as the new Messiah was Jewish and based in Jerusalem. Distinguishing the growth of this church from other

Jewish sects were memes that would be spread by a few gentile believers in Antioch. In a momentous decision by the ruling council of Jewish Christians in Jerusalem, gentiles were accepted into the faith without having to fulfill the initiation requirements to become a Jew (Acts 15; all references to the New Testament are to the Revised Standard Version.). Thus the uncircumcised gentile and the observant Jew were equal members in this new religion. Because the religion was so appealing to pagans in the Hellenized world and to the diasporic Jews, many became converts. Soon the vast majority of Christians were gentiles and, when the Jerusalem temple was destroyed by the Romans in 70 ce, all formal links to Judaism were severed. However, the Christian church was always to retain the central theistic meme, the belief in the God of Israel.

The Christians kept the theistic memes associated with the Hebrew God, Yahweh, as creator of the universe, earth, and humanity. In Matthew 5:17-20, Christ teaches his disciples:

*Think not that I have come to abolish the law, or the prophets: I am not come to abolish them but to fulfill them. For truly, I say unto you, until heaven and earth pass away, not an iota, not a dot, will pass from the law until all is accomplished. Whoever then relaxes one of the least of these commandments, and teaches men so, shall be called least in the kingdom of heaven, For I tell you, unless your righteous exceeds that of the scribes and Pharisees, you will never enter the kingdom of heaven.*

Also retained were the theistic memes detailing Israelite history and the interpretations of this history by the Israelite God. These included the history of Moses and the giving of the Law and the history of the prophets and the kings. From the large amount of Jewish memetic law, the Christians appropriated the memes related to human equality and equal justice; theistic memes restricting holiness to those who could provide the elaborate, sacrificial rituals were selectively minimized. In Matthew 23:23 Christ expounds, "Woe to you, teachers of the law and Pharisees, you hypocrites! You give a tenth of your spice-minth, dill and cumin. But you have neglected the important matters of the law - justice, mercy and faithfulness. You should have practiced the latter, without neglecting the former."

Thus the theistic memes to be embraced and acted upon in order to achieve holiness were reduced in number, simplified, and in effect made available to any individual. The Christian theistic memes included behavior commandments attributed to Jesus (You must love your fellow human as yourself) and to his proselytizers. The Lord's prayer is a central Christian theistic meme, which appeals to the Jewish God, our Father, with adoration; the will of God is accepted, with a request for the basic necessities of life and a plea for forgiveness and the promise to forgive. Many Christian religions were to evolve from the initial followers of Christ in Jerusalem. The Church of Rome, the Byzantine or Eastern Orthodox Church, and the Protestant Church are some of the major Christian populations that adopted mutant theistic memes dealing with differences in beliefs, rituals, and behaviors.

Spong (1996) has catalogued the same process of theistic memetic mutation (Midrash) in the development of Christian canon. The Gospel writers, in attempting to legitimize the new Jewish messiah, used the Midrashic technique to ground Christ firmly in the prophecies and traditions of the Scriptures, or Old Testament. Many examples are evident (Table 5). Christian Midrash also

included the alteration of existing Gospel texts by later Gospel authors to harmonize and further legitimize Christ. The earliest written Synoptic Gospel was that of Mark. In writing of the parentage of Jesus, Mark does not mention an earthly father, referring to Jesus thus: "Is not this the carpenter, the son of Mary" (Mark 6:3). The lack of any reference to a father in any part of this Gospel would have opened Jesus to charges of being born a bastard. For Jews who considered Jesus the messiah this would have been a major obstacle: "A bastard shall not enter into the congregation of the Lord" (Deut. 23:2). The modifications or mutations made by Matthew when he copied Mark's text are obvious and transparent. When Matthew incorporated Mark's text on the parentage of Jesus, the text is transformed from "Is not this the carpenter, the son of Mary" into "Is not this the carpenter's son, whose Mother is called Mary" (Matt. 13:55). Thus, in this remarkable mutation of a theistic meme, Joseph is created as Jesus' father.

<i>Crucifixion</i>				
Passage	Old Testament Verse	Passage	New Testament Verse	Significance
Psalm 22 :1	My God, my God, why hast thou forsaken me? Why art thou so far from helping me, from the words of my groaning?"	Mark 15 :34 Matt. 27 :46	'And at the ninth hour Jesus cried with a loud voice, "E'lo-I, E'lo-i, la'ma Sabach-tha'ni?" which means, "My God, my God, why hast thou forsaken me?"' 'And about the ninth hour Jesus cried with a loud voice, "Eli, Eli, la'ma sabach-tha'ni?" that is, "My God, my God, why hast thou forsaken me?"'	Midrashic description of the crucifixion based on Psalm 22. Nearly every element of the Psalm is used to depict the crucifixion seen. Here the cry, "My God, my God, why has thou forsaken me has been lifted word for word from the Psalm.
Psalm 22 :7-8	'All who see me mock at me, they make mouths at me, they wag their heads; "He committed his cause to the LORD; let him deliver him, let him rescue him, for he delights in him!"'	Mark 15: 29 Matt. 27 :43	'And those who passed by derided him, wagging their heads, and saying, "Aha! You who would destroy the temple and build it in three days,' 'He trusts in God; let God deliver him now, if he desires him; for he said, "I am the Son of God."'	Here the concept of someone divine in dire consequences mocked by others to save himself has been copied and mutated from Psalm 22.
Psalm 22 :17	'I can count all my bones--they stare and gloat over me;'	John 19 :32-36	'So the soldiers came and broke the legs of the first, and of the	Jesus, who had come to symbolize the perfect sacrifice as the

			<p>other who had been crucified with him; but when they came to Jesus and saw that he was already dead, they did not break his legs.'</p> <p>... 'For these things took place that the scripture might be fulfilled, "Not a bone of him shall be broken."</p>	<p>paschal lamb, had to be physically perfect as an offering and could not have any broken bones.</p>
Psalm 22 :15	' my strength is dried up like a potsherds, and my tongue cleaves to my jaws; thou dost lay me in the dust of death.'	John 19 :28	'After this Jesus, knowing that all was now finished, said (to fulfill the scripture), "I thirst."	The theme of thirst is continued to parallel Psalm 22.
Psalm 22 :18	'they divide my garments among them, and for my raiment they cast lots.'	John 19 :23-24	'When the soldiers had crucified Jesus they took his garments and made four parts, one for each soldier; also his tunic. But the tunic was without seam, woven from top to bottom; so they said to one another, "Let us not tear it, but cast lots for it to see whose it shall be." This was to fulfill the scripture, "They parted my garments among them, and for my clothing they cast lots."	The text of Psalm 22 is further developed with the division of clothes and the casting of lots. The symmetries among the stories is now complete

### *Virgin Birth*

Psalms 2 :7	'I will tell of the decree of the LORD: He said to me, "You are my son, today I have begotten you.'	Romans 1 :4	'and designated Son of God in power according to the Spirit of holiness by his resurrection from the dead, Jesus Christ our Lord.'	The earliest of scriptural writings from Paul which designate Jesus as God's son, a title that had been used for the Kings of Israel
-------------	---	-------------	--	--

Isiah 7:14	Behold a young woman shall conceive and shall bring forth a son and you shall call his name Emmanuel, which means God is with us.	Matt 1:23	Behold a virgin shall conceive and shall bring forth a son and you shall call his name Emmanuel, which means God is with us.	Midrashic use of Isaiah 7:14 by Matthew to signify Jesus' virgin birth. Matthew did not realize that he had copied from a Greek translation that had improperly used "virgin" for "young woman".
Isaiah 42:1	'Behold my servant, whom I uphold, my chosen, in whom my soul delights; I have put my Spirit upon him, he will bring forth justice to the nations.'	Mark 1:11	'and a voice came from heaven, "Thou art my beloved Son; with thee I am well pleased."	The use of Midrash by Mark to signify the inauguration of Christ's as the "Son of God" to occur at his baptism
Psalm 110:4	'The LORD has sworn and will not change his mind, "You are a priest for ever after the order of Melchiz'edek."	Hebrew 7:1, 17	'For this Melchiz'edek, king of Salem, priest of the Most High God, met Abraham returning from the slaughter of the kings and blessed him;'  'The LORD has sworn and will not change his mind, "You are a priest for ever after the order of Melchiz'edek."	Because Jesus was not part of the recognized priesthood, a Midrashic attempt was made to link Jesus with another scriptural figure who was not of the regular Jewish priesthood, Melchizedek.
<b><i>Passion</i></b>				
Zechariah 9:9-11	'Rejoice greatly, O daughter of Zion! Shout aloud, O daughter of	Matt 21:7-9	And [the disciples] brought the ass, and the colt, and put on them their clothes, and	Misdrashic interpretation of Zechariah's account of the King entering

	Jerusalem! Lo, your king comes to you; triumphant and victorious is he, humble and riding on an ass, on a colt the foal of an ass.		they set him theron....And the multitudes that went before, and that followed, cried, saying Hosanna in the highest. And when he was come into Jerusalem, all the city was moved, saying Who is this? And the multitude said, this is Jesus the prophet of Nazareth of Galilee.	Jerusalmem. The story was muted to convey the entry of Jesus in the Holy City.
Zechariah 14:21	'and every pot in Jerusalem and Judah shall be sacred to the LORD of hosts, so that all who sacrifice may come and take of them and boil the flesh of the sacrifice in them. And there shall no longer be a trader in the house of the LORD of hosts on that day.'	Matt 21:12-13	And Jesus went into the temple of God, and cast out all them that sold and bought in the temple, and overthrew the tables of the moneychangers, and the seats of them that sold doves, and said unto them, It is written, My house shall be called the house of prayer; but ye have made it a den of thieves.	Another story from Zechariah that was modified to convey the idea that Jesus had cast out the moneychangers and cleansed the temple.
<b><i>Judas</i></b>				
Zechariah 11 :13	Then the LORD said to me, "Throw it to the potter, that magnificent price at which I was valued by them" So I took the thirty shekels of silver and threw them to the potter in the house of the Lord..	Matt. 27 :3-7	'When Judas, his betrayer, saw that he was condemned, he repented and brought back the thirty pieces of silver to the chief priests and the elders, saying, "I have sinned in betraying innocent blood." ... And throwing down the pieces of silver in the temple, he departed; ..And they conferred	Lifting of the story from Zechariah, with acknowledgement, of thirty pieces of silver to be used for the burial plot for the poor.

			together and with the money bought the Potter's Field as a burial place for strangers.	
Psalms 41:9	' Even my bosom friend in whom I trusted, who ate of my bread, has lifted his heel against me.'	John 13 :26	'Jesus answered, "It is he to whom I shall give this morsel when I have dipped it." So when he had dipped the morsel, he gave it to Judas, the son of Simon Iscariot.'	The Midrashic use of the story from Psalm 41 of betrayal during the act of eating.
2 Samuel 20: 9-10	"And Joab said to Amasa, Art thou in health, my brother? And Joab took Amasa by the beard with the right hand to kiss him, But Amasa took no heed to the sword that was in Joab's hand; so he smote him therewith in the fifth rib, and shed out his bowels to the ground,..and he died.	Mark 14 :44, 45	'Now the betrayer had given them a sign, saying, "The one I shall kiss is the man; seize him and lead him away under guard." And when he came, he went up to him at once, and said, "Master!" And he... kissed him.'	The act of giving a kiss to signify an act of betrayal lifted from 2 Samuel is used to further flesh out the story of Judas.

**Table 5:** Theistic Memetic Mutation Old Testament to New Testament (Spong 1996)

In another remarkable theistic memetic mutation, Matthew wishes to anchor the story of Jesus' divine birth in midrash. He chose a passage from a Greek translation of Isaiah 7:14: "Behold a virgin shall conceive and shall bring forth a son and you shall call his name Emmanuel, which means 'God is with us.'" Matthew did not realize or chose to ignore that the word *virgin* had been mistranslated from the original Hebrew text. Properly translated the Isaiah verse reads, "Behold a young woman shall conceive and shall bring forth a son and you shall call his name Emmanuel, which means 'God is with us.'" Thus, the virgin birth story, like the story of Joseph as father, was the creation of a mutated theistic meme. Other examples of extensive midrashic alteration of texts are found in the depiction of the crucifixion, the story of Judas, and the Passion (Table 5).

In a study of the chronological relationships and textual criticisms of early Christian religious texts, Ehrman (1993) has revealed interrelationships and internal corruptions that demonstrate a remarkable selection process leading toward a uniform and non-contradictory religious canon. The confusion created with Jesus' virgin birth and the question of the nature and timing of his divinity led to the development of a number of mutant theistic memes. All Christians believed

that Christ was the Son of God. But what did this mean? New "Christians" wondered, "Was he born God or did he become God at his baptism or did he become God at his crucifixion and death?" Some early Christians believed that Christ was born a man then was "adopted" by God at the time of his baptism and became divine.

Anti-adoptionist theistic mutations by scribes wishing to embrace the orthodox position focused on several key issues in Jesus' life and ministry. The anti-adoptionists wished to support the belief of a divinely inspired virgin birth and to minimize any suggestion that Jesus had a real father (Table 6). In nearly every theistic meme originally written to convey that Joseph was Jesus' father, the text was mutated to enforce the emerging orthodox position that Christ was the Son of God and born of a virgin.

Passage	Change from	Change to	Significance
<b><i>Anti-Adoptionist</i></b>			
Luke 2:33, 48 Luke 2:43 Luke 2:48	"father and mother began to marvel" "his parents" "Look your father and I have been grieved, searching for you."	" <i>Joseph</i> and his mother began to marvel." " <i>Joseph and his mother</i> " " <i>Your relatives</i> and I have been grieved, searching ...."	Jesus the unique Son of God. Orthodox affirmation of the Virgin Birth and that Joseph was in fact not Jesus' father.
Luke 3:22,/ Mark 1:11	"You are my Son, today I have begotten you"	"You are my <i>beloved</i> Son, <i>in you I am well pleased.</i> "	Orthodox opposition to an adopted Jesus. Supports the view that Jesus was the Son of God at birth.
Mark 1:1	." "The beginning of the Gospel of Jesus Christ."	"The beginning of the Gospel of Jesus Christ, <i>the Son of God</i>	Added corruption that supports the orthodox view that Jesus was the Son of God at birth.
<b><i>Anti-Separationist</i></b>			
Mark 1:11/Luke 3:22/Matthew 3:16	" the spirit as a dove descending unto him"	"the spirit as a dove <i>upon</i> him"	Jesus, the Christ at His baptism. Divinity of Jesus at his baptism.
Mark 15:34	"My God, my God, why have you forsaken me?"	" My God, my God why have you <i>reviled</i> me."	Jesus not left behind as the crucified Christ is resurrected. Text harmonized with Psalm 22.
Hebrews 2:9	"that he apart from God should taste death for every man"	"that he <i>by the grace of God</i> should taste death for every man."	An orthodox corruption to obscure an interpretation that Christ did not die on the cross
<b><i>Anti-Docetic</i></b>			
Luke 22:19-20	"And taking bread,	"And taking bread,	Orthodox interpolation that

	giving thanks, he broke it and gave it to them saying, This is my body. But behold, the hand of the one who betrays me is with me on the table"	giving thanks, he broke it and gave it to them saying, This is my body <i>that is given for you. Do this in my remembrance. And likewise after supper, saying, this cup is the new covenant in my blood that is poured out for you.</i> But behold, the hand of the one who betrays me is with me on the table"	Christ died for atonement of sins and that real blood was shed for the sins of the world to affirm the reality of his body at the time of the crucifixion.
Luke 22:43-44	Saying, Father, if thou be willing, remove this cup from me: nevertheless not my will, but thine, be done. And when he rose up from prayer, and was come to his disciples, he found them sleeping for sorrow.	Saying, Father, if thou be willing, remove this cup from me: nevertheless not my will, but thine, be done. <i>And there appeared an angel unto him from heaven, strengthening him. And being in an agony he prayed more earnestly: and his sweat was as it were great drops of blood falling down to the ground.</i> And when he rose up from prayer, and was come to his disciples, he found them sleeping for sorrow.	Orthodox interpolation that counters the docetic belief that Christ could not suffer and was not a human of true flesh and blood.
Luke 24: 11-12	"And their words seemed to them as idle tales, and they believed them not"	"And their words seemed to them as idle tales, and they believed them not. <i>Then arose Peter, and ran unto the sepulchre; and stooping down, he beheld the linen clothes laid by themselves, and departed, wondering in himself at the which was to come to pass.</i> "	Orthodox interpolation that Christ's literal body had been raised from the dead and this was recognized by the chief disciple and was not a silly tale by hysterical women.

**Table 6:** Theistic Memetic Mutations in Early Christian Canon (Ehrman 1993)

Marcion of Pontus (second century CE) was the son of an orthodox bishop of the church of Sinope. After disagreements with his father about church doctrine, he was forced to leave the

church. He eventually became a wealthy ship owner and once again tried to influence the church. This time he went to Rome and, after making a large donation to the Church, attempted to establish a dominant Christian canon. Marcion seized upon the writings of Paul as a rejection of Jewish Law and scripture, and he proposed a canon consisting of Pauline letters with Luke's gospel, purged of Hebrew scriptural references. So complete was the rejection of the Jewish God that Marcion believed Christ was not descended from the Old Testament creator god but was sent by a "stranger" god. Since Christ was not of the Creator God, he could not have been born and thus could not have been a flesh-and-blood human being. Marcion's views thus expressed the basic tenets of *Gnosticism*. When presented to a council of the Church of Rome, his work was squarely rejected and his donation was returned; he was again excommunicated. Undaunted, Marcion traveled through Asia Minor, where his message and canon were well received. His movement was so successful throughout the Mediterranean region that the Church was provoked to combat the perceived heresy. In reaction to Marcion's teachings and "canon," the Church of Rome convened councils to formulate an "orthodox" canon. The formulation of this canon again involved a conscious "corruption" or mutation of theistic memes in accepted religious texts: the Gospels (Mark, Matthew, Luke, John), Acts, and the Letters of Paul. These mutations countered other interpretations, both anti-adoptionist and Gnostic, from rival churches.

Gnostics believed that Jesus the man and Christ the divine were not the same. Two basic arguments defined Gnosticism. The *Docetics* believed Christ was always divine and only assumed the semblance of a man while on earth. To counter the Docetics, orthodox scribes inserted passages in the early canon that made Jesus appear as a man of real flesh and blood (Table 6). The *Separationists*, the Gnostic majority, believed the divine Christ descended from heaven to inhabit the human being called Jesus at his baptism and gave the saving knowledge (*gnosis*) during Jesus' ministry; being only a temporal inhabitant, the divine Christ left Jesus' body at the crucifixion. To counter this perceived heresy, orthodox scribes mutated text that suggested Jesus and Christ were separate before his baptism or before his death on the cross (Table 6). These variants were selected and eventually would prevail, establishing the orthodox view. Each variant interpretation had profound effects on the final canon. The Christ embodied in orthodox canon was born divine of a virgin birth; he was both God and a man during his ministry; and after crucifixion his body was resurrected.

## 6 Theistic Memetic Mutation in Creation of the Islamic Canon

The basic Islamic theistic memes are present in the Qur'an, as given by the Prophet Muhammad in a series of revelations by God over a period of ten years (622-632 ce). Like Jesus, Muhammad was reinterpreting the Jewish Torah and its laws (theistic memes) for his people. But unlike Christ, Muhammad was not a Jew. Although Muhammad knew and respected the "People of the Book," he formulated for the Arabs a monotheistic religion with altered Jewish theistic memes. He claimed the god of the Jews for his religion and Abraham as the first true believer. Any theistic memes attributed to other prophets (including Moses and Christ) that Muhammad did not like were deemed false and discarded. The revelations to Muhammad were eventually transcribed as written theistic memes in the Qur'an - the indisputable and authoritative source for all future followers.

The five structural theistic memes essential to the faith of Muslims are (a) the *shahada*, a public profession of faith: "There is no god but Allah and Muhammad is the Messenger of God"; (b) the *salat*, the performance of a ritual prayer; (c) the *zakat*, obligatory alms; (d) sawm, fasting during the month of Ramadan; and (e) the *Hajj*, a pilgrimage to Mecca. These obligatory behaviors are the Pillars of Islam. The Qur'an is the undisputed word of God, and all believers are referred to this source for any questions about appropriate rituals and behaviors. If the memes in the Qur'an are silent or ambiguous on a particular topic, the memes in the *hadiths*, the sayings or reported sayings of Muhammad (collected in the *Sunna*), are consulted. If the *hadiths* are contradictory, the memetic laws of Islamic jurists (the *ijma*) are used to produce uniform behavior. Examples of the Prophet's sayings in the *hadiths* include

*Every Muslim has six obligations toward his fellow Muslim: he greets him whenever they meet; answers his call; wishes him well when he sneezes; visits him when he is ill; follow in his funeral when he dies; and wishes for him what he wishes for himself.*

*The strong man is not the one who knocks people down; the truly strong man is the one who can control himself in anger.*

*God has no mercy for those who have no mercy for their fellow men.*

These Islamic theistic memes, directly related to those of the Old Testament, prescribe how individuals should treat each other. They are in effect Islam's Golden Rules of behavior. In contrast to what we know about the Old and New Testaments, our knowledge about the canonization of the word of God according to Muhammad is remarkably complete. God reportedly revealed his words directly to Muhammad. The verses and phrases were recorded on separate pieces of bone and parchment and, at Muhammad's death (610 ce), three successive caliphs gathered, compiled, and authenticated the text. The Caliph Uthman (644-656 ce) proclaimed an official version of the Qur'an, and all others were deemed false and ordered destroyed. This contrasts with Hebrew and Christian authors spending centuries composing and editing their theistic memes. For the Qur'an, no period was allowed for harmonization of the text. Because only one edition was allowed, tampering with the text was impossible. Thus the Qur'an was "orthodox" from its beginnings; there is no evidence of processes to "corrupt" or mutate the text. However, because the Qur'an was compiled from the unedited recorded sayings of the Prophet and in a relatively short period by one author, inconsistencies would be expected and indeed are present. The most conspicuous inconsistency is the reference to other deities -in conflict with the monotheism so prominently expounded upon later in the text. These deities, al-Lat, al-'Uzza, and Manat, were goddesses present in contemporary religious culture. Because no alteration of the text has been allowed, this glaring inconsistency remains in the Qur'an, creating an unharmonized set of theistic memes.

Perhaps those who constructed the Islamic canon, who must have witnessed the difficulties encountered in formation of the Christian canon, ensured that no such tampering could occur with the Qur'an.

## 7 Discussion

Religious ideas, concepts, and dictated behaviors, referred to here as theistic memes, have a historical context in the written record beginning about five thousand years ago. Memes have been proposed to vary and undergo selection. The examples provided in this paper show that theistic memes, studied chronologically, have been transmitted from generation to generation (vertical transmission), have varied, and with time have been selected and incorporated into a number of different religious texts. The proposed underlying mechanism for the successful incorporation of a selected theistic meme and its successful propagation in a religious canon is the ability of that meme to enhance human fitness.

A study of the evolution of the *Epic of Gilgamesh* shows how fitness-enhancing memes were progressively incorporated into the canon. In the Sumerian tales, creative scribes infused mutant theistic memes into fanciful stories with little message. The author(s) of the *Epic* took a story that primarily focused on the hero-god himself and transformed the focus into a message about death and the afterlife. They created the first-known recorded theistic memes that directly affected reproductive success by promoting marriage and family ("Pay heed to a little one that holds your hand. / Let a spouse delight in your bosom") and reducing death anxiety (life should be lived for the moment, not in obsessive thoughts of death and the afterlife).

If theistic memes randomly varied, some would, by chance, be fitness minimizing; incorporated into a canon and taken as rules for behavior, these could lead to extinction of a population. The American Shakers are a most notable example of this phenomenon. This Christian sect adopted the fitness-minimizing theistic meme of celibacy. The sect was otherwise not unlike the Quakers of that era, who believed in religious equality of the sexes, a pending apocalypse, and condemnation of established religious hierarchies. Shakers migrated to America in the nineteenth century, and one of these immigrants, Ann Lee, organized the sect into a group of interrelated churches throughout New England. In response to attacks on the religion by the press, the sect adopted a written code of uniform behavior, the Millennial Laws (Shakers 1848), which formalized beliefs and practices for its members and helped define the religion for the general public. The distinctive behavior codified in these Laws was celibacy and separation of the sexes. In Chapter II, the "Principles of the Church of Christ, as revealed to practice in the United Society" list the seven moral duties of the Believer: ". Duty to God 2. Duty to man 3. Separation from the world 4. Practical peace 5. Simplicity of language 6. Right use of property 7. A virgin life." Paragraph 48 of the seventh moral duty reads,

*A virgin life, therefore, means a life of purity, a life undefiled with sinful indulgences, unmixed practices, unadulterated with carnal gratifications and impure desires and pursuits. But the sense to which we immediately confine it, and indeed which is the most obvious, is to express a life of continency, or a life of abstinence from carnal gratifications of the flesh, from sexual coition, and from all lascivious indulgences.*

In paragraph 51, the Shakers use arguments from the past to justify the virgin life; for example: "Josephus informs us that the Essenes, who maintained the virtue of continence, were many of them favored by divine revelation." Thus the Shaker restriction on marriage and sexual behavior may have had its origins in a historical meme about the Essenes, a sect that also became extinct.

The practicalities of this theistic meme meant that to gain new members the Shakers had to rely on adoptions and conversions. The number of Shaker societies grew in the early eighteenth century, but by the time of the Civil War the number of "Believers" began to decline. With fewer members, inadequate management of Shaker assets and land ensued and contributed to the economic decay of the church. Between 1880 and 1900 the Shaker population declined by half. Burns and colleagues (1996) report a newspaper account of 1905 that described the root of the Shaker demise:

*It was a case of love at first sight on both sides when they met four years ago, but neither dared to speak to each other . . . One day, as Audette was passing his lover's window, a thimble dropped to the ground. Looking up he saw Miss Thayer in the window. He immediately tore a piece of paper from a notebook, wrote a proposal of marriage, tucked it in the thimble, and tossed it to his sweetheart. Last November Miss T. forsook the Shaker religion and went to Winsted, Ct . . . where they were married in April.*

By the mid-twentieth century the membership had dwindled to forty. Soon thereafter the last male member died. The Shakers, unable to sustain their population, became extinct.

The search for the origins of *Homo sapiens* has led to two extremes of beliefs. Either the human species had a divine origin or, like any other species, it evolved from previous life forms.

Although there are those who believe strictly in one position or the other, many have adopted a hybrid position in an attempt to reconcile scientific facts with a belief in divine human creation. The most notable of the latter was Alfred Russell Wallace, co-discoverer with Charles Darwin of the theory of evolution by natural selection. Wallace believed that all species evolved according to the Darwin-Wallace theory, except the human species. But as the accumulating scientific evidence provided overwhelming validation of the theory, exclusion of humans from a history of evolution by natural selection became untenable. However, large communities still hold to religious beliefs denying human evolution.

In the scientific community, the present-day debate has polarized into those who categorically deny any divine existence or intervention in life's evolutionary origins and those who believe in human evolution but allow the coexistence and credibility of a formalized religious belief.

According to the former group, best represented by Richard Dawkins (1976), all life forms and their respective behaviors have been made possible by evolution through the natural selection of genes over the earth's four billion year history. The latter group, best represented by Stephen Jay Gould, acknowledges the fact of evolution but does not concede that religion is irrelevant. Gould (1987) writes, "Unless at least half my colleagues are inconsistent dunces, there can be -on the most raw and direct empirical grounds -no conflict between science and religion. I know hundreds of scientists who share a conviction about the fact of evolution, and teach it in the same way. Among these people I note an entire spectrum of religious attitudes."

Dennett (1996), in reflecting on this passage, concludes that some evolutionary biologists have in effect drawn a *cordon sanitaire* around the topic of religion and evolution. In this paper I have proposed that *religion is a part of human evolutionary history*.

Another polarizing issue for evolutionary biologists is the question of the extent to which genes influence human behavior. Dawkins (1976) and Wilson (1989) presume that genes play a major

role; Gould (1977) and Dobzhansky (1963) minimize genetic influence on human behavior. Dobzhansky (1963) states,

*The first, basic fundamental fact about human evolution is that mankind is simultaneously engaged in two kinds of evolutionary development -the biological and the cultural. Human evolution can be understood only as a product of interaction of these two developments . . . Culture is not inherited through genes, it is acquired by learning from other human beings. In a sense, human genes have surrendered their primacy in human evolution to an entirely new, non-biological or superorganic agent, culture. However, it should not be forgotten that this agent is entirely dependent on the human genotype.*

The classic argument used to define the debate between these two groups, as proposed by Gould (1977), is the self-sacrificing behavior of Eskimos. In some family groups of Eskimos, when food becomes scarce, older family members willingly sacrifice themselves to enhance the survivability of their children and grandchildren. The gene theorists believe this altruistic trait is under direct genetic control. Families with the gene for altruism, whose grandparents increased the survivability of younger generations by sacrificing themselves in times of hardship, have increased fitness. Those lacking the gene are less fit and will likely perish.

That such a complex social strategy could be under genetic control -implying that translation of DNA into a specific protein leads directly to a behavior specific to a particular population -seems implausible. Although genetic influence in certain human behaviors is incontrovertible (Steen 1996), genetic control of all social behaviors and strategies would seem far beyond the functional capability of the genetic code. But if genes do not directly control social behavioral strategies, what does?

Gould's answer, using the scenario of altruism among Eskimos, is that the behavior is an adaptive, nongenetic cultural trait. Some families celebrate sacrifice in song or story, venerating aged grandparents who sacrificed themselves and thus maintaining the capacity for this fitness-enhancing behavior through family generations. Families without such a tradition, passed through stories or legends, could become extinct.

This proposed role of religion in human evolution is in agreement with Benedict's (1989) assumption that the human species has a range of genetically determined behaviors, evolved through natural selection. Within this range of plastic social strategies and behaviors, specific strategies and behaviors are executed not under genetic control but under some sort of societal influence or religious mandate. Religious laws are subject to selection and can evolve to provide the species with greater fitness and reproductive success. The agent affecting fitness is the meme, as conceived by Dawkins: a unit of cultural transmission, propagated from generation to generation as the spoken or written word. Memes resemble genes in that they carry information, can randomly vary, and are subject to selection. Most importantly, they can affect human survivability and fitness. The culturally transmitted songs or stories of which Gould speaks in his explanation of self-sacrifice in Eskimos, and which Dobzhansky calls "new, nonbiological or superorganic agent[s]," are in fact memes.

The model proposed in this paper takes the meme concept further: in some populations, memes have been selected to produce a religious literature that enhances the population's fitness by providing strict codes of conduct and social strategies. The phenotypic expression of behavior is determined first by a broad spectrum of possibilities provided by a genotype; but the specific behavior executed is determined by a theistic meme incorporated in a religious canon.

This model stipulates that religion evolved as a consequence of humans' sophisticated self-awareness. The model recognizes the important contribution of Maser and Gallup (1990) in describing humans' self-awareness as a trait that enhances fitness but also creates the epiphenomenon of an awareness of finiteness and eventual death. At some point in human evolutionary history, self-awareness created a void in humans' interpretation and understanding of their world. At that juncture, the species needed answers to alleviate fears and anxieties arising from awareness of the unknown and death. The god concept filled this void. With the advent of writing some 6000 years ago, the specifics of god and the relationship between humans and god began to take on a permanent, indelible form. These written forms could have three different effects on the fitness of the gene pool: positive, negative, or neutral (no effect). What has been demonstrated here to support this model is that theistic literature has been subject to selection. Fitness-enhancing theistic literature has survived to become canon; fitness-minimizing theistic literature has not survived selection and has become extinct.

The model described here is supported by the unique and remarkable fitness provided by theistic memes in the Jewish canon for several millennia (Table 7). The Hebrew religious canon underwent profound evolutionary changes, providing increasing fitness at four distinct periods in its history. In the formative period, after the Exile (around 1300 BCE), the Covenant code and the Decalogue provided the Hebrews with a unique understanding of the relationship with God: all men were equal under a covenant made with a just God. The later Deuteronomic code affirmed this relationship, excluding foreign religious influences and incorporating fitness-enhancing memes about sexual behavior that promoted the nuclear family. Of profound importance in limiting Hebrew fitness, however, was the practice of child sacrifice. Addition to the canon of a strict prohibition of this practice (including the story of Abraham and Isaac), with sacrifice transferred to animals, markedly raised overall fitness. The Holiness code further affected reproductive fitness by severely limiting behaviors such as prostitution, incest, adultery, sodomy, and bestiality and encouraging fitness-enhancing behaviors that perpetuated the nuclear family.

	Middle Babylonian 1750 BCE	Canaanites 1400 BCE	Hebrew 700 BCE	Jewish 1100 CE	Christian (Church of Rome) 1100 CE	Islam 1100 CE	Christian (Cathars) 1100 CE	Christian (Shakers) 1800 CE	Christian (The Peoples Temple) 1977 CE
<b><i>Self-awareness</i></b>									
Creation	yes	yes	yes	yes	yes	yes	yes	yes	yes
God	yes	yes	yes	yes	yes	yes	yes	yes	yes
Universe	yes	yes	yes	yes	yes	yes	yes	yes	yes
Man	yes	yes	yes	yes	yes	yes	yes	yes	yes
Death	yes	yes	yes	yes	yes	yes	yes	yes	yes

Suicide	NS	NS	against	against	against	against	for	against	for
Afterlife	yes	yes	yes	yes	yes	yes	yes	yes	yes
Unknowns	yes	yes	yes	yes	yes	yes	yes	yes	yes

#### *Behavior / God*

Ritual									
Sacrifice	yes	yes	yes	yes	no	no	no	no	no
-Animal	for	for	for	for	against	against	against	against	against
-Adult	against	against	against	against	against	against	against	against	against
-Child	against	practiced	practiced	against	against	against	against	against	against
Worship	yes	yes	yes	yes	yes	yes	yes	yes	yes

#### *Behavior / Fellow man*

Principle of interaction	yes								
Culpable Negligence	yes								
Murder	yes								
War	yes	yes	yes	yes	yes	yes	no	no	yes

#### *Behavior/ Nuclear Family*

Marriage	for	for	for	for	for	for	against	against	for
Betrothal contracts	for	for	for	for	for	for	against	against	for
Foreign intermarriage	NS	NS	NS	against	NS	NS	NS	NS	NS
Divorce	permitted	NS	permitted	permitted	against	permitted	against	against	against
Parental respect	for	for	for	for	for	for	for	for	for

#### *Behavior / Sexual*

Adultery	NS	against							
Polygamy	for	for	for	against	against	for	against	against	against
Heterosexual sex	for	for	for	for	for	for	against	against	for
Rape	against								
Incest	against	for	against						
Homosexual sex	NS	for	against						
Bestiality	NS	for	against						
Castration	NS	NS	against						
Religious	NS	for	against						

Prostitution									
<b>Fitness Index</b>	Low /Moderate	Low	Moderate	High	High	High	Low	Low	Nil

**Table 7:** Theistic Memetic Types and Fitness Index in Various Religious Canon

Most of the Torah is related to outlining a credible world order: the nature of God, God's credibility, God's relationship to man, and the behaviors necessary to obtain holiness. The great majority of the 613 commandments in the Torah are devoted to these topics. However, in the long and tumultuous history of the Hebrews, the Law had to be modified (mutated) to fit the times. The remarkable stability of the Jewish canon that emerged from the multitude of ancient Near Eastern populations is due to the incorporation of fitness-maximizing theistic memes and the rejection of fitness-minimizing memes. However, the Hebrew population, in its highly stable state of fitness, was confined by laws against intermarriage. Other reproductive populations with less stable fitness were, by self-awareness, at a disadvantage.

The Christian canon enabled gentiles and disaffected Jews to rise to the same level of fitness as the Jewish population (Table 7). The Christian canon, accepted by many populations, was able to overwhelm less stable fitness strategies throughout the Roman Empire. For Arabic populations living with a polytheistic canon, Muhammad created an Arabic Jewish monotheism, incorporating in the Qur'an essentially the same fitness-maximizing theistic memes found in the Hebrew Torah. Through human history, some populations have assumed mutant theistic memes that were detrimental to fitness. The Shakers are a vivid example of a religious population that incorporated the lethal theistic memetic mutation of celibacy.

## References

- Benedict, R. (1989). *Patterns of Culture*. Boston: Houghton Mifflin Company.
- Blackmore, S. (1999). *The Meme Machine*. Oxford: Oxford University Press.
- Burns, K., Clay, L., Liebling, J., and Burns, S. A. (1996). *The Shakers: Hands to Work, Hearts to God; Their History and Visions from 1774 to the Present*. New York, London: Aperture, Robert Hale.
- Childs, S. B. (1965). The Birth of Moses. *Journal of Biblical Literature* **84**:109-122.
- Dawkins, R. (1976). *The Selfish Gene*. Oxford: Oxford University Press.
- Dennett, D. C. (1996). *Darwin's Dangerous Idea*. New York: Simon and Schuster.
- Dimont, M. I. (1962). *Jews, God, and History*. New York: Signet Press.
- Dobzhansky, T. (1963). Anthropology and the Natural Sciences -The Problem of Human Evolution. *Current Anthropology* **4**:138-148.

- Ehrman, D. B. (1993). *The Orthodox Corruption of Scripture*. Oxford: Oxford University Press.
- Encyclopaedia Judaica* (1971). s.v. Commandments, The, 613.
- Finegan, J. (1979). *Archaeological History of the Ancient Middle East*. Boulder, CO: Westview Press.
- Finkelstein, J. J. (1981). The Ox That Gored. *Transactions of the American Philosophical Society* **71**:5-47.
- Gould, S. J. (1977). *Ever Since Darwin: Reflections in Natural History*. London: Norton and Company.
- Gould, S. J. (1987). Darwinism Defined: The Difference between Fact and Theory. *Discover*, January, 64-70.
- Goulder, M. D. (1974). *Midrash and Lection in Matthew*. London: SPCK.
- Hamblin, J. D. (1987). Has the Garden of Eden Been Located at Last? *Smithsonian* **18**(2):127-135.
- Heidel, A. (1942). *The Babylonian Genesis*. Chicago: University of Chicago Press.
- Heidel, A. (1963). *The Babylonian Genesis: The Story of Creation*. Chicago: University of Chicago Press.
- Heidel, A. (1970). *The Gilgamesh Epic and Old Testament Parallels*. Chicago: University of Chicago Press.
- Heylighen, F. (1998). *What Makes a Meme Successful? Selection Criteria for Cultural Evolution*. Paper presented at 15th International Congress on Cybernetics.
- Johnson, P. (1988). *A History of the Jews*. New York: Harper and Row.
- Kramer, S. N. (1972a). Sumerian Mythology: A Review Article. *Journal of Near Eastern Studies* **51**:136-141.
- Kramer, S. N. (1972b). *Sumerian Mythology: A Study of Spiritual and Literary Achievement in the Third Millennium B.C.* Philadelphia: University of Pennsylvania Press.
- Maccoby, H. (1982). *The Sacred Executioner: Human Sacrifice and the Legacy of Guilt*. New York: Thames and Hudson.
- Maser, D. J., and Gallup, G. G. Jr. (1990). Theism as a By-product of Natural Selection. *Journal of Religion* **10**:515-532.
- Mellor, B. E. (1972). *The Making of the Old Testament*. Cambridge: University Printing House.

- Patrick, D. (1985). *Old Testament Law*. Atlanta: John Knox Press.
- Schmandt-Besserat, D. (1986). An Ancient Token System: The Precursor to Numerals and Writing. *Archaeology*39(Nov./Dec.):32-39.
- Segal A (in press). *Charting the Undiscovered Country: Western Views of the Afterlife* New York: Doubleday and The Anchor Bible Reference Library.
- Shakers (1848). *A Summary View of the Millennial Church or United Society of Believers, Commonly Called Shakers*, 2nd ed. Albany, NY: Van Benthuysen C Press.
- Spong, J. S. (1996). *Liberating the Gospels*. New York: Harper San Francisco.
- Steen, R. G. (1996). *DNA and Destiny: Nature and Nurture in Human Behavior*. New York: Plenum Press.
- Tigay, H. J. (1982). *The Evolution of the Gilgamesh Epic*. Philadelphia: University of Pennsylvania Press.
- Wilson, E. O. (1989). Sociobiology and the Theory of Natural Selection. *American Association for the Advancement of Science Symposia* 35:257-269.

## MEMES: INTRODUCTION

By

**Glenn Grant, Memeticist**

*"An idea is something you have;  
an ideology is something that has you."*

--Morris Berman

What if ideas were viruses?

Consider the T-phage virus. A T-phage cannot replicate itself; it reproduces by hijacking the DNA of a bacterium, forcing its host to make millions of copies of the phage. Similarly, an idea can parasitically infect your mind and alter your behavior, causing you to want to tell your friends about the idea, thus exposing them to the idea-virus. Any idea which does this is called a "meme" (pronounced 'meem').

Unlike a virus, which is encoded in DNA molecules, a meme is nothing more than a pattern of information, one that happens to have evolved a form which induces people to repeat that pattern. Typical memes include individual slogans, ideas, catch-phrases, melodies, icons, inventions, and fashions. It may sound a bit sinister, this idea that people are hosts for mind-altering strings of symbols, but in fact this is what human culture is all about.

As a species, we have co-evolved with our memes. Imagine a group of early Homo sapiens in the Late Pleistocene epoch. They've recently arrived with the latest high-tech hand axes and are trying to show their *Homo erectus* neighbours how to make them. Those who can't get their heads around the new meme will be at a disadvantage and will be out-evolved by their smarter cousins.

Meanwhile, the memes themselves are evolving, just as in the game of "Telephone" (where a message is whispered from person to person, being slightly mis-replicated each time). Selection favors the memes which are easiest to understand, to remember, and to communicate to others. Garbled versions of a useful meme would presumably be selected out.

So, in theory at least, the ability to understand and communicate complex memes is a survival trait, and natural selection should favor those who aren't too conservative to understand new memes. Or does it? In practice, some people are going to be all too ready to commit any new meme that comes along, even if it should turn out to be deadly nonsense, like:

*"Jump off a cliff and the gods will make you fly."*

Such memes do evolve, generated by crazy people, or through mis-replication. Notice, though, that this meme might have a lot of appeal. The idea of magical flight is so tantalizing -- maybe, if I truly believed, I just might leap off the cliff and...

This is a vital point: people try to infect each other with those memes which they find most appealing, regardless of the memes' objective value or truth. Further, the carrier of the cliff-

jumping meme might never actually take the plunge; they may spend the rest of their long lives infecting other people with the meme, inducing millions of gullible fools to leap to their deaths. Historically, this sort of thing is happening all the time.

Whether memes can be considered true "life forms" or not is a topic of some debate, but this is irrelevant: they behave in a way similar to life forms, allowing us to combine the analytical techniques of epidemiology, evolutionary science, immunology, linguistics, and semiotics, into an effective system known as "memetics." Rather than debate the inherent "truth" or lack of "truth" of an idea, memetics is largely concerned with how that idea gets itself replicated.

Memetics is vital to the understanding of cults, ideologies, and marketing campaigns of all kinds, and it can help to provide immunity from dangerous information-contagions. You should be aware, for instance, that you just been exposed to the Meta-meme, the meme about memes... The lexicon which follows is intended to provide a language for the analysis of memes, meme-complexes, and the social movements they spawn. The name of the person who first coined and defined each word appears in parentheses, although some definitions have been paraphrased and altered.

## Sources

Richard Dawkins, *The Selfish Gene*.

Keith Henson, "Memetics", Whole Earth Review #57: 50-55. Douglas Hofstadter, *Metamagical Themes*.

Howard Rheingold, "Untranslatable Words", Whole Earth Review #57: 3-8.

For a fictional treatment of these ideas, see my short story, "Memetic Drift," in *Interzone* #34 (March/April 1990).

# MEMETIC LEXICON

By

Glenn Grant

PO Box 36 Station H, Montreal, Quebec, H3C 2K5 (1990)

## **Auto-toxic**

Dangerous to itself. Highly auto-toxic memes are usually self-limiting because they promote the destruction of their hosts (such as the Jim Jones meme; any military indoctrination meme-complex; any "martyrdom" meme). (GMG) (See exo-toxic.)

## **bait**

The part of a meme-complex that promises to benefit the host (usually in return for replicating the complex). The bait usually justifies, but does not explicitly urge, the replication of a meme-complex. (Donald Going, quoted by Hofstadter.) Also called the reward co-meme. (In many religions, "Salvation" is the bait, or promised reward; "Spread the Word" is the hook. Other common bait co-memes are "Eternal Bliss", "Security", "Prosperity", "Freedom".) (See hook; threat; infection strategy.)

## **belief-space**

Since a person can only be infected with and transmit a finite number of memes, there is a limit to their belief space (Henson). Memes evolve in competition for niches in the belief-space of individuals and societies.

## **censorship**

Any attempt to hinder the spread of a meme by eliminating its vectors. Hence, censorship is analogous to attempts to halt diseases by spraying insecticides. Censorship can never fully kill off an offensive meme, and may actually help to promote the meme's most virulent strain, while killing off milder forms.

## **co-meme**

A meme which has symbiotically co-evolved with other memes, to form a mutually-assisting meme-complex. Also called a symmeme. (GMG)

## **cult**

A sociotype of an auto-toxic meme-complex, composed of membots and/or memeoids. (GMG) Characteristics of cults include: self-isolation of the infected group (or at least new recruits); brainwashing by repetitive exposure (inducing dependent mental states); genetic functions discouraged (through celibacy, sterilization, devalued family) in favor of replication (proselytizing); and leader-worship ("personality cult"). (Henson.)

## **dormant**

Currently without human hosts. The ancient Egyptian hieroglyph system and the Gnostic Gospels are examples of "dead" schemes which lay dormant for millennia in hidden or untranslatable texts, waiting to re-activate themselves by infecting modern archeologists. Some

obsolete memes never become entirely dormant, such as Phlogiston theory, which simply mutated from a "belief" into a "quaint historical footnote."

### **earworm**

"A tune or melody which infects a population rapidly." (Rheingold); a hit song. (Such as: "Don't Worry, Be Happy".) (f. German, *ohrwurm*=earworm.)

### **exo-toxic**

Dangerous to others. Highly exo-toxic memes promote the destruction of persons other than their hosts, particularly those who are carriers of rival memes. (Such as: Nazism, the Inquisition, Pol Pot.) (See meme-allergy.) (GMG)

### **hook**

The part of a meme-complex that urges replication. The hook is often most effective when it is not an explicit statement, but a logical consequence of the memeUs content. (Hofstadter) (See bait, threat.)

### **host**

A person who has been successfully infected by a meme. See infection, membot, memeoid.

### **ideosphere**

The realm of memetic evolution, as the biosphere is the realm of biological evolution. The entire memetic ecology. (Hofstadter.) The health of an ideosphere can be measured by its memetic diversity.

### **immuno-depressant**

Anything that tends to reduce a personUs memetic immunity. Common immuno-depressants are: travel, disorientation, physical and emotional exhaustion, insecurity, emotional shock, loss of home or loved ones, future shock, culture shock, isolation stress, unfamiliar social situations, certain drugs, loneliness, alienation, paranoia, repeated exposure, respect for Authority, escapism, and hypnosis (suspension of critical judgment). Recruiters for cults often target airports and bus terminals because travelers are likely to be subject to a number of these immuno-depressants. (GMG) (See cult.)

### **immuno-meme**

See vaccine. (GMG)

### **infection**

1. Successful encoding of a meme in the memory of a human being. A memetic infection can be either active or inactive. It is inactive if the host does not feel inclined to transmit the meme to other people. An active infection causes the host to want to infect others. Fanatically active hosts are often membots or memeoids. A person who is exposed to a meme but who does not remember it (consciously or otherwise) is not infected. (A host can indeed be unconsciously infected, and even transmit a meme without conscious awareness of the fact. Many societal norms are transmitted this way.) (GMG)

2. Some memeticists have used 'infection' as a synonym for 'belief' (i.e. only believers are infected, non-believers are not). However, this usage ignores the fact that people often transmit memes they do not "believe in." Songs, jokes, and fantasies are memes which do not rely on "belief" as an infection strategy.

### **infection strategy**

Any memetic strategy which encourages infection of a host. Jokes encourage infection by being humorous, tunes by evoking various emotions, slogans and catch-phrases by being terse and continuously repeated. Common infection strategies are "Villain vs. victim", "Fear of Death", and "Sense of Community". In a meme-complex, the bait co-meme is often central to the infection strategy. (See replication strategy; mimicry.) (GMG)

### **membot**

A person whose entire life has become subordinated to the propagation of a meme, robotically and at any opportunity. (Such as many Jehovah's Witnesses, Krishnas, and Scientologists.) Due to internal competition, the most vocal and extreme membots tend to rise to top of their sociotypeUs hierarchy. A self-destructive membot is a memeoid. (GMG)

### **meme**

(pron. 'meem') A contagious information pattern that replicates by parasitically infecting human minds and altering their behavior, causing them to propagate the pattern. (Term coined by Dawkins, by analogy with "gene".) Individual slogans, catch-phrases, melodies, icons, inventions, and fashions are typical memes. An idea or information pattern is not a meme until it causes someone to replicate it, to repeat it to someone else. All transmitted knowledge is memetic. (Wheelis, quoted in Hofstadter.) (See meme-complex).

### **meme-allergy**

A form of intolerance; a condition which causes a person to react in an unusually extreme manner when exposed to a specific semiotic stimulus, or 'meme-allergen.' Exo-toxic meme-complexes typically confer dangerous meme-allergies on their hosts. Often, the actual meme-allergens need not be present, but merely perceived to be present, to trigger a reaction. Common meme-allergies include homophobia, paranoid anti-Communism, and porno phobia. Common forms of meme-allergic reaction are censorship, vandalism, belligerent verbal abuse, and physical violence. (GMG)

### **meme-complex**

A set of mutually-assisting memes which have co-evolved a symbiotic relationship. Religious and political dogmas, social movements, artistic styles, traditions and customs, chain letters, paradigms, languages, etc. are meme-complexes. Also called an m-plex, or scheme (Hofstadter). Types of co-memes commonly found in a scheme are called the: bait; hook; threat; and vaccine. A successful scheme commonly has certain attributes: wide scope (a paradigm that explains much); opportunity for the carriers to participate and contribute; conviction of its self-evident truth (carries Authority); offers order and a sense of place, helping to stave off the dread of meaninglessness. (Wheelis, quoted by Hofstadter.)

**memeoid, or memoid**

A person "whose behavior is so strongly influenced by a [meme] that their own survival becomes inconsequential in their own minds." (Henson) (Such as: Kamikazes, Shiite terrorists, Jim Jones followers, any military personnel). hosts and membots are not necessarily memeoids. (See auto-toxic; exo-toxic.)

**meme pool**

The full diversity of memes accessible to a culture or individual. Learning languages and traveling are methods of expanding one's meme pool.

**memetic**

Related to memes.

**memetic drift**

Accumulated mis-replications; (the rate of) memetic mutation or evolution. Written texts tend to slow the memetic drift of dogmas (Henson).

**memetic engineer**

One who consciously devises memes, through meme-splicing and memetic synthesis, with the intent of altering the behavior of others. Writers of manifestos and of commercials are typical memetic engineers. (GMG)

**memeticist**

1. One who studies memetics. 2. A memetic engineer. (GMG)

**memetics**

The study of memes and their social effects.

**memotype**

1. The actual information-content of a meme, as distinct from its sociotype.  
2. A class of similar memes. (GMG)

**meta-meme**

Any meme about memes (such as: "tolerance", "metaphor").

**Meta-meme, the**

The concept of memes, considered as a meme itself.

**Millennial meme, the**

Any of several currently-epidemic memes which predict catastrophic events for the year 2000, including the battle of Armageddon, the Rapture, the thousand-year reign of Jesus, etc. The "Imminent New Age" meme is simply a pan-denominational version of this. (Also called the 'Endmeme'.)

**mimicry**

An infection strategy in which a meme attempts to imitate the semiotics of another successful meme. Such as: pseudo-science (Creationism, UFOlogy); pseudo-rebelliousness (Heavy Metal); subversion by forgery (Situationist détournement). (GMG)

**replication strategy**

Any memetic strategy used by a meme to encourage its host to repeat the meme to other people. The hook co-meme of a meme-complex. (GMG)

**retromeme**

A meme which attempts to splice itself into an existing meme-complex (example: Marxist-Leninists trying to co-opt other sociotypes). (GMG)

**scheme**

A meme-complex. (Hofstadter.)

**sociotype**

1. The social expression of a memotype, as the body of an organism is the physical expression (phenotype) of the gene (genotype). Hence, the Protestant Church is one sociotype of the Bible's memotype. 2. A class of similar social organisations. (GMG)

**threat**

The part of a meme-complex that encourages adherence and discourages mis-replication. ("Damnation to Hell" is the threat co-meme in many religious schemes.) (See: bait, hook, vaccine.) (Hofstadter)

**Tolerance**

A meta-meme which confers resistance to a wide variety of memes (and their sociotypes), without conferring meme-allergies. In its purest form, Tolerance allows its host to be repeatedly exposed to rival memes, even intolerant rivals, without active infection or meme-allergic reaction. Tolerance is a central co-meme in a wide variety of schemes, particularly "liberalism", and "democracy". Without it, a scheme will often become exo-toxic and confer meme-allergies on its hosts. Since schemes compete for finite belief-space, tolerance is not necessarily a virtue, but it has co-evolved in the ideosphere in much the same way as co-operation has evolved in biological ecosystems. (Henson.)

**vaccine**

(pron. vak-seem) Any meta-meme which confers resistance or immunity to one or more memes, allowing that person to be exposed without acquiring an active infection. Also called an 'immuno-meme.' Common immune-conferring memes are "Faith", "Loyalty", "Skepticism", and "tolerance". (See: meme-allergy.) (GMG.)

Every scheme includes a vaccine to protect against rival memes. For instance:

Conservatism: automatically resist all new memes.

Orthodoxy: automatically reject all new memes.

Science: test new memes for theoretical consistency and (where applicable) empirical repeatability; continually re-assess old memes; accept schemes only conditionally, pending future re:-assessment.

Radicalism: embrace one new scheme, reject all others.

Nihilism: reject all schemes, new and old.

New Age: accept all esthetically-appealing memes, new and old, regardless of empirical (or even internal) consistency; reject others. (Note that this one doesn't provide much protection.)

Japanese: adapt (parts of) new schemes to the old ones.

### **vector**

A medium, method, or vehicle for the transmission of memes. Almost any communication medium can be a memetic vector. (GMG)

### **Villain vs. Victim**

An infection strategy common to many meme-complexes, placing the potential host in the role of Victim and playing on their insecurity, as in: "the bourgeoisie is oppressing the proletariat" (Hofstadter). Often dangerously toxic to host and society in general. Also known as the "Us-and-Them" strategy.

# **MEMES AND THE CREATION OF NEW PATTERNS OF MOVEMENT IN DANCE**

**By**

**Prof. Dr. Christine Greiner**

Catholic University of São Paulo, Center of Dance Studies of the Department of Semiotics  
(PUC-SP)

This paper aims to study the influence of cultural dialogue in the creation of new patterns of movement in dance and to develop a hypothesis to understand how it works, presenting a practical example: the evolution of butoh dance in the West.

Butoh is a radical artistic manifestation born in Japan in the late 50s; after the 80s it spread through the Western world, suffering successive degradations. Besides Pina Bausch's Tanztheatre, butoh thought can be indicated as one of the powerful memes developed in the second half of our century.

However, we have to be careful to avoid hasty conclusions, questioning whether we can really understand butoh as a meme and, if so, we have to clarify what kind of instruction it really offers to the replication process.

We will take as fundamental theoretical instruments: the Consciousness Thesis of Daniel Dennett (1991;1995;1997), the Meme Thesis (1976) of Richard Dawkins and Charles Darwin's (1859) idea of evolution by natural selection, interpreted by Dawkins and Dennett.

Such a methodological choice transfers theoretical and conceptual matrices of the cultural dialogue problem from the outer scope of sciences that have considered aesthetical changes of choreographies as simple products of socio-cultural stimulus of the environment to the understanding of cultural dialogue as an action of consciousness (and not its product). Therefore, perceiving an environment is not to make the brain receive stimuli in a passive way.

According to Dawkins (1982), a meme should be regarded as a unit of information residing in a brain. It has a definite structure, realized in whatever physical medium the brain stores information. This means: as a pattern of synaptic connections or in a distributed form. In the latter, it would not be detectable on a microscope, but, Dawkins insists, it still reside in the brain.

These mental representations and the movements implemented in the body are a necessary integration (brain-body-environment) that affords not only the artistic making, but the survival.

Through cultural dialogue, mental representations (memes) are changed because the brain receives new information through different modes of perception. To understand how this transformation of mental representations influences the creation of new patterns of movement in dance, it is necessary to think about the implementation of this information in the dancing body.

According to Dawkins, a cake is the consequence of obeying a series of instructions when to mix the various ingredients but it is not those instructions translated into another coding medium. A

body too, is the result of obeying instructions, but we cannot reconstruct an individual's genome by studying his body. And the same happens to the dancing body. When you learn a new dance technique and receive the instructions, you will not translate them into another medium, thereby recovering the "original". In the new environment, the instructions will be re-organized according to the brain-body continuum, establishing new relationships, and different mental representations that once implemented in the body, will give birth to new patterns of movement.

Time develops an important role in this process. In the dancing body, to develop a new pattern of movement, it is necessary to conquer a minimum of stability. This stable thought, that Dennett called conscious thought, will permit a dance movement (implemented thought) to exist as a habit, i.e. a movement already incorporated to create a specific dancing body.

According to Dennett (1991:210): "human consciousness is itself best understood as the operation of a 'von Neumannsque' virtual machine implemented in the parallel architecture of a brain that was not designed for any such activities". Dennett (1996:155) explains that mental contents become conscious not by entering some special chamber in the brain but by winning the competition against others.

In dance, memes, as mental representations and the process of replication of these representations of ideas (or other imagined constructs), are phenotypically mediated by the dancing body. In this transit between mental patterns and body implementation, natural selection occurs at different levels. There is a big difference between replicators and vehicles.

## The example of butoh evolution

To understand butoh evolution we ask a question to identify this relationship: which is the information of this universe of knowledge that is able to replicate?

As we said before, a vehicle is not a replicator. Therefore, the dead body or even the idea of the dead body (the fundamental of butoh dance) cannot be replicated. Only the instructions that make it happen.

At this point, it is important to review some basic principles to think about their evolution in the West.

Semantics professor Yoshie Yoshida discusses some of the main translations of the butoh ideogram, explaining that in general butoh is understood as the harmony between elevation (*bu*) and being earth bound. However, he interprets *bu* (from *buyo*) as *mai* (another name for dance in Japanese). In this sense, *bu* could mean *ma-i* or the action of *ma* (known as the conception of the time-space interval that is a fundamental aspect of Japanese culture). The choreographer Akaji Maro, who worked for many years with Tatsumi Hijikata, the creator of butoh dance, said that the body is supported by something invisible. Butoh is about capturing the spirits that live in *ma*. It is beyond outside time.

To complete this reasoning, we can already mention another butoh choreographer, Ko Murobushi, who explains the dead body as the possibility of giving existence to anything. Like

*ma*, which can be understood not as an empty place in the sense of nothingness, but in the sense of emptiness in which anything can happen.

Therefore, we think about *ma* as a kind of instruction of the butoh meme. Or, more specifically, of the dead body meme, as an imaginary construction.

If this instruction is able to replicate in other bodies, it will not reproduce or translate butoh in another medium (a western body, for example). But it will exist as another possibility of experiencing life.

After many years of studying butoh, we have concluded that a good definition for butoh is to think about this dance as a map of one of the possibilities of the state of being alive.

Butoh is just one example. The Memetics approach can contribute to the study of dance by reorganizing Dance History and the understanding of the evolution of the different dances.

It permits to think about what kind of instructions can or cannot replicate, according to the relationship present at the continuum body-brain-enviroment.

Thinking about the creation of new patterns of movement in dance, based on cultural dialogue, Memetics explains that it is not enough to study the enviroment, neither the mental representations apart from their extended phenotypes. It is in the transit between mental representation and its implementation in the dancing body, that natural selection is responsible for the aesthetical result. Therefore, cultural dialogue in dance is often responsible for the creation of new patterns of movement in the dancing body, but not by a sociological effect of the external world on the inner world of the dancer, but by natural selection of the competition among mental contents, meme-effects in brains. In this sense, it's the same process that is present in all nature: an algorithmic process (Dennett, 1995, 1996; Dawkins, 1986, 1995, 1996) creating new designs.

## Bibliography

- Dawkins, Richard (1976) The Selfish Gene.NY: Oxford Univesity Press  
\_\_\_\_\_(1989) The Blind Wachtmaker.NY: W.W.Norton  
\_\_\_\_\_(1995) River Out of Eden.Orion Publishing Books  
\_\_\_\_\_(1996) Climbing Mount Improbable.NY:W.W.Norton & Company
- Dennett, Daniel (1991) Consciousness Explained.Boston:Little, Brown and Company  
\_\_\_\_\_(1995 ) Darwin Dangerous Idea.NY:Touchstone Books  
\_\_\_\_\_(1996) Kinds of Minds. NY:Basic Books
- Gabora, Liane (1997) The Origin and Evolution of Culture and Creativity  
Journal of Memetics, Evolutionary Models of InformationTransmission,1
- Greiner, Christine (1997) Butoh Dance in the West, a thought in evolution. Doctoral Thesis defended at Catholic University of São Paulo (BR)

**MEMES AND GRINNING IDIOT PRESS  
OR  
WHY I HAVE BEEN STUDYING VAMPIRES SINCE 1972**

By  
Grinning Idiot Press 1993  
716 Unit A 21 St N  
St Petersburg FL 33713  
(813)321-2120

**Why GRINNING IDIOT PRESS?**

At one time I was involved in a somewhat extreme version of Christianity; I joined in my innocent and ignorant teenage years when I was vulnerable. The group I fell under the sway of is in the outer territory of cultdom; I will not name names. Somewhere along the line I forgot to turn off my bright and inquiring mind, which led to no end of trouble for me. At first I simply wanted to learn all about the faith, but after I learned a thing I thought about it and what it meant and how it related to other things in the faith I was taught.

After discovering that when I criticized certain ideas and doctrines and tried to analyze them logically in debate with respected authorities in the faith, I elicited irrational behavior from them, which surprised me immensely. At the time, I associated with this group, and considered myself one of the good guys, like them. My motive was to dispel apparent contradiction; it never dawned on me (how naive) that there were things thou shalt not question. When I criticized or questioned certain things or pointed out apparent contradictions, the results were amazing and VERY traumatic. I thereafter did my research in private and slowly came to certain conclusions and discoveries which were NOT approved right thinking. I came to the conclusion that they were not very sane, and that a few of them were downright irrational. Further research led me to understand the compartmenting behaviour, where a cultist leads one part of their life by one definition of logic, which is NOT applied to all parts. If I fixed my brakes by this sort of sanity, the ride would be exciting, to say the least. In short, I burned out.

This led to a year long bout of depression and internal conflict as my rational mind came to grips with habitual mental behavior that was difficult to stop. I came to understand that I had to some extent been programmed, and that I had assisted it, if only by default. It took some time to come to trust and like myself after that, and to understand what happened and how. I had been ABSOLUTELY CERTAIN I was right, only to later discover I had been in error. Self trust, and faith in general, became VERY iffy things.

It was easy to ascribe evil intentions to these people, but on examination of several with carefully worded questions and long term observation of their behavior, I satisfied myself that they in fact were as programmed as was I, or more so. This left me wondering how such things come to be, until I read certain works by Robert A. Wilson, Dawkins, and others.

After I had learned and discovered many things, got my facts and head (a little more) straight, and developed a measure of understanding, I engaged in a series of discussion groups and

debates with several members of this community of like minded individuals, some of whom were respected scholars in their system of thought. At this time, my theories were still quite primitive, and I was not very certain about what would happen, which was part of why I did it. The results were remarkable. They displayed an amazing selective blindness to evidences, a fascinating selective memory concerning evidences produced, and utterly defective recall of the evenings in question several months later when I spoke to the organizer at a party. I remain convinced they were being spontaneous and NOT intending to deceive me.

An associate insists they are liars, and dishonest men pursuing an agenda of power. Concerning 90% of them, I disagree. People do not react that quickly when their reactions are not preplanned (as they did not know what to expect), and are deliberately chosen to play a role supporting some ulterior motive. Their emotions do not become so intense when faked as to raise blood pressure (with attendant flushing), produce sweat, fidgeting, and other body english. These people really meant it, however irrational their favored persuasions. They are in effect captives of coherent information systems which have subverted their ability to deliberate, to THINK; they are 'possessed' by what are called MEMES, information viruses. A popular counter culture term for such is robots. Those with limited or total immunity who use them to control robots, and are to some extent aware they are doing it, are termed vampires, a term invented by an author. PSYCHIC vampires, as they drain off free will and enslave minds. I refine such terms in a moment. Cult lords are vampires.

When I spoke to these robots, they seemed to ignore anything I said and spoke and acted as if I was somehow mentally impaired (improperly programmed?), and had to be led to the truth gently by the hand, like some sort of a grinning idiot (a term I heard descriptive of simpletons in my child hood from my friend next door Carl Cleback). Fascinating. Were they truly liars, or fools who really and truly mean it? (Yes, I later found out.)

Investigating further, I came across more of the writings of Robert Anton Wilson which pointed out to me that there is in men a thing Leary calls the reality grid, a system of information whereby we interpret all raw information coming in from the senses as important or trivial; it also helps us determine WHAT the information means. I was on to something. Another work, *The Selfish Gene* by Dawkins introduced an idea and a word, the MEME, a coherent system of information which tends to 'hang together' and propagate through society, existing pretty much for its own sake. Some are harmless, or even beneficial, such as the 'stop, drop, and roll' poem and song invented by the Canadian government. (IN BRIEF, it reminds people whose clothing is on fire to NOT run in panic, but instead, stop, lie down, and roll the fire out. Cool?)

I came to believe that many people are robots controlled by memes. A few are partially or totally immune to their chosen virus, but continue to associate with them and the participating robots for the bennies of prestige, power, money, and sometimes even nookie, these I term vampires. Highly successful ones I call Vampire Lords. The ones who are infected without knowledge of it are the robots. Some admitted the programming but did not know what they were doing. They are still robots. They are not particularly awake. Vampires have some partial awareness of what is up at least; robots do not. They are oblivious, and are controlled by whatever got into their minds. If that data included some defensive software, they will resist further programming by unauthorized sources. The control loop for a robot is externalized with random sources and

programming, or a specific source and program for the 'possessed' robot, the lackey of a vampire, a predatory neuromancer. Some of them are very good at it. Jim Jones was one such. Unfortunately, he was not immune to his virus, and succumbed to the self destruct program he created. He would not have been executed; we would have certainly found him insane. Half assed neuromancers are dangerous. To themselves.

People who control the programming of their own minds are called occultists or heretics or neuromancers or free spirits. People whose programs are sourced and defined by outside agents are ROBOTS. People WITHOUT programs in their subconscious minds are totally insane, or dead. People who define other people's programming may be consulting technicians, but if they do it to define that other person to their tastes, particularly to their benefit, they are practicing vampires.

Preprogrammed behavior which preserves life is VERY useful when you are filled with panic and the deliberative mind is running in circles screaming. The meme which allows you to remain standing erect while lost in thought is a complex and carefully structured meme. You invested years of work in developing it. It is a benign, unaggressive, and useful program dwelling now in the subconscious mind. Such skills are very useful. Everything they teach you in boot camp is ideally engraved on your brain stem for when seconds are far too long to take to decide and the wind from the grenades is blowing your sweat soaked hair dry and you're so scared the urine is running down your pant leg -- THAT is why you drill and practice until you DREAM how to field strip a rifle blindfolded. MAYBE your D.I. knew what he was doing?

Others have NOTHING to do with the well being of persons they infect, and are simply self replicating for their own sake. The simplest was invented by Dawkins, and is titled 'Say me'. SAY ME! There; I've done it; you are now infected with an information virus, a harmless one. Say me. Granted, it's so simple there is little to no reward for playing the sayme mind game, but there it is, and even so simple a meme has surprising persistence and propagation. Others are indeed seductive, even enjoyable to the mind they are subverting, but are irrational say me and harmful, either to the host, those around the host, or both.

Obviously there are many sorts of religion (and politics) which are irrational systems of self propagating information systems 'riding' host humans. Nazi fascism is such a meme, a large and complex one; so is irrational compulsive screaming emotional manipulative socialism, the sort that is sacred and therefore may not be criticized; it is politically incorrect to question it, disagree with it, or question it critically. (There is a conservatism like this too, a feminism, a macho masculine male superiority thing, a white thing, a black thing, an Islamic thing, an Atheistic thing, and so-on; each is equally irrational, and is also a meme; I am sure you can name religious groups like this. They are sometimes thought of as cults, but ALL of them are memes.) These memes invade using pseudo-rational trickery, outright brain washing, punishment for criticism, and anything else that bypasses rational examination, such as mental ignorance or sloth.

Many are subverted simply because they do not know better, such as the uninformed, the naive, cowards who prefer not to perceive than to resist emotional pressure, those who cannot stand the emotional stress of rejection, and children. The politely courageous one who knows that there are no propositions which may not be doubted and examined, who is willing to risk rejection, who

trusts in the process of thought, who knows that ideas have consequences and that words have meanings, such a person is highly resistant to memes -- and very rare. They are declared nonexistent, called every name in the book, and lied about by meme infested persons. The meme victim even believes their own dishonest statements, because they are being enslaved to illusions, and the more so the easier it is to believe one's own lies. Technically, they are insane; they cannot differentiate between reality and unreality, good and evil. As their compartmenting of the mind progresses, actual multiple personality disorder or schizophrenia is quite possible, although few push it this far. Instinct is a powerful counterbalance, and keeps many souls alive mostly on autopilot, while they keep repeating approved programs and mantras. Lest I be received as an arch-Atheist, I also mean programmed robots such as slavish fans of many popular figures, such as Rush Limbaugh, Ayn Rand, Anita Bryant, Jim Jones, Ronald Reagan, and Jane Fonda.

Rational analysis, PAYING ATTENTION, and an unwillingness to accept on a basis of emotional pressure are strong serums say me against these viruses. Say me. Of course, some say me viruses are harmless and even fun SAY ME to play with. Maybe. JUST REMEMBER: If you know you're doing it, you're not schizophrenic.

Some memes are without a doubt destructive, such as those associated with Jim Jones, Herr Hitler, Bulimia, Anorexia, 'BoB', and Alcoholism rooted in psychological disturbances. Almost everyone is to a limited extent controlled by memes not of their deliberate choosing, just as we all carry viruses we are partially resistant to (and some we are not but are benign; they simply get in, reproduce, we spread them, we evict the current crop, they take root elsewhere, and life goes on). There are many types of memes. They exist for their own sake. Usually.

As in organic life vs. viruses, destructive memes which are too toxic destroy their hosts so well they edit themselves out of the informational 'gene pool'. However, if they could be engineered, contained, and released in local areas, such information structures would make excellent weapon systems. I believe this is being inefficiently explored by many parties who only partially understand what they are about. I believe this has been more or less the case for some time. The cult outbreak of the 60's and 70's was simply a particularly toxic and obvious example of memes, in this case memes spread by persons who to some extent knew what they were doing.

Several of these leaders are now in jail or are deceased, yet their memes live on, with the organization attendant being shaken out and purged of corruption; the vampire lord has died, so the robots cleanse the order of the traces of corruption left behind, inventing new memes to explain these puzzling evidences (one of the sacred ideas is that the leader was not in it for the money, egoboot, power, or nookie).

Someday these 'cults' will be old established religions, with long traditions, carefully worked out theologies, and fictitious but very seriously affirmed and believed histories. The meme at such future date will only superficially resemble the one the founder was teaching. If Rabbi Joshua Bar Nazareth were to pop up today in a time warp and go to a church (ANY CHURCH), he would not be approved of; nor would he approve of them. Krishna would be appalled with people talking about him, Siddhartha (the Buddha) would either laugh or cry, at what his people

have made of what he left behind, Bless Me if I know which.

Now, before I get lynched, I do not mean to suggest any of these people were corrupt. Into suffering maybe. No one dead more than 500 years is of anything but the highest order of saintliness or the most vile of evilness, which these were not, unless they are simply no longer recalled at all. These most excellent teachers were simply edited and spin doctored and interpreted to beat hell after they departed. But I digress.

Memes can be useful tools to govern human behavior, yours or mine. They can be harmless, beneficial, or toxic, slightly or greatly. Best of all, they could be disseminated to the target audience with leaflet drops and radio broadcasts on shortwave or AM radio from outside the target audience. And Freedom of Speech protects it. If you think this is nuts and that you cannot talk people to death, please go refresh your memory of Jim Jones' body count. Then visit Auschwitz. When some people shoot their mouths off, they get a body count.

The more I learned, the more I liberated myself, and the more annoying to robots and dangerous to vampires I became. Both tend to shun me like the plague once they begin to know me. I guess they like their viruses more than mine. Perhaps censorship in time of war is simply a sort of public hygiene?

Since then I have been exploring the dynamics of this, keeping track of the performance record of vampires who make public declarations and promises. I also explore how to deprogram the robotic thinking in myself and others, and enable them and myself to replace it with chosen planning which is more enabling, more centered on personal well being and liberation, and more resistant to infection by information viruses by society. This generally pisses off robots and their vampire lords, and authoritarians in general. Free spirits leaning to reading and writing and words (left brain dominant) love this sort of trip. Wanna play?

Anything which radically challenges the popular reality interpretation reference grid of ideas is dangerous, because it could destroy it, and without a grid we would be lost in a sea of meaningless information, the plight of a new-born. We all have such grids. However, not all of us are emotionally involved in them; we are willing say me to revise and critically examine them, remove defects as we find them, improve them, even discard useless or harmful parts of them. Those who do not want to are the vampires. Some vampires understand all of this, if in different terms; they use this lore for their own benefit. These are vampire lords (and ladies, if I may abuse the term?); those who do not even understand this are robots, deeply, even hopelessly, under the sway of a meme.

It is possible to reprogram the grid with calm and deliberate rational thinking and meditation and self monitoring, although old habits die hard. It is generally best to do this carefully and only concerning specific subjects or subject groups, so the grid is restructured one section at a time; a total melt down is both very difficult to achieve and totally paralyzing, probably causing a psychotic episode if achieved. The idea is self liberation, not immolation. You have to have the subconscious programs you use to live; us neuromancers redesign them to suit our tastes and needs, redesign them to be more appropriate and fun and efficient for our particular lives. This of

course makes us heretics to all the robots and vampires. We do well to keep a low profile.

Final shot: a quote from Robert A. Wilson: ``If you are not programming your own trip, who is?''

## **EPIGENETIC SOLUTIONS TO THE ADOPTION PROBLEM IN EVOLUTIONARY PSYCHOLOGY OR THE ADOPTION MEME**

**By**  
**Dave Gross**  
(1993)

For those who try to examine how the pressures of natural selection have affected human behavior (an area of thinking known as "evolutionary psychology," "Darwinian psychology," "selection thinking," and probably several other terms), there have been some significant gains in understanding, but also some persistent problems.

Evolutionary psychologists try to examine human behavior from the standpoint of the theory of evolution -- to explain a human behavior by examining how that behavior contributes to the reproductive success of the individual exhibiting that behavior, and therefore to the proliferation of the genetic material which, presumably, at some level is responsible for causing that behavior. For these theorists, the problem of adoption is a thorny one -- at least when the adoption occurs outside of genetic relatives. Adoption requires a person to devote resources and energy to the upbringing of a child that carries unrelated genes.

While it has been shown that non-biological (i.e. step- or adoptive-) children are statistically treated in an inferior way by their parents than biological children are, the question remains in the case of adoption -- why are they treated well at all? Why are adoptive parents willing to put out so much time, energy and resources to help the reproductive success of a stranger?

A standard theory is that adoption, pursued mainly by otherwise childless couples, is a pathological hijacking of a parental instinct that would normally go towards encouraging the parents to raise biologically related offspring. If the parents are infertile, the psychological desires to raise a child remain and are transferred to an adopted child.

I believe that meme theory can take the pathology out of adoption, and eliminate the above awkward theory for one which uses more straightforward natural selection thinking.

The theory of evolution is based on the following facts:

- Individuals have traits which are specified genetically
- Some of these genes, and therefore traits, are passed by individuals to their offspring.
- There are variations in the genes in a population.
- Some of these genes will cause traits which make the individuals carrying them better able to propagate the genes.

Genes which have such traits will expand in the gene pool to the expense of less adapted genes. In evolutionary psychology, the restriction of natural selection thinking to the genetic level is a drawback. Memes, which can be thought of as "ideas," or "idea systems," or "systems of processing ideas," or indeed all three, also follow similar rules.

Individuals have traits (in this case, behaviors, but fashion is also a meme, so memes have the potential to shape the body as well -- the "big boobs are good" meme has shaped more than one woman's body through surgery) which are specified memetically (which isn't to say that they are permanent -- a big difference between memes and genes is that memes can be, and usually are, subject to mutation throughout the organism's lifetime. Some memes are more durable than others -- a lifetime Catholic is not unheard of, but a lifetime New Kids on the Block fan is unlikely).

Some of these memes, and therefore traits, are passed on to others. (Note: Not necessarily offspring; and in fact biological relation is only indirectly connected to meme propagation. This point is especially important to the adoption issue, and I'll go into more detail later.)

There are variations in the memes in a population (are there ever!)

Some of these memes will cause traits which make the individuals carrying the memes to be better able to propagate the meme. (A good example is a religion which encourages missionary work [read: causes the trait that we know as "missionary work"], compared to another religion which does not).

Memes which cause these traits will expand in the "meme pool" to the expense of less adapted memes.

Memes are passed epigenetically, and can be passed to people other than biological offspring.

Memes can be passed from person to person in conversation or below the level of communication "by example." Memes can even be propagated by the dead; the authors of the gospels continue to have their very successful memes read.

One way memes spread is from primary caretakers to children. The science of psychology tells us that a great deal of personality and thought is fixed in humans within the first five years of life. The primary caretakers of a child have a great opportunity to pass durable memes into their children. These memes can range from a religion or a language which is explicitly taught, to the location of the silverware drawer, which is absorbed less directly.

If the trait which encourages adoption is a genetic trait, we have the following situation (greatly simplified, of course, for illustration):

MOTHER      FATHER  
AA            BB        <- Genes

(BIOLOGICAL CHILD) (BIOLOGICAL CHILD) (ADOPTED CHILD)  
AB            AB           MN        <- Genes

If the gene that encourages adoption is, for instance, a "B" gene (again, forgive the simplification) -- that gene has been passed to two biological children in the above example. Clearly, from that gene's point of view, it would be more likely to succeed if the resources

allocated to the adopted child not carrying the gene were instead allocated to the two children carrying the gene, or to the production of another child carrying the gene.

If we imagine, instead, that adoption is a memetic trait, and one which can be passed to children in the care of an adult with the trait (through religious or moral teaching, for instance), then we have the following situation:

MOTHER	FATHER		
AA	BB	<- Genes	
cc	dd	<- Memes	
(BIOLOGICAL CHILD)	(BIOLOGICAL CHILD)	(ADOPTED CHILD)	
AB	AB	MN	<- Genes
cd	cd	cd	<- Meme

So in this case, if we assume that the "c" meme was the one that encouraged the trait of adoption, that meme has not been hurt in the slightest by adoption -- in fact it has been helped. And the adoption of a child already born may be a more reliable way to get a subject for a meme than waiting for a pregnancy and the attendant possibilities of miscarriage, infant death or defect, etc.

The meme that encourages the trait of adoption, if that meme also encourages passing itself on to the possessor's children, is a meme that has a built in method of propagation to a wider and wider segment of the "meme pool." It doesn't require the possessor of the meme even to be biologically capable of having children.\*

One would expect, if this meme theory of the adoption trait is true, that adoptive parents would be more likely to prefer to adopt children which A) will be more likely to absorb the adoption meme themselves, and B) will be likely to pass that meme on to other children.

As a child passes infancy and early childhood, it becomes harder to implant memes on the child. This theory would suggest, then, that adoptive parents would prefer to adopt younger children.

If a child is unlikely to reach adulthood, or if when it reaches adulthood it is unable to care for children, it will be unable to pass on memes to its children. This theory would suggest, then, that adoptive parents would prefer physically and mentally healthy children.

\* Note that a possible complication of this meme becoming prominent in the "meme pool" is that it would encourage a genetic adaptation to it -- i.e. a gene which would encourage the possessor to have children and place them up for adoption, without adopting any children of the possessor's own. A gene which encouraged the possessor to put all of his/her energy into producing children which are adopted by other families would be at an advantage in such a world and would tend to counteract the adoption meme.

In fact a gene which had the effect of immunizing the bearer from the adoption meme would be a successful gene, because the possessor of that gene would be encouraged to put its resources into raising only its genetic children. On the other hand, a meme which was able to overcome this immunization would also be successful, because it would pass itself to more children (if, for

example, the possessor of that gene was unable to have biological children, the meme would survive while the gene would not)

Genes and memes are likely to be at odds like this often.

## THE BLUE STAR MEME: APPLYING NATURAL SELECTION THINKING TO URBAN LEGENDS

By  
**Dave Gross**

How does an urban legend like the "Blue Star" LSD Tattoo legend spread, why do people believe it and want to encourage other people to believe it, and what does this say about the rest of the information we encounter and form opinions about? Perhaps surprisingly, insights from Darwin's discovery of natural selection may help us answer these questions.

The principle of **natural selection** is best known for its success in explaining the evolution and variety of life. But it is not restricted in its application to living things. In short,

- if a population is composed of mortal individuals that reproduce,
- and different varieties of individuals in this population are more or less able to reproduce relative to each other,
- natural selection will take place, favoring those individuals with the greater ability to successfully reproduce versions of themselves that also successfully reproduce.  
Individuals with greater **reproductive success** will extend the reach of their form into future generations, at the expense of less-successful individuals.

If the individuals in the population reproduce with imperfect fidelity - if not all of the reproductions are perfect reproductions - there may occasionally be novel forms of individuals that have advantages over individuals in the existing population.

The genes that code for attributes of biological organisms fit this model well - the model was, after all, developed to explain the emergence and variety of biological organisms - but in fact the model is more general than this. Any group of **replicators** that satisfies the above criteria will be subject to natural selection - whether they are biological in nature or not.

Good ideas, slang terms, chain letters, recipes and **urban legends** are among the non-biological entities that qualify and that are subject to natural selection. They reproduce with imperfect fidelity, and some reproduce better than others, so an evolution of sorts will take place, in which certain of these cultural replicators will flourish and others will die away.

### Is there an *unnatural* selection?

The adjective "natural" was put in the phrase "natural selection" to differentiate the sort of evolution caused by the selective breeding of race-horses or prize roses from the sort of evolution that happens without such conscious direction by people.

Clearly, recipes and chain letters are spread because of individual people who make conscious decisions to spread some but not others. The term "natural selection" may seem inappropriate to describe this process.

This distinction - between "natural" and "unnatural" or "directed" selection - was useful initially in explaining the theory of evolution, but it is no longer a useful distinction. *All* selection is natural selection. We are all part of nature, and if we select and breed a certain kind of sheep, it is no less natural that we should do so, than if bees, by landing on more colorful flowers, select and breed a certain kind of plant.

Furthermore, what the "Blue Star" urban legend shows us is that while the transmission of ideas may be a voluntary pursuit, people are not always conscious of the ideas they are spreading - they may believe that they are spreading an important message about a real danger when in fact they are spreading a collection of durable and virulent untruths. While a person who decides to spread such a message is acting deliberately and is selecting the message consciously - it is more accurate to say that the message has evolved into its form in order to encourage people to spread it, than to say that people have selected or "bred" the message into its form.

In Richard Dawkins's book *The Selfish Gene*, he coined the word "**meme**" to serve as the counterpart of "gene" when talking about natural selection as it applies to, in Dawkins's words, such things as "tunes, ideas, catch-phrases, clothes fashions, ways of making pots or of building arches." The name caught on, and an incipient science of "memetics" has formed around it. The "blue star" urban legend is a wonderful example of a meme, and can be compared to a **virus**, in that it briefly takes over the information reproducing abilities of a human mind - which we like to think are usually devoted to reproducing accurate, useful or at least intentionally amusing information - and uses them to reproduce more instances of itself.

## How the "blue star" meme mutates

It has evolved to do this by mutating over time - the **mutations** being caused by errors, modifications and elaborations introduced when the warning message is copied. Those mutations that made people more eager to spread the warning survived and expanded in the meme pool. Examples of mutations include:

**"The drug is absorbed through the skin simply by handling the paper"** - This is a recent version of a phrase which several years ago read "contact with moisture and skin could cause the same effect as taking a dose of acid orally." Note how the "could" has been changed to "is" in the new version. This increases the urgency - it's not enough to tell your child not to lick the blue stars; even casual handling of an innocent-looking tattoo will cause the child to be contaminated.

**"A young child could happen upon these and have a fatal 'trip'"** - At some point, the warning was embellished by this note that not only are these tattoos dangerous - they can be deadly. A few years later, versions of the warning would include the phrase "young lives have already been taken" which would in turn mutate to "tattoos have already taken the lives of many young children!" At each step, the warning gets incrementally more frightening, and the importance of noticing and spreading the meme becomes more urgent.

**"Please copy and post it at your work, give to friends..."** - Most all of the warnings have a plea for people to reproduce the warning and to spread it. Sometimes, multiple, overlapping pleas for reproduction exist. For instance, a typical e-mail warning included the following pleas to reproduce the meme:

- "Please pass it on via e-mail."
- "Please feel free to reproduce this article and distribute it as widely as possible in your community, schools, churches and the workplace."
- "Please post it in public places and bulletin boards. Please! Get the word out as quickly as possible. This is growing faster than we can warn parents and professionals."
- "Please do your part in educating our young by being an informed parent or guardian."

**"They are known to react quickly and some are laced with strychnine"** - The strychnine-in-acid motif is its own urban legend, believed by many people both inside and outside of the drug culture independently from the "blue star" urban legend. This appears to be a case of cross-pollination, whereby an especially vulnerable host for information viruses was holding both of these legends simultaneously, and, noticing the commonality of LSD in both memes, combined the two. This proved to be a very popular combination and almost all warnings circulating today mention strychnine.

**Sources** - A number of sources and authorities have been listed as originators of the warning over the years, including specific institutions ("Beth Israel Medical Center in New York"), generic institutions ("various military organizations"), and even specific people. Almost all of the warnings found today, for whatever reason, are attributed to a "J O'Donnell" (O'Donnell being spelled sometimes with one 'n' or one 'l' or both) of "Danbury Hospital's Outpatient Chemical Dependency Treatment Service" (words sometimes rearranged or missing). Danbury Hospital in Connecticut claims that they've never had a patient or staff member by that name, but that doesn't reduce the value of the supposed authority to the meme.

**Tattoo types** - The various types of tattoo (in the earliest known versions just "Mickey Mouse" - but now including "blue star," Superman, Red Pyramid, Butterflies, Clowns, etc) are specifics that add credibility to the meme. Many of these are innocent-seeming images associated with child aesthetics which magnifies the horror of their hidden contamination and the likelihood that a child will be attracted to them. Many also (e.g. "Mickey Mouse," "blue star") are actual designs found on LSD blotter, which gives the legend more credibility among people who have heard about busts involving blotter acid but are confused about the difference between this and "tattoos."

Some of the mutations are less adaptive, and seem more random. "Window pane" is a term used to describe LSD distributed in a flat, translucent gelatin medium divided into a grid of squares. Sometimes mutated descriptions of "window pane" acid work their way into the flyers. Some that I've seen include:

- "Another sort is 'WINDOW PAIN' which has a grid which can be cut out."
- "...and another, that can be cut out, called WINDOW PANE which has an acid."
- "...and another called WINDOW PANE which has an acid that can be cut out."

- "Another kind called "WINDOW PANE", which has a grit that can be cut out..."

These mutations seem to be just random errors and misconceptions, resulting in nonsense that adds little to the reproductive success of the meme. The sentences that include information about "Window Pane" acid are not often found in versions of the warning seen these days, and probably got left out during reproduction when reproducers stopped being able to make any sense out of them (what's "a grit that can be cut out" anyway?).

## Why people spread the meme

In the good old days, before the meme found email, it was generally reproduced without explanation or disclaimer, and so it was difficult to discover the motivations of the people doing the duplication, except through news articles about the legend in which its spreaders were interviewed. When people spread the message in e-mail or on the usenet newsgroups, however, they frequently tack on a sentence or two about why they have decided to spread the meme. These fall into a few categories. First, the "maybe it's not true, but in case it is true, I thought I should tell you:"

- "I don't know how much apparent danger there is. We just felt like we could not ignore if it was occurring. It was enough of a sufficiently gray area that I felt like I didn't want to take a chance."
- "I have not personally substantiated it, but the topic is too important to ignore out of hand."
- "[W]hether it is true or not, I thought that its content was serious enough that it warranted posting."
- "I'm just passing it on because if there is an inkling of truth, then it's important!"
- "I hope it isn't true, but I will talk to the boys about it anyway, just in case."
- Another variation of this category is "I'm going to pass it on because it sounds important" (its accuracy is not a consideration):
- "The content is so important that I want to post it here."
- "I know nothing else about these drugs, I simply posted this message when I read about it today."
- "When we hear about these things, we don't attempt to confirm or deny them. We simply send it out to emergency rooms across the region..." [*Los Angeles Times* 9 Dec 1987]
- "I felt that if it was something that concerned the safety and well-being of our students, then the parents ought to know about it."
- Then there's the "this is true" assertion sometimes masked by "we haven't seen it... yet:"
- "I don't want to start a panic wave, but I would like to point out a serious problem!" (Translated from the German)
- "We're not saying it is present in any particular community, location or school. We just want people to know that the possibility exists."
- "This may not be in your area yet!!!! But I am passing it along as an awareness process."
- "I got this through the agency so it must be really going on in the streets."
- And there are others that are in the same neighborhood as the above two categories:
- "I do not know the source or if it is true but it sounds real enough to me."

- "You feel like if it's happening, you want to let parents know. We didn't make a big issue of it, but we wanted to pass it along."
- "I was really concerned about this. I photocopied it and gave it out to some parents."
- "With drugs, if you're going to err, it's better to do so on the side of extreme caution." (given as a justification for spreading the warning)
- "I thought about the youngsters and the children who are entrusted to me. My spontaneous reaction prevented me from verifying the veracity of this 'information.' My good faith was abused and I may have been careless." (translated from the French)

The meme is so influential that some people, even when the urban legend status of the meme is pointed out to them, refuse to abandon belief. Some point out police busts involving blotter acid and insist that this proves the legend to be true. Others attack the motives of the people debunking the legend, like one mother who wrote: "...thank you for pointing out that it is an urban legend. I'm sure all the drug dealers out there will be happy to hear that."

Others who acknowledge the mythical status of the warning, still insist that it should be spread. A participant on a parents' newsgroup wrote: "Parents aren't necessarily concerned with truth, where it could be a gray area. We are concerned with our children's safety."

People are generally baffled as to why an almost completely inaccurate warning message like this keeps spreading. From the point of view of meme theory, it makes perfect sense. Without meme theory, though, people have less satisfying theories:

**Hoax** - Somebody is spreading this rumor as a sick joke, spreading panic through society and laughing behind our backs.

**Conspiracy** - People are spreading the rumor to further demonize drug users and dealers as a way of promoting escalation of the war on drugs. (Or alternately, drug enthusiasts are spreading blatantly false anti-drug information as a way of undercutting the credibility of anti-drug messages).

**Truth** - Maybe the legend *is* true after all, just slightly exaggerated.

**Prejudice** - People are hostile toward drug users and dealers and make up stories about their evil ways in the same way that anti-semites will concoct elaborate conspiracy theories about Zionists.

**Human nature** - Exchanging myths is a recognized aspect of culture and this is just another example.

While some of these have an element of truth to them - prejudice does play a part in making the meme effective, elements of truth can sometimes be found in the legend itself, and spreading urban legends does seem to be a part of human nature - looking on the "blue star" urban legend from a memetic point of view makes a lot more sense and leads to much more satisfying conclusions.

The main shift in thinking that needs to take place is to look at the spread of the legend not so much from the point of view of the people who propagate the warning, but from the point of view of the warning itself. Instead of asking "what made this person want to spread this warning," ask "how did this warning acquire elements that make people want to spread it." Those memes that include elements that interest people and encourage them to spread the word will survive, reproduce and flourish at the expense of less attractive versions.

What we learn from urban legends is that people will assist in the spread of a meme regardless of its usefulness or accuracy if the meme is well-constructed and virulent - that usefulness and accuracy are not necessary elements of a successful meme.

## **Take notice of your own vulnerability**

This legend gives us all a good opportunity to try to discover which viral memes we have been infected with. Everyone spreads some memes, and everyone has certain criteria by which they decide which ones to spread and which ones not to spread. Nobody I've ever met personally verifies the truth of every piece of information they pass on, so we're all vulnerable to memes that satisfy our reproduction criteria despite being untrue.

Some of the criteria I use when deciding whether to pass on a piece of information are:

**Truth** - I like to think that the information I'm sharing is accurate. If it comes from a source I respect, if I've personally verified it, or if it seems very unlikely to be false, I'm more likely to pass it on. For some memes, like jokes or puns or opinions of style, truth is not an issue at all.

**Prestige** - I'm more likely to spread a piece of information if spreading it will make the people who receive it think more of me (or less of my enemies or competitors). Clever jokes or anecdotes - things which paint me as a witty person - fall into this category, too.

**Interest** - I like to spread memes that I find to be novel, intriguing and challenging.

**Correctness** - I feel good about spreading memes that support my political or social viewpoints.

**Importance** - If I think that by passing on the meme, I'm helping to avert a crisis, or to improve the world in some way, I'm more likely to do so.

Looking at this list, and this paper, I see that memetics is a subject that seems, from personal introspection, to have a close fit with reality (**Truth**). I hope that people will read this paper and think, "Wow, what a brilliant way of looking at information!" (**Prestige**). It is a philosophy of epistemology, linguistics and information that strikes my curiosity (**Interest**). It demonstrates that popular delusions, for instance the "blue star" meme, can have a strong affect on public perceptions, and thereby a detrimental effect on government policy, for instance the drug war (**Correctness**). I think that if more people looked upon information as memes to which they may be vulnerable, they would add a level of skepticism to their information processing that would be healthy for them and for society in general (**Importance**).

I'm happy to spread the memetics meme, and I recognize it as such. I have to be aware though, that if a meme offers me enough in the prestige, interest, correctness, and importance categories, I'm vulnerable to help its spread regardless of its truth value. I also have to note that such a meme, in order to have survived long enough to encounter me, will probably be somewhat plausable and attractive (and will have fooled other people successfully), and that to the extent that I do not personally verify a meme with strengths in these categories, I will be successfully targeted by these memes.

My response to this is to do the opposite of what my inclinations are and what most people seem to do. The natural thing is, if a meme promises a great deal in non-truth-related categories, to lower the threshold of acceptance in the truth-related categories. The justifications for spreading the "blue star" meme quoted above (e.g. "I have not personally substantiated it, but the topic is too important to ignore...") are examples of this.

In light of meme theory, however, one should do just the opposite, since an untrue meme is more likely to survive and propagate if it promises a great deal in non-truth-related areas (the same goes for advertising and campaign promises) - therefore the untrue memes you encounter will probably be very tempting.

Another meme I like to spread is the "George Washington grew cannabis" meme. I strongly suspect that it's true; I've read sources that sound resonable that claim that entries in Washinton's diaries (May 12-13 and August 7, 1765 for instance) explicitly say that he was growing cannabis hemp, but I've never verified this personally, and if I were to be honest with myself, I'd have to say that the sources I've seen this information in are probably biased. But I'm tempted to take this on faith because it supports my own views that cannabis isn't so bad (**Correctness**), I think it's a curious bit of history (**Interest**), and I think it may influence people to support a more enlightened public policy toward marijuana users (**Importance**). To follow my own advice, then, I would have to be extra-skeptical about the truth of this meme.

A brief look through the alt.folklore.urban FAQ will probably show you a number of untrue bits of information that you yourself have enjoyed spreading.

For each one, try to remember why you decided to spread the information despite not being sure of its truth. By doing this, you will be able to map out the range of your vulnerability to deceptive memes. This is nothing to be ashamed of - we are **all** susceptible to this sort of memetic infection - but it is good to be aware of which categories of ideas are likely to slip past your defenses.

An important lesson that meme theory teaches that might not be learned from a non-memetics-based set of critical thinking skills is this: **An untrue idea does not have to be a lie in order to spread. Even if everybody who spreads it believes it - even if nobody benefits from people believing an untrue idea - it still may just be an effective, deceptive, memetic virus.**

## Conclusion

Meme theory compells us to look upon ideas as almost independent creatures, in symbiotic relationships with our minds and our cultures. All shared ideas are memes in the cultural meme pool, and all compete for survival. Importantly, an idea's usefulness and truth are only two of several elements that lead to the survival and propagation of memes - it is possible for memes that are both useless and false to survive and prosper.

Even memes that are essentially true are frequently embellished with false details that help that meme survive in the meme pool. (For instance, I've seen newspaper articles describing LSD busts in which the seized blotter acid was described as "tattoos" - this in articles which never mentioned the "blue star" warnings. The "tattoo" reference was just a meme held by the reporter or one of the reporter's sources that found its way into the article - thereby reenforcing the "LSD comes in tattoos" meme held by anyone who read the article)

By understanding how memetic evolution works, it should be easier for you to examine those memes you encounter in your day-to-day life with a critical eye - using extra skepticism on elements of a meme that seem likely to assist in its spread.

This may help you gain a more accurate view of the world, of culture, and of human nature.

## Further reading:

- alt.folklore.urban
- alt.memetics
- Brunvand, Jan Harold *The Choking Doberman And Other "New" Urban Legends* W.W. Norton & Company, New York 1984.
- Burkardt, John *Meditations on the Chain Letter*
- Dawkins, Richard *The Selfish Gene* Oxford University Press, New York 1976.
- Dawkins, Richard "The 'St. Jude' mind virus" *Nature* 371, 23, 1994.
- Gross, Dave The 'Blue Star' LSD Tattoo Urban Legend
- Lynch, Aaron *Thought Contagion: How Belief Spreads Through Society* (In press).
- Mikkelsen, Barbara & David P. Faxlore/Netlore
- Rollins, K.A. The Little Kimberly Anne FAQ
- Rosenberger, Rob Computer Virus Myths
- VanArsdale, Daniel W. Chain Letter Evolution

## **SOME REFLECTIONS ON CREATION VERSUS EVOLUTION OF MEMES**

By  
Dave Gross

There's a "creation vs. evolution" debate due to hit the world of memetics any time now. I can't seem to shake the impression that true individual creativity, and not just selection operating on recombinations and mutations in the social arena, is responsible for the great variety and complexity of memes that surround us.

I suspect that the counterintuitive alternative -- that the world's memes are the product of natural selection and not the accumulated treasure of creative consciousnesses -- may be an unwelcome truth in the same way that Darwin's assassination of God was unwelcome.

One potential solution I came up with to explain individual creativity from a selection-oriented view is this:

People engage in an inner dialog -- a thinking process that involves the conceit of an exchange of words between actors on the subjective stage. In my case (one I don't suspect to be atypical), this dialog takes the form of different genres of drama -- sometimes arguments between advocates of opposite arguments on a position on which I am, as a whole, ambivalent; sometimes predictive role-playing in which I set up a series of actors to simulate real people from my environment confronted with different situations; etc.

My theory is that in the course of these inner dialogs, the same sort of mutations and recombinations take place as do in multi-person conversation. Each person becomes a unique evolutionary arena for memes, as some versions survive and thrive in this inner dialog and some die away.

So a person doesn't just reflect memes that they've acquired, but each person subjects each meme to a fertile inner memetic "ecosystem" before releasing it again into the social environment. This is the process by which new, novel and complex memes seem to be suddenly "created."

Unsolved is the mystery of why people think in a dialog form. It requires brains to take a thought that presumably is already in a form suitable for brain processing, compress it into a form which adds inaccuracy and ambiguity to the thought, subjectively simulate the process of hearing this compressed version out loud, translate this compressed version back into a thought, and then analyze the thought from a second point-of-view.

Perhaps this process is a form of memory, holding a thought in the delay of linguistic processing while a second personality (with a potentially valuable second opinion) is swapped in. Perhaps this is just a way of alpha-testing a linguistic formulation of a thought prior to release to see if it is as convincing as a collection of words as it is as a pre-coded thought. Or perhaps this process is a pathological form of behavior that is reinforced by parasitic memes (to whom it could potentially be very helpful).

# MEMETICS: A SYSTEMS METABIOLOGY

## Version 950220

By  
**Ron Hale-Evans**  
rwhe@ludism.org

### Note:

This paper is close to completion, but there are still rough spots where elaboration is needed; in some places you will see notes from and to myself on what needs to be changed. I was asked by a friend to put a "pre-release" version online. Comments are very welcome and may shape the final version of this paper.

In this paper, I will present an application of general systems theory to memetics, the study of memes. Evolutionary biologist Richard Dawkins coined the word "meme" to describe the similarity of ideas to genes. Dawkins says of memes in his 1976 book *The Selfish Gene*: Examples of memes are tunes, ideas, catch-phrases, clothes fashions, and ways of making pots or of building arches. Just as genes propagate themselves in the gene pool by leaping from body to body via sperm or eggs, so memes propagate themselves in the meme pool by leaping from brain to brain via a process which, in the broad sense, can be called imitation.

Researchers have also likened memes to viruses or germs in that they often seem to come and go in waves, much as epidemics of biological organisms do. One can take Nazism as perhaps the most familiar example of what can happen when a fanatic ideology spreads and gathers many followers. The memetic plague at Jonestown proved just as fatal for the followers of Jim Jones as any biological plague would have, possibly more so. (Henson 1987)

We can use memetics as an important method for studying the spread of ideologies, religions, cults, fads, and so on. One may find it worthwhile to so consider the witch hunts of the Middle Ages, the genocide in Kampuchea, the takeover by Fundamentalist Islam in Iran, as well as the already-mentioned Nazi and Jonestown massacres, all of them ideologically or religiously (hence memetically) oriented. (Henson 1987) Fundamentalist Christian groups exist throughout the United States holding beliefs that could prove extremely dangerous in the future, especially as we approach the years 1999 and 2000. In those "odometer years" we can confidently expect to see an epidemic of apocalyptic millenarianism, just as happened in the years 999 and 1000 (Mackay 1841) and has happened to a lesser extent in recent times (Chamberlin 1975). Should an individual affected with an apocalyptic millenarian meme come into power then, as Pat Robertson threatened to in the 1988 election, the (nuclear) results could prove disastrous for the world and the human race.

Memetics, though valuable, until now has for the most part used a model of linear causation, effectively saying memes *cause* information disease, much as early germ theory said germs *cause* biological disease. This seems a stage through which all sciences seem to go, but one that we need not prolong, since we already have the foundations of a general systems theory that we can apply to most disciplines. To that end I have written this paper.

Dawkins and others usually define memes as the informational equivalent of genes. Do there exist, then, any memetic forms that act like higher biological organisms? It would appear so. Memes can propagate in isolation, but they more usually propagate in larger units which Dawkins calls "meme complexes", Hofstadter calls "schemes", and Grant calls "m-plexes". Hereafter, for simplicity, I will use the term "meme complex" to refer to this sort of organism, and the term "metabiological" to refer to its class and to the class of mental and informational phenomena of all sorts; in other words, I will define "metabiological organisms" as a kind of organism that requires a biological system for its substrate, and the more general "metaorganisms" as a kind of memetic organism that uses biological, electronic, or nanotechnological system for its substrate. (Salk 1985)

Here I will defer debate over whether we should consider given metaorganisms as memes or meme complexes by noting that just as self-replicating biological organisms can exist at many different levels of organization and complexity, from genes to cells to multicellular organisms to societies to the biosphere, so can metabiological organisms, from words to ideas to ideologies to cultures to the nöosphere. For the purposes of this paper, therefore, I will define a meme as a component of a meme complex, and note that almost any meme can amount to a meme complex and vice versa. I define bits (1 and 0) as the exception; we cannot call them meme complexes since they constitute the simplest individual memetic components.

We can refer to some metaorganisms as "metaelectronic" as well as "metabiological"; in other words, they rely on electronic as well as biological substrates. This holds true for countless metaorganisms today -- for example, one might find such in the comedy routines of Monty Python. Many college and high school students know the classic routines, such as "Argument Clinic" and "The Philosophers' Song" entirely by heart and recite them on the slightest notice. Moreover, those sketches have spread from the brains of the Monty Python troupe, through television, movies, and audio recordings to millions of people around the world, into "cyberspace"; readily available computer files and programs exist that contain the material, and the Yale mainframe has a special facility for accessing them, from which they gradually migrate throughout the Internet.

The metabiological (or metaelectronic; for simplicity's sake I will continue to use only the term "metabiological") organisms known as meme complexes have many similarities to biological organisms: they mutate, they spread, evolutionary pressures select for them, and so on (Miller 1978, Lumsden and Wilson 1981, Swanson 1983, and many others). On the other hand, simple memes, as components of meme complexes, behave more like the individual units of biological heredity, genes. The analogy between memes and genes seems so striking that it tempts one to construe an actual systems isomorphism of the biological and the metabiological. Earlier proponents of memetics have acted rather cautiously in this regard. For example, Dawkins, one of the originators of memetics, believed that memes had no equivalent to alleles, as a strict isomorphism with genes would indicate they ought to. Swanson (1983), however, has pointed out that one can think of memes (or "sociogenes" as he calls them) that conflict, such as the rival ideas of Protestantism and Catholicism, as alleles.

Another biologist, Miller (1978) believes that, among other differences, there can exist no parallel to sex in memetic evolution, so that one must construe the similarity between memes and

genes as a weak analogy rather than an isomorphism. Yet perhaps one can find a parallel to sex in memetics. The avant-garde writer William S. Burroughs once said, "Language is a virus from outer space." While a language normally reproduces itself by contagion and asexually, as a virus does, languages can "mate" and produce new languages. For example, to oversimplify a bit, one can see in English the result of a cross between Old English (a Teutonic language) and early French, a Romance language. This cross occurred as a result of social intercourse between the Anglo-Saxons originally inhabiting England and the Norman invaders. There also appears to exist a memetic equivalent to the sex drive. One can see curiosity and ecumenism as urges toward mingling two or more memetic structures.

Since, as the ongoing Human Genome Project shows, we can store genetic patterns digitally (the biological becoming metabiological), one can see that genes seem merely a particular chemical reification of certain memes (namely those for building biological organisms). Both what we have been calling "memetics" and genetics, then, become subsets of the more general study of what Dawkins calls "replicators". It seems therefore something of a mistake to build memetics only by analogy to genetics; we need not a mirroring but a broadening, a generalization of both memetics and genetics to something we might call "replicatorics".

While most readers will see meme complexes as clearly metabiological, individual human personalities or egos may seem less obviously metaorganisms. Nevertheless, I will argue for this case. Perhaps one day through the emerging engineering discipline of nanotechnology, the possibility will exist of our mapping our brains in detail and uploading our minds into new, even inorganic bodies. If this proves the case, then nothing will stop individual human minds from reproducing asexually in their entirety by making "backup copies" to store in other biological or electronic media. People might see such a process as desirable. (Moravec 19??) In this way, one can see human minds as meme complexes themselves. *Les mèmes seraient memes*, so to speak. One can imagine a megalomaniac reproducing without limit and infecting new host bodies, like a virus; or perhaps a superintelligence using human personalities as "masks" for its processing capabilities so that it may interact better with intelligences of a lower order. (Lem 1984) Such superintelligences might well keep a store of human personalities and even exchange them with one another. In this way, personalities would come to differ very little from the memes we trade in contemporary commerce.

We may discover that this already occurs to a certain degree. Some metaorganisms other than human minds seem sapient and self-aware. For example, the *loa* of Haitian voodoo seem a sort of metaorganism that uses humans as *horses* or biological substrates. The same loa reappear in ceremony after ceremony: Papa Legba, Baron Samedi, and so on. More than one "copy" of a particular loa can attend a ceremony simultaneously.

These ancient meme complexes completely dominate the consciousness of their horses during the ceremony, and the horses will respond to the names of the particular loa "riding" them. Perhaps then, we can say the "software" that comprises the loa responds to interaction with others. One can argue that a loa cannot interact with humans without a human brain, and therefore does not seem intelligent of itself, but one can say the same of human minds: where can one possibly meet a human mind not presently reified in a human brain?

A similar phenomenon seems to occur in other religions and philosophies. In Christianity, it sometimes happens that a believer will become "possessed by a demon" or "filled with the Holy Spirit." In the nineteenth century there existed many "trance mediums" who would claim to speak for a "spirit"; today we know the same phenomenon as "channeling." Perhaps one can say that in a manner of speaking, there really does exist a "war in Heaven," an all-out battle for supremacy between "Christ" and "Satan," two quasi-real entities. If so, it certainly seems less clear-cut than most Christians would have it. Besides "the body of Christ" (Christianity) and "Satan" (that metaorganism which can infect or "possess" the suggestible, such as the nuns of Loudun, there exist many other metabiological entities all fighting for the ecological niches in the ideosphere: Buddha, Allah, the *kami* of Shinto, various *versions* of Christ, and so on.

Perhaps the loa and "spirits" exist independently of humans in the "ideosphere" or "nöosphere". Would the Magna Carta exist even if we destroyed all paper and electronic copies of it and everyone who had ever read it had died? Answers to this question seem more the result of one's taste in ontology than falsifiable scientific hypotheses. Nevertheless, some prominent psychologists have speculated on this topic. Jung believed in a "collective unconscious", which contained all human ideas. If such does exist, then metabiological organisms must live there. Jungian archetypes may manifest themselves as mega-memes in the ideosphere or collective unconscious. Jung said that archetypes have no content. An archetype may therefore function as a sort of memetic skeleton, a meme that itself binds together clusters of memes with informational content, which may include directives of the sort discussed above.

Why do some of the metabiological phenomena discussed above manifest themselves as separate, self-conscious entities capable of reproduction, while others do not? Part of the answer to this question may lie in the way in which these metaorganisms insinuate themselves into the memetic structure of the individual human. Most often people who carry these metaorganisms become infected by them during a sudden, almost catastrophic "conversion process". This goes for the Baptist speaking in tongues as well as for the Haitian speaking as Baron Samedi. When a trance channeler seems to manifest a metaorganism but has not become infected in a traumatic conversion process, one perhaps ought to question the veracity of the experience. While the horses of Voodoo and the speakers in tongues of evangelical Christianity gain little but the temporary attention of an audience during their experience, many trance channelers make a great deal of money from their "infections," perhaps motive enough for fakery.

As an interesting sidelight, one might here want to consider a book that came out in various editions in the 1980s, titled, in fact, *War in Heaven*. This book, which author Kyle Griffith claims he "channeled," propounds the doctrine that there exist in a very real sense spiritual beings masquerading as the demons and deities of Earth's religions and that these beings compete for human souls after death, draining the *élan vital* from them for sustenance, literally eating said souls. The author seems to have a strong working knowledge of memetics; the meme complex he has constructed makes promises (of beating the bad guys) and threats (of having one's soul eaten), seems self-consistent, seems to "explain everything," (remarking upon that feature of the meme in several places so we won't fail to notice it) and has catchy names for the villains (the Theocrats) and heroes (the Invisible College). Griffith as much as admits he cribbed the idea for War in Heaven from a hoax popular in 1950s science fiction magazines known as the Shaver Mystery, which enjoyed enormous popularity.

Biologists have a motto: "Biology is destiny." Yet our memes determine our behavior no less than our genes, and becoming less "mechanical" (in the Russian mystic G.I. Gurdjieff's sense of the word (Ouspensky 19??)) entails freeing ourselves from the tyranny of both. We can probably consider some meme complexes as inherently less dogmatic and therefore making their carriers less mechanical. For example, we can probably consider agnosticism ("God may or may not exist") as less dogmatic than either fundamentalism ("God exists and He told me that...") or atheism ("God does not exist").

For every sapient being, there exist certain "unthinkable" thoughts, and these unthinkable thoughts depend on the existing memetic structure of the brain in question. One can metaphorically consider these "unthinkable thoughts" as Gödel strings for certain thinkers, in other words, patterns of information that the organism in question simply cannot express. For example, I have sometimes referred to the number 666 (the numerological equivalent of the name of the Great Beast in the biblical Book of Revelations, which one will supposedly find written on the forehead or right hand of Satan's legions during the time of the Apocalypse) as the Gödel number for Christians; many of them regard it with superstitious awe, and one hates to think what some fervent fundamentalists would do if someone stamped 666 on their foreheads as a practical joke while they slept -- suicide seems a possibility.

Let us continue with this example. Suppose a memeticist multimillionaire offers a fundamentalist Christian, an agnostic, and an atheist a million dollars apiece if they will have the number 666 tattooed on their foreheads. While the Christian would find it unthinkable to wear such a tattoo, the agnostic would find it less so, and presumably the atheist would find it a trivial matter.

I assume here that the reader does not hold to fundamentalist Christianity, and therefore does not find the prospect of such a tattoo unthinkable. Since the atheist seems the only person practically guaranteed to end up with a million dollars in pocket at the end of the experiment, does that mean that we should consider the atheism meme complex as fundamentally more liberatory than the agnostic meme complex? Not necessarily. The atheist in our *gedankenexperiment* would give the matter little thought, and in this way differs little from the Christian, who would also give the matter little thought. Only the agnostic would bother thinking through the possible consequences of the tattoo consciously, and therefore one can say that only the agnostic has come to a rational, "conscious," "non-mechanical" decision. While we can predict the responses of the Christian and the atheist beforehand with a fair degree of accuracy, the agnostic's response remains unpredictable. I therefore consider the agnostic meme complex the most liberating of the three, since it allows for the greatest leeway in action.

If we consider some memes more or less liberatory than others, then we have another, basic and personal reason for studying memetics: to achieve the greatest freedom from the meme complexes to which we humans seem no more than "horses."

According to those who work within the discipline of general semantics, one can show a correspondence between use of the verb "to be" in a text and the text's dogmatism. For example, Adolf Hitler said that Jews "are" animals, and Christians say that atheists "are" damned. The former type of use ("Jews are animals") has been termed by general semanticists the "'is' of identity" and the latter ("atheists are damned") the "'is' of predication".

Since we have as yet no objective measure of a text's dogmatism, perhaps we should temporarily take the percentage of a text that consists of forms of the verb "to be" as a measure of its dogmatism and generate some hypotheses. If one can estimate a meme's dogmatism by the number of times it uses the verb "to be," and if dogmatism contributes to a meme's virulence, then perhaps we can experimentally test the virulence of memes by counting the number of times they use the verb "to be."

I would like to venture that, all else equal, a meme that often uses the "is" of identity and the "is" of predication will prove more virulent than a meme with the same informational content that uses them less often. Perhaps one could test this by counting the number of times various versions of a meme with the same informational content use the verb "to be" and testing how randomly chosen groups rate them for plausibility and persuasiveness

In general, if the hypothesis holds, it should prove possible to create a virulence-measuring program along the lines of the grammar checker through which I ran this paper, which will count the number of times memes use various forms of the "is" of identity and the "is" of predication and generate a rough estimate of their reproductive ability thereby. (Similarly, Francis Heylighen has proposed **formality** as a measure of memetic virulence; in other words, more formal speech is more virulent than less formal speech. One could use similar methods such as computer analysis, in this instance as well.

One researcher has already performed a rough relative measure of the dogmatism of the Declaration of Independence and the Communist Manifesto by counting the appearance of the verb "to be" in (???? 198?). He suggests that we adopt the use of a subset of the English language that he calls English Prime (E-Prime or E' for short), which completely avoids the use of the verb "to be". So that the reader may judge the accuracy of this paper more dispassionately, I have written it entirely in E-Prime. Only quotations from other authors use any form of the English verb "to be".

I present, below, some of the basic factors, as I understand them, which contribute to the spread of memes. These include both factors intrinsic to the memes themselves ("hooks" (Hofstadter 1985), probably so named after the musical motifs that make hit tunes so catchy), and extrinsic, societal, political, individual and world factors.

**TABLE 1: Intrinsic factors in memetic reproduction: "hooks"**

1. intrinsically rewarding factors
  - 1.1. promises of reward
  - 1.2. interestingness of meme
    - 1.2.1. promises of reward
    - 1.2.2. aesthetic qualities
    - 1.2.3. humor
    - 1.2.4. strangeness/novelty
  - 1.3. feelings of superiority
  - 1.4. scapegoating
  - 1.5. explanation of everything

- 1.6. usefulness**
- 1.7. self-directed reward**
- 1.7.1. unconditional reward**
- 1.7.2. for obeying**
- 1.8. other-directed conditional reward**
- 2. claiming intrinsically rewarding qualities**
- 3. intrinsically punishing factors**
- 3.1. threats of punishment: fear**
- 3.2. guilt**
- 3.3. self-directed conditional physical harm**
- 3.4. other-directed conditional harm**
- 4. mechanical factors in reproduction**
- 4.1. ease of reproduction**
- 4.1.1. simplicity**
- 4.1.2. comprehensibility**
- 4.1.3. reproductive ability of medium**
- 4.1.4. copying fidelity**
- 4.1.4.1. cohesiveness**
- 4.1.4.2. noteworthiness**
- 4.2. redundancy**
- 4.3. longevity**
- 4.4. plausibility**
- 4.5. adaptability**
- 4.5.1. syncretism**
  - 4.5.1.1. subsuming other memes**
  - 4.5.1.2. consistency with other memes: cooperation**
- 4.5.2. intolerance to other memes**
- 4.6. precise marketing**
- 5. claiming mechanical factors**
- 6. general commands to host mind**
- 6.1. faith**
- 6.2. commands to explore the meme complex**
- 6.3. commands to spread the meme**
- 6.4. duty**

**1. Intrinsically rewarding factors:** There seem to exist, from my survey, more uses of the carrot than the stick in memetic hooks, perhaps because people seem more likely to avoid memes that punish than memes that reward. People seem more likely to seek out memes that reward than memes that punish. However, memes can offer a "variable-ratio schedule of reinforcement" (Skinner 1974) that punishes more often than it rewards, yet which rewards often enough that people believe the meme and even spread it. One might see the meme that says "You can win money at Las Vegas" as a prime example.

**1.1. Promises of reward:** One of the ploys that memes use the most to get intelligent systems to spread them seems that of promises of reward. (Hofstadter 1985) Whether memes fulfill these promises or not seems of little consequence. (If they do, so much the better; then the meme

proves itself useful, a memetic trait we shall discuss later.) Many memes promise rewards in the indefinite future, or after death. One can see the meme complexes of many organized religions as perfect examples of this. The Christian meme says, "If you believe this meme, you will be rewarded by going to Heaven after you die." This promises reward and thus, in a way, proves intrinsically rewarding because the carriers of the meme can, in times of duress, think about the lovely, comfortable, and joyous existence promised them while in this vale of tears.

**1.2. Interestingness of meme:** Some memes prove successful partly because they seem interesting. A meme can seem interesting because it promises reward, of course, but can also seem interesting because carriers find it intrinsically rewarding in certain other ways: because they find it aesthetically pleasing, humorous, novel, or strange.

**1.2.1. Promises of reward:** See 1.1.

**1.2.2. Aesthetic qualities:** One might see Leonardo's *La Gioconda* (the "Mona Lisa") as a somewhat less trivial example of a meme that has spread by its aesthetic appeal, it having appeared in many unexpected places, artistic and otherwise, from Duchamp's *LHOOQ* to candy bar commercials.

**1.2.3. Humor:** Good examples of memes spread because of their humorous appeal seem the folk art photocopies which adorn offices and places of work all over the country.

**1.2.4. Strangeness/novelty:** Memes spread through strangeness or novelty probably include urban legends and ghost stories, the Discordian and SubGenius pseudo-cults of recent years, and the many Eastern religions and cults that appeared in America in the 1960s and have continued to appear and disappear to the present day.

**1.3. Feelings of superiority:** Another hook memetic organisms use to reward their host organisms seems the generation of feelings of superiority. *We* "are" saved; *they* "are" heathen. *We* "are" the Master Race; *they* "are" animals. The meme, of course, usually defines "we" as "vectors of this meme" and "they" as "non-vectors of this meme." The formula thus runs almost invariably, "People who carry this meme are good. People who don't carry this meme are bad. Therefore, vectors of this meme are better than non-vectors of this meme."

**1.4. Scapegoating:** Scapegoating seems a closely related memetic hook. By scapegoating, a group both defines itself as good and another group as bad, thus generating feelings of superiority, and seems to solve its problems at the same time. Thus Germany in the 1930s could blame all its economic problems on the Jews. Genocide, of course, solves nothing but the problem of dissent -- a problem only for those in power.

**1.5. Explanation of everything:** The memetic ploy of *explaining everything* seems related to scapegoating and feelings of superiority. Many of the meme complexes that have been most successful, including Marxism and Psychoanalysis as well as meme complexes more commonly recognized as religions, tend toward the all-encompassing. These "theories" and cosmologies can explain essentially all phenomena with which meme carriers come into contact in everyday life. As the philosopher of science Karl Popper pointed out (Popper 1965):

A Marxist could not open a newspaper without finding on every page confirming evidence for his interpretation of history; not only in the news, but also in its presentation--which revealed the class bias of the paper--and especially of course in what the paper did *not* say. The Freudian analysts emphasized that their theories were constantly verified by their "clinical observations".... Naturally, this all-encompassing explanation tends to have a dramatic stress-reducing effect, and in this sense, this hook (an emergent property of the meme complex itself and not usually a single meme) seems intrinsically rewarding. Moreover, the façade of usefulness that such a meme complex presents functions as a promise of reward.

More specifically, Marxism, Psychoanalysis, Christianity, Ayn Rand's currently popular philosophy of Objectivism, et al. explain too well *deviance from the meme complex*. Science seems perhaps the first meme complex in history that does not (ideally at least) explain deviance from its memes as faulty perception or reasoning that the deviant could correct by acceptance of the meme. Rather, the scientific meme complex contains the meme of the fallibility of its carriers.

**1.6. Usefulness:** Truly useful meme complexes exist also, as evidenced by the proliferation of life-enhancing technology since the adoption of the scientific method meme complex several hundred years ago. Moreover, useful memes extolling exercise, proper diet, hygiene, and so on have become widespread as well. Nor do scientifically based meme complexes seem the only useful ones. Through my own experience I can add that meme complexes that promote meditation seem useful in establishing a sense of well-being in the individual; and the various yoga memes also have something to recommend them. With less certainty, considering tracts such as Teresa of Avila's autobiography, I include meme complexes that promote prayer here as well, since it seems prayer may also prove useful in establishing well-being in the individual entirely apart from their promised contact with deity. Note that while certain memes in these meme complexes may prove useful, other memes within the meme complex may prove detrimental, yet they all come linked.

**1.7. Self-directed reward:** Some memes seem rewarding in a very concrete sense, yet one cannot call them "useful." For example, the meme that says that people ought to strive for material goods seems materially rewarding, but useful only in a trivial sense.

**1.7.1. Unconditional reward:** The materialism meme and other memes in the contemporary conservative meme complex seem unconditionally rewarding, at least *immediately*, and so it does not seem surprising that the complex spread so far in the 1980s. On the other hand, memes that promote kindness and development of oneself toward a spiritual ideal seem also rewarding, but not so immediately, so it should come as no surprise that these memes have traditionally been poor competitors for the memes that promote greed and self-interest (the laity has always outnumbered the relatively more ethically-advanced monks and nuns in any religion one cares to mention).

**1.7.2. Reward for obeying:** One would suspect that there exist also conditionally-rewarding meme complexes, i.e., memes that say, "If you obey this meme complex, you can reward yourself by method *x*," yet I can think of none. There exist, however, conditionally punishing meme complexes, as we will see below.

**1.8. Other-directed conditional reward:** Many memes do not reward the carriers, or do not reward the carriers alone, but reward others. They say, "Reward others who carry this meme." This way, the population as a system rewards itself for carrying the meme, and the meme gets spread. One can see this mechanism at work in the confirmation parties held by Christian parents for their children, in Bar and Bas Mitzvahs, and in any celebration organized by those people who carry the meme for those who do not or in whom the meme has not yet completely matured, such as the feast days organized by the Catholic Church for the laity.

**2. Claiming intrinsically rewarding qualities:** Memes need not only have any of the above hooks to possess an intrinsic reward; they need merely claim that they have these hooks, in other words, claim that they *are* useful, that they explain everything, that they *are* amusing or aesthetically interesting, and so on. The pseudoscientific meme complex of phrenology, for instance, makes claims of usefulness, and Warhol's Campbell's Soup "art" makes claims of aesthetic value. These and other claims, including the promises of reward mentioned above, seem analogous to camouflage in biological species; just as biological species have evolved to seem like other species to protect themselves, so have memetic species like the pseudosciences adopted the coloration of their legitimate cousins, the better to garner respectability, thereby attracting more followers to carry and spread their memes.

### **3. Intrinsically punishing factors:**

**3.1. Threats of punishment: fear:** Not all memes seem intrinsically rewarding. Some memes punish. Analogous to the promise of reward seems the threat of punishment. Just as some memes promise you will go to Heaven if you swallow the meme, so do some, often the same ones, threaten that you will go to Hell if you do not, thus punishing disbelief by fear. Again, whether the meme can justify these claims or not has only a partial bearing on the success of the meme.

**3.2. Guilt:** Again analogously, memes which have adopted the strategy of "We are good; they are bad" often punish by inducing feelings of guilt ("I am bad") in the carrier should the carrier deviate from the program set forth in the meme. One suspects that this factor contributes as least as much to church attendance as does the allegedly uplifting nature of the services.

**3.3. Self-directed conditional physical harm:** Some memes induce the carriers to inflict physical punishment on themselves should the carriers disobey the meme; the epidemic of flagellation during the Middle Ages in Europe seems witness to this.

**3.4. Other-directed conditional harm:** Memes can also achieve success by inflicting punishment on others who do not carry the meme, just as they can achieve success by rewarding others who do. Jihads, crusades, witch-hunts and witch-burnings, the McCarthy-era persecution of "Communists," government persecution of "political criminals," ostracization of "nerds" by junior high-schoolers, and so on and on, seem all examples of this punishment. (It would seem unusual if a meme could achieve success by merely claiming to have the punishing qualities above; usually organisms tend to avoid entities that seem punishing.)

**4. Mechanical factors in reproduction:** Dawkins (1976) acknowledges three qualities of memes that make for high survival value: "longevity, fecundity, and copying-fidelity." We have

already discussed what seems to make for a fecund meme; longevity and copying-fidelity **seem**, roughly speaking, mechanical factors in the reproduction of the meme, and we will discuss them below.

**4.1. Ease of reproduction:** One of the primary mechanical factors in memetic spread seems that of ease of reproduction, Dawkins's "copying-fidelity" being only one aspect of this ease. Ease of reproduction has many aspects: simplicity, comprehensibility, the reproductive ability of the memetic medium, and cohesiveness. If a meme lacks any of these factors, it will lose ground and minds to the more easily duplicated memes.

**4.1.1. Simplicity:** That simplicity constitutes an important factor in memetic reproduction seems fairly easy to demonstrate: compare how many people can sing "Twinkle, Twinkle, Little Star" to the number of people who can recite the *Odyssey in toto*. Clearly, the simpler a meme, the less time, space, matter, and energy it will take to reproduce and the farther it will spread. Moreover, this simplicity seems a form of intrinsic reward, a hook generated as an emergent property of the whole meme or meme complex.

**4.1.2. Comprehensibility:** Comprehensibility also seems an important factor. One does not publish a Hindi newspaper in Sweden if one intends to reach and influence the majority of the population. Similarly, one does not put signs solely in Spanish in a Japanese airport. In fact, we seem to be evolving iconic (picture) languages for international use, computer interfaces, and so on.

**4.1.3. Reproductive ability of medium:** The reproductive ability of the medium that expresses the meme also has much to do with the success of the meme. One sees many of the same pieces of xerox art repeatedly in offices partly because of the spectacular reproductive ability of photocopying. People will duplicate almost any meme if it appeals to them, if there exist few or no restrictions on its duplication, and if the medium that carries it (such as 8 1/2" x 11" paper in this case) has a high reproductive ability. One acute observer of information spread, Stewart Brand, dubs this principle "information wants to be free." (Brand 1987) Other good examples of this principle include computer software piracy and the flap over Digital Audio Tape (DAT). DAT, would have allowed consumers of music to copy music they like with the fidelity of a compact disk and for the cost of a blank tape. The music industry lobbied the US federal government rather successfully to prevent DAT from being used in this way, even to the extent of placing anti-copy chips in DAT recording equipment. (Brand 1987)

**4.1.4. Copying fidelity:** Dawkins's "copying-fidelity" seems a fourth important factor in memetic reproduction. Copying fidelity seems crucial in preventing mutation in memes. Mutation seems not always a bad thing, of course, since through innovation (mutation), we have a rich world culture -- but in the case of useful information, any mutation seems potentially dangerous.

**4.1.4.1. Cohesiveness:** Poetic form seems an important early form of copying fidelity. Rhyme and meter help prevent information from decaying, and act as a primitive form of error checking. Take, for example, the well-known folk heuristic meme:

*Thirty days hath September,  
April, June, and November*

Probably this meme would not have spread so far without rhyme and meter. "September" and "November" check each other in the rhyme and prevent one from substituting, say, "August."

**4.1.4.2. Noteworthiness:** Moreover, anything that renders a meme sufficiently noteworthy that one repeats it acts as a form of copying fidelity, another stratagem being humor, such as "My Very Educated Mother Just Served Us Nine Pumpkins" (the planets of Sol system in order away from the sun: Mercury, Venus, Earth, Mars, Jupiter, Saturn, Uranus, Neptune, Pluto).

**4.2. Redundancy:** Redundancy seems another mechanical factor in memetic reproduction that acts as a form of copying fidelity, but it also functions in another way, and so deserves a word of its own. Constant repetition of a meme increases the likelihood that those thus repeatedly exposed to it will catch it, just as someone repeatedly exposed to an influenza virus becomes more likely to catch it. "What I tell you three times is true," the Queen told Alice. This phenomenon seems partly behind one of the extrinsic factors in memetic reproduction, namely isolation of the targeted group, which we will discuss later. (Henson 1987)

**4.3. Longevity:** Longevity of the memetic medium seems also a factor in memetic reproduction. The more durable the medium, the longer the information it contains will survive to infect intelligent organisms. Dawkins (1976) gives as an example of the manifestation of this principle the Torah, which has survived perhaps partly so long because of the durability of written material.

**4.4. Plausibility:** Plausibility certainly seems a crucial factor in the reproduction of memes, at least inasmuch as they present themselves to their intended host organisms as fact and not fiction. Memes that seem true will more likely evolve and spread than memes that seem false. We have all heard the urban legends about people who took LSD in the 1960s, jumped out of windows thinking they could fly, and fell to their deaths, yet it seems unlikely, all else equal, that a meme that told of someone who took LSD and flew would spread very far in contemporary urban society, since it would seem highly implausible to listeners.

**4.5. Adaptability:** As with biological organisms, adaptability seems an important factor in memetic success. Memes have two main strategies available upon meeting allelic memes: they can attempt to subsume the other memes syncretistically, or they can attempt to destroy the other memes in competition.

**4.5.1. Syncretism:** Syncretism seems a fairly common memetic stratagem.

**4.5.1.1. Subsuming other memes:** Christian doctrine adopted many pagan cult figures as saints and many pagan celebrations as Church feast days, including Christmas and Easter. Buddhism subsumed many native religious symbols during its evolution in India, and Hinduism, which took over the ecological niche (or allelic position) formerly occupied by Buddhism, after the Muslim invasion destroyed much of Buddhism, syncretistically adopted the figure of the Buddha as an avatar of Vishnu, with one small catch: it said the Buddha was an avatar of Vishnu on

Earth intended to ***mislead*** people. In all cases of subsumption, the new meme at least relegates the old memes to an inferior position, or, to use the genetic analogy, they shift the old memes to another position so that they no longer amount to alleles of the new meme. Both the new and the old memes may then travel together.

**4.5.1.2. Consistency with other memes: cooperation:** Memes not occupying an allelic position can even prove helpful to new memes; they can then evolve a strategy of cooperation. In order for memes not to occupy an allelic position, they must prove consistent, or apparently consistent, with the new memes. Consistency with preexisting memes seems an important factor in a meme's success. If a meme seems consistent with old prejudices, it simply becomes another prejudice. If a person has already contracted the "We are good; they are bad" meme, and believes that their racial group (A) "is" superior to racial groups B and C, a meme that claims that racial group A "is" superior to racial group D seems likely to find easy acceptance.

**4.5.2. Intolerance to other memes:** Competitive destruction of other memes, however, seems at least as common as syncretism. Keith Henson points out, "From a meme's viewpoint, toleration of other memes is ***not*** a virtue. It is, in fact, a fatal characteristic for a particular meme, as memes inducing intolerance to other memes would soon displace it." (Henson 1987) "Monomemes" (Henson 1987) like Fundamentalist Christianity and Islam have become intolerant to most other memes and thus achieve more-or-less complete domination of the host organism, an extreme example of this domination being the "memoid", who carries a meme that overrides biological survival programs inherent in the human brain, and will risk their life to further the meme. (Henson 1987)

**4.6. Precise marketing:** One final characteristic commonly found in successful memes seems precise marketing. A Fundamentalist Christian meme that induces followers to recruit members from National Science Foundation meetings seems not likely to prove very successful. A meme such as that propagated by the Unification Church (the "Moonies") seems more likely to succeed when, as this meme does, it induces carriers to attempt to spread the meme among vulnerable first-year college students in a strange new environment.

**5. Claiming mechanical factors:** Again, it seems not always necessary for memes to have the mechanical virtues mentioned above, for example, simplicity, comprehensibility, etc., to achieve some success. The nineteenth-century planned language Volapük was enormously complex compared to its successor language, Esperanto, yet in its day Volapükists propagated it widely, merely ***claiming*** that it was simple and comprehensible.

**6. General commands to the host mind:** Some of the most successful strategies used by memes to reproduce themselves seem only peripherally related to reward and punishment. They include exhortations to blind faith, commands to spread the meme, and commands to explore the full meme complex.

**6.1. Faith:** Faith, says Richard Dawkins, means blind trust, in the absence of evidence, even in the teeth of evidence. The story of Doubting Thomas is told, not so that we shall admire Thomas, but so that we can admire the other apostles in comparison. Thomas demanded evidence. Nothing is more lethal for certain kinds of meme than a tendency to look for evidence. The other

apostles, whose faith was so strong that they did not need evidence, are held up to us as worthy of imitation. The meme for blind faith secures its own perpetuation by the simple unconscious expedient of discouraging rational inquiry. The meme for faith ("Believe this meme and do not question!") usually seems complemented by memes that threaten Hell and promise Heaven, so to speak, based on whether the person meeting the meme believes the meme or not.

**6.2. Commands to explore the meme complex:** Similar to memes which command people to believe the meme complex seem memes which command people to explore the meme complex, for example "Study the Bible with a friend this Sunday." If the person meeting this meme believes it and follows it, they will become exposed to a more complete version of the meme complex, in this case Christianity, and it seems much more likely that they will then contract it.

**6.3. Commands to spread the meme:** Commands to spread the meme, in other words, sentences to the effect of "spread this meme," seem one of the most common ways a meme gets its host organism to spread it. Christianity, for example, has a long history of proselytisation, and this seems partially because it has always had a memetic component that urged its spread.

Commands to spread the meme need not always be explicit. Donald Going, in (Hofstadter 1985) gives an example meme complex in which this hook develops implicitly from two explicit memes:

**System X**

*Begin:*

*X1: Anyone who does not believe in System X will burn in hell.*

*X2: It is your duty to save others from suffering.*

*End.*

Emerging from the two sentences in System X one can see a third, implicit sentence, "It is your duty to spread System X." Nor does the hook have to emerge from the meme complex proper. People who already believe they have a duty to save others from suffering will relatively readily spread a meme that claims that anyone who does not believe in it will burn in hell. ---- and the same holds for any claims they make, etc!!!!

**6.4. Duty:** Finally, the meme of duty itself has probably been a material influence on the spread of certain meme complexes, notably militaristic ones. This meme, linked with others, such as the faith meme, as in "ours is not to wonder why," makes for a particularly virulent and dangerous meme.

**TABLE 2:**

**EXTRINSIC FACTORS IN MEMETIC REPRODUCTION**

- 7. input transducer
- 7.1. isolation
- 7.2. events consistent with meme
- 8. internal transducer

- 8.1.** stresses in population increasing suggestibility
- 9.** channel and net
  - 9.1.** fitness of channel and net --
  - 9.1.1.** high information capacity
  - 9.1.2.** signal-to-noise ratio
  - 9.1.3.** ease of use
- 10.** decoder
  - 10.1.** availability of decoder units for population
  - 10.2.** ability of population to interpret memes
- 11.** associator
  - 11.1.** suggestibility of population
- 12.** memory
  - 12.1.** presence of consistent memes
  - 12.2.** durability of memory
  - 12.3.** ease of storage and recall
- 13.** decider
  - 13.1.** favor of central authority
- 14.** encoder
  - 14.1.** intelligence of those propagating the meme
- 15.** output transducer
  - 15.1.** attractiveness of those propagating the meme

## Extrinsic factors in memetic reproduction

Memes cannot propagate in a vacuum. They need intelligent, or at least living, systems to spread them, and conditions prevailing in these systems can greatly influence their reproduction. Miller (1978) has enumerated the major information-processing components of living systems. He calls them the input transducer, internal transducer, channel and net, decoder, associator, memory, decider, encoder, and output transducer. Below we will examine how factors related to each of these components of living systems, sometimes at the organismic level, sometimes at the societal, contribute to memetic spread.

**7. Input transducer:** The input transducer consists of the component of a living system that brings information into the system. In a human being this consists of the various sense organs. In a society it consists of the telecommunications devices that link it with other societies.

**7.1. Isolation:** As Henson points out (Henson 1987), the inputs left to a living system isolated from other inputs will tend to influence it relatively more strongly. Thus, cults such as the Unification Church tend to isolate their members from the outside world, and authoritarian governments such as the former Soviet Union, but also to some extent the US, tend to restrict the flow of memes across their borders from societies that differ greatly from their own. (Henson 1987)

**7.2. Events consistent with meme:** The input transducer subsystem also brings information about stressful events such as eclipses and other celestial phenomena, plagues, famines, and so on. These, particularly those phenomena that seem to fall in line with prophecy, particularly

apocalyptic prophecy such as that in the Book of Revelations in the New Testament, often have the tendency to increase suggestibility in the population, and increase susceptibility to apocalyptic memes. For example, during the Black Death in Europe there was a corresponding epidemic of millenarianism. (Barkun 1974)

**8. Internal transducer:** The internal transducer consists of the subsystem that gathers information about other subsystems within the system. In a human being, the nerves leading from the various organs and internal subsystems to the brain serve this function. In society, news crews, reporters, and so on serve this function.

**8.1. Stresses in population increasing suggestibility:** The society's internal transducer would tend to relay stresses within a population to the rest of the population. Intrasocietal stresses often make it more likely that new memes will find a niche; for instance, it seems unlikely that the Nazi meme would have found such acceptance in Germany in the 1930s had Germany not been undergoing extensive economic crises at the time.

**9. Channel and net:** The channel and net (or "net" for short) consist of the subsystem of a living system that transmits information to all parts of the system. In the human body, of course, this means the nervous system; on a larger scale, this means the telecommunications available to a society, including telephone, radio, television, mail, email (electronic mail), and so on.

**9.1. Fitness of channel and net:** The chief factor influencing memetic spread in the channel and net subsystem seems its fitness. This fitness includes, among other factors, a high information capacity, a high signal-to-noise ratio, and ease of use.

**9.1.1. High information capacity:** According to Miller (1978), information capacity seems one important variable in function of the net. High information capacity means that the channel and net can carry many signals at once. For example, the Pony Express could carry very many fewer letters cross-continent than today's fleets of mail planes; thus, the net can carry more signals in this way and spread more memes: propaganda, religious tracts, scientific information, and so on.

**9.1.2. Signal-to-noise ratio:** Signal-to-noise ratio seems another important variable in function of the net. (Miller 1978) This refers to the amount of desirable information that the channel and net subsystem can carry over the amount of undesirable information. A high signal-to-noise ratio means that the net can transmit much desirable information relative to undesirable information. Note that, for a given meme, any information that does not foster its own spread amounts to noise, no matter how important that information may seem to the recipients. Thus, if information about a recent important event overwhelms the net (death of a central political figure, moon landing, outbreak of war, announcement of a plague cure), a meme may find it a very hostile environment in which to spread.

**9.1.3. Ease of use:** Another important factor in memetic spread consists of easy use of the net. The easier the net to use, the more people will use it to spread memes. For example, whereas I rarely receive letters by regular US mail, I find it not uncommon to receive several personal messages in one day by the much easier to use email, or electronic mail, often bearing important new memes about technological advances. This advance in ease of use of communication has

also brought an increase in undesirable memes, of course; whereas I have only received a chain letter perhaps twice in my life by postal mail (commonly called "snailmail" by email users), I have received perhaps ten in the past two years by email.

**10. Decoder:** The decoder consists of the living system component that translates information input to it into a form usable internally by the system. In a human being, this component consists of part of the human brain; in society, it consists of both human brains with their literacy and language capabilities and the technology they have evolved (for instance, television sets) which allow them to receive information transmitted.

**10.1. Availability of decoder units for population:** One factor in the propagation of memes seems availability of decoding technology. The more technology available, the more information will individual organisms in the society prove able to receive, and memes will spread faster.

**10.2. Ability of population to interpret memes:** Literacy seems also an important factor in memetic reproduction, as do the factors of which and how many languages the population speaks. The better educated a population, the more memes it can interpret. (This does not necessarily mean that all memes that a well-educated population can interpret will necessarily fare well in that population; see "associator" below.)

**11. Associator:** The associator consists of the subsystem that carries out the first stage of learning within the system, that of forming associations between items of information. In the human being, the associator subsystem consists of the brain, and in society, it consists of schools, scientific institutions, and so on.

**11.1. Suggestibility of population:** The frequency with which the associator forms associations seems the major factor in its reproduction of memes; to put it another way, the suggestibility of the population seems a major factor in the spread of memes. The more suggestible the members of a society, the more memes they will contract and spread. Education can mitigate suggestibility; university students can read the *National Enquirer* but rarely do, and even then usually reject what they find within.

**12. Memory:** Memory consists of that subsystem by which a living system carries out the second stage of learning, storing associations and other sorts of information. In individual human beings, the brain carries out this function; in human society, both individual brains (as individuals store oral and procedural knowledge), and physical technologies, including books, tape, digital computer storage, and so on, carry out this process.

**12.1. Presence of consistent memes:** The presence of consistent memes in the memory of the targeted living system seems an important factor in the spread of memes. We have already discussed this factor from an intrinsic standpoint above, in the passages on adaptability of memes. Presence of inconsistent memes in memory contributes negatively to the spread of certain memes also, as already mentioned; education can have a drastically negative effect on the spread of memes that require high suggestibility in the target population.

**12.2. Durability of memory:** Durability of the memory also affects spread of memes, as discussed above under "longevity." The Odyssey propagated quite well as an oral tradition in ancient Greece, but it seems unlikely it would have survived until today had it not been written down at some point. Paper seems more durable a form of memory than oral repetition.

**12.3. Ease of storage and recall:** Ease of storage and recall seems also likely to affect the spread of memes. One can instantiate a meme in as durable a medium as one likes -- say, carving it in stone -- but if it remains inaccessible -- say, carved in stone in a Himalayan cave -- it will likely not propagate far.

**13. Decider:** The decider consists of the component subsystem that receives inputs of information from all other subsystems and outputs information to them that controls them. In the human being, this subsystem undoubtedly consists of the brain; at levels larger in scope, it consists of the board of directors, the papacy, the government, etc..

**13.1. Favor of central authority:** One of the main factors on any level influencing the spread of memes seems whether the memes being spread have already infected central authority (the decider component). If it has, the decider seems likely to put its considerable force behind the spread of the meme, and since it has information outputs leading to all systems, the meme can then spread very far. If the decider has not contracted the meme, it seems likely either to ignore the meme being spread, or to attempt to repress it, usually by spreading conflicting memes which interfere with the meme being spread (and incidentally acting as noise on the channel and net), or by punishing those who carry the meme, or both.

**14. Encoder:** The encoder consists of that subsystem that changes the form into it from a private form used by a system internally to a public form that other systems can use. This consists of the language centers in an individual human, and on a societal level might consist of the brain of a spokesperson for an organization, or a priest in a church.

**14.1. Intelligence of those propagating the meme:** The intelligence and skill of the encoder seems the main factor influencing memetic spread by encoder subsystems. Eloquence and tactfulness seem more likely to spread memes than do stammering and verbal clumsiness.

**15. Output transducer:** Finally, the output transducer consists of the subsystem that outputs information from the system to the environment. In a human being, this consists of the lungs and vocal tract, or the hands if writing or typing; in a larger system, it consists of the person or people proper who speak for the group, organization, or culture.

**15.1. Attractiveness of those propagating the meme:** One of the main factors in the spread of memes contributed by the output transducer seems the attractiveness of the output transducer. No one wants to listen to an ugly person with a screeching voice deliver homilies about Brand X Toothpaste or Brand Y Ideology, but people will listen long to a beautiful person with a mellifluous voice. Moreover, the attractiveness of the person sometimes acts as an implicit sort of promise of reward: "If you believe the meme I am spreading now, you can become as attractive as I."

## **Afterword**

Someone once sagely said that all science consists of either physics or stamp collecting. I hope the reader has found a few interesting specimens in these pages. They hardly present a complete systems theory of memetics, but memetics has not yet outgrown its infancy, and I hope this paper has contributed to its growth.

I would like to offer my thanks to my wife Marty Hale-Evans for her extensive helpful criticism and suggestions. I would also like to thank E. Jay O'Connell... Alex Chislenko...

# **AN OPEN MIND IS NOT AN EMPTY MIND: EXPERIMENTS IN THE META-NOOSPHERE**

**By**

**David Hales**

*Journal of Artificial Societies and Social Simulation* vol. 1, no. 4, 15 Oct 98

<http://www.soc.surrey.ac.uk/JASSS/1/4/2.html>

## **Abstract**

Using the "meme" conception (Dawkins 1976) of cultural transmission and computer simulations, an exploration is made of the relationship between agents, their beliefs about their environment, communication of those beliefs, and the global behaviours that emerge in a simple artificial society. This paper builds on previous work using the Minimeme model (Bura 1994). The model is extended to incorporate open-mindedness meta-memes (memes about memes). In the scenarios presented such meta-memes have dramatic effects, increasing the optimality of population distribution and the accuracy of existing beliefs. It is argued that artificial society experimentation offers a potentially fruitful response to the inherent problems of building new meme theory.

**Keywords:** Artificial Societies, Computer Simulation, Memetics, Meta-memes

## **Introduction**

1.1

If ideas are seen as replicating, mutating entities (replicating through people's minds via communication) then they can be viewed as "viruses of the mind" (Dawkins 1993). The analogy is that ideas spread through a population by "infecting" brains in a similar way to the spread of a virus. It is argued that memes are often successful because they induce their hosts to replicate them. Since Dawkins' seminal work, several other writers have used the meme concept to explain various cultural and social phenomena (Bonner 1980; Dennett 1995; Gabora 1997).

1.2

In order to investigate the implications of this conception of belief spread I have constructed a computational simulation of a simple artificial society. Several experiments have been conducted and some interesting and counter-intuitive results have been observed.

1.3

Various computational models of cultural transmission have been advanced (e.g. Axelrod 1995; Epstein and Axtell 1996; Reynolds 1994; Gabora 1995). "Cultural Algorithms" introduced by Reynolds (1994) augment standard Genetic Algorithm techniques with a belief structure of hierarchically organised beliefs and their generalisations. In the simulations he presents, group level fitness values are used to update the belief structure. The belief structure is then used to bias the selection of chromosomes for reproduction into the next generation. The model shows how cognitive abilities (generalisation in this case) and intergenerational cultural knowledge (the

belief structure) can be used to improve the performance of GAs when applied to tricky group co-operation scenarios. It is a high-level model which assumes the existence of shared cultural knowledge and group level selection. The model was not designed to address issues such as the spatial aspects of cultural transmission or the emergence of stable shared cultural characteristics from micro-level asynchronous cultural transmission and innovation. It is this latter aspect that memetic models attempt to address.

#### 1.4

The Sugarscape (Epstein and Axtell 1996) model uses strings of binary flags to represent cultural transmission units. Each agent randomly propagates flags to local neighbours. A function is then applied to the string in order to ascribe cultural identity [1]. The spread of such identities (or "tribes") can then be monitored and behaviours can be influenced by them (e.g. combat).

#### 1.5

It would seem that individuals in real societies are much more active in their selection of ideas, practices and beliefs. They often reject or "repel" new ideas and beliefs, particularly if they are currently strongly attached to contradictory ones. Attachment or confidence in a particular belief may grow if many others with whom the individual has come into contact also share such a belief (a form of "reinforcement" or "frequency dependant bias" (Boyd & Richerson 1985).

#### 1.6

A multiple agent model of meme spread which attempts to address these issues has been proposed by Bura (1994). This model is extended to incorporate meta-memes (ideas about ideas). A comparison is then made between three simulation scenarios with and without a particular meta-meme.

#### 1.7

Gabora's ( 1995) "Meme and Variation" Model is a more fine grained model which utilises neural networks to capture a form of meme evaluation. In contrast, the Minimeme model described in this paper is a higher-level model. It does not make assumptions concerning agents' cognitive functioning, specifically, agents do not evaluate memes. In this sense the agents presented here are more passive than those presented by Gabora. However, evaluation mechanisms should be seen as meta-memes. The Minimeme model can accommodate evaluation mechanisms but these would be implemented as meta-memes within the system. The argument is that evaluative functioning can be "boot-strapped" from a simpler meme model without building in such functions initially. However, this is not a dogmatic statement: in general artificial societies are constructed to answer specific questions and explore particular phenomena and as such different architectures are dictated by the requirements at hand.

### **The Minimeme Model**

#### 2.1

The model is composed of two parts: (a) the environment and agents, (b) the meme level (or noosphere [2]). The environments and agents differ from simulation to simulation, but the rules governing the noosphere do not.

## **The Noosphere**

### **2.2**

According to Dawkins (1993) any idea capable of transmitting itself from one person to another (replicating itself) is a meme. In Minimeme only memes that define behaviour are considered. Such memes can be "executed" by their hosts to produce an effect (e.g. movement, fighting, socialising etc.). In order to be successful and continue to exist, a meme must satisfy three conditions:

- it must find at least one host (an agent that stores it in its memory);
- the "execution" of the meme must not endanger the host's life (at least not before the meme has been able to reproduce itself);
- the meme must be able to resist the attack of opposing memes (termed "concurrent" memes) in the meantime.

### **2.3**

The sum of the memories of all the agents in the environment constitutes a space called the *noosphere*. Memes inhabit the noosphere in the same way that agents inhabit the simulated environment.

## **How Memes Evolve and Spread**

### **2.4**

To simulate the ability of the memes to conquer a part of the noosphere two parameters are associated with each meme: "change", which is a measure of the meme's propensity to mutate or to succumb to attacks by other memes and "aggression", which is a measure of the meme's propensity to try to reproduce itself. These parameters take real values in the range [0...1]. It is important to note that these parameters do not take into account the ability of the meme to keep its agent host alive.

### **2.5**

Memes evolve and spread in three stages:

- Satisfaction test: update change and aggression values;
- Mutation: mutate the meme in some way;
- Replication: attempt to spread the meme to other agents.

### **2.6**

First a satisfaction function is evaluated for each host. This function is simulation dependent. It may involve an estimation of the correct accomplishment of a task or the inspection of state variables in the host (e.g. is it hungry, ill etc.). The function should return an all-or-nothing result. Either the host is or is not satisfied. If the host is satisfied, it increases the aggression of each of its memes by 25 per cent and decreases their change by 25 per cent. Conversely, if the host is not satisfied, it decreases its memes' aggression and increases their change.

## 2.7

After this stage the memes may mutate and reproduce. A mutation occurs when a random draw in the range [0...1] gives a number lower than the meme's change. The actual nature of the mutation is simulation dependent.

## 2.8

If a meme was not mutated and if another random draw in range [0...1] is lower than its aggression, replication may take place. A random number of individuals are chosen among the host's neighbours (i.e. the ones it can communicate with) and the meme is proposed to each of them. If any of the neighbours are hosts to concurrent memes then a random draw in the range [0...1] is made. If this is lower than the attacked meme's change the meme is overwritten by the attacking meme (*replicated*) otherwise it stays in the host's memory (*repelling* the attacking meme). If a meme tries to infect a host that already possesses the same meme it is *reinforced* (its change is decreased and its aggression is increased).

## 2.9

It is important to note that hosts can learn new memes only by interacting with each other. Memes can not be coded into the environment or learned by experience.

## 2.10

These mechanisms are the same for all the simulations using the model. The characteristics to be defined for a given simulation are: (a) The satisfaction function for the hosts; (b) The nature of the mutations each meme can undergo; (c) The range of communication between hosts (i.e. how to find the "neighbours" of a given host).

# The World of the Grazers

## 3.1

"Grazers" are very simple agents who live in a very simple environment. They can move, feed (accumulate energy), die and communicate with others in their territory. The environment they inhabit consists of just four territories. Each territory can feed a fixed number of grazers during each cycle (a "carrying capacity"). Any number of agents can occupy a territory. Grazers have one decision to make in each cycle: whether to move to a new territory<sup>[3]</sup> or "stay put". Grazers try to maximise their energy (if it falls below a minimum, they die). The desirability of a territory is a function of its carrying capacity and the number of grazers that already occupy it. The grazers do not have knowledge of the carrying capacities of the territories, but they do have knowledge of the distribution of the population in each territory and as "grazers" they have a natural propensity to herd. They determine the desirability of each territory based on the number of grazers already occupying it. A grazer makes a decision with reference to a meme which tells it the ideal number of grazers that should occupy a territory. It makes a rational decision using its current meme<sup>[4]</sup>. This "herding" meme is represented by a single integer in the range 1 to 10. If a grazer possessed a '1' meme it would look for an empty territory (or the most empty if none were empty). Grazers mutate their memes by increasing or decreasing them by one.

## **Accumulating and Consuming Energy**

### **3.2**

Movement from one territory to another costs a grazer one energy point. If a grazer can not feed during the system cycle it loses an energy point. If a grazer can feed it gains an energy point (up to a maximum of 5 energy points). If there are more grazers in a territory than the specified carrying capacity, the grazers that will go hungry are selected at random. When the energy level of a grazer falls below 1 it dies instantly [5]. Newly born grazers start with a maximum energy level of 5 and take their memes from a random neighbour or generate them randomly if no neighbour exists. At the start of a simulation, the locations and memes of grazers are generated randomly. All energy levels are set to the maximum.

## **The System Cycle**

### **3.3**

- One pass through the following phases constitutes a single system cycle [6]:
- Action Phase: Each grazer gets a chance to move to a new territory.
- Environment Phase: Predators attack any vulnerable grazers (see below).
- Feeding Phase: Each grazer tries to eat from its current location.
- Meme Phase: Each grazer tests its satisfaction and then updates, mutates or spreads its memes (see above).

## **Three Simulation Scenarios**

### **4.1**

The following grazer simulation scenarios were implemented:

#### **A**

"Just Enough Food": The carrying capacities of all territories are set to 3. This means that there is one optimal distribution of agents: a 3-3-3 population distribution (three grazers in each territory). Such a distribution is optimal since it allows each grazer to harvest enough energy to stay alive. Intuitively one would assume that a noosphere dominated by the "3" meme would produce such an optimal solution. On reflection though, it can be seen that any noosphere totally populated with memes less than "4" should be optimal.

#### **B**

"Too Much Food": The carrying capacities of all territories are set to 4. An environmental constraint has been relaxed. This means that there are many possible optimal distributions. One might expect that such a scenario would give the grazers a better chance of finding an optimal distribution.

#### **C**

"Too Much Food with Predators": The carrying capacities are as scenario B but any territory which is occupied by less than 4 grazers is "attacked" by "predators" during the environment phase. Practically this means that all the grazers within such a territory have 2 energy points

deducted. There are four possible optimal distributions (4-4-4-0, 4-4-0-4, 4-0-4-4 and 0-4-4-4). Intuitively such a scenario seems to place heavy constraints on the possible composition of the noosphere. For example, if agents were distributed in one of these optimal arrangements it would not be stable if any one of the agents held a meme greater than "4" or less than "3". An agent holding a lower meme would move to the empty territory. An agent holding a higher meme would move to a territory holding four agents.

## Experimental Methods and Presentation

### 5.1

For the purposes of analysis, the model is iteratively executed until a stable noosphere is attained (termed equilibrium). A stable state is one in which the composition of the noosphere stays constant over time [7]. In such a situation "deviant" memes (those which destabilise the noosphere) will tend to be repelled and replaced by non-deviant memes through the process of replication. Such a state has parallels to the concept of an evolutionary stable strategy (Dawkins 1982). The noosphere defines the social behaviour of every agent. Any stable state could be said to be a viable social organisation (or "culture") since it persists over time even though agents may die and be replaced. Noosphere stability does not indicate the stability of other properties of the population such as death rates or the population distribution (which could be stable, seemingly random or periodic).

### 5.2

For each of the three scenarios, two experiments were performed, one without meta-memes and one with meta-memes (described below). Each experiment consisted of 100 simulation runs. The summary presented below (see Table 2) is therefore a synthesis of 600 individual simulation runs.

### 5.3

Results are presented for each experiment in the form of general observations based on a synthesis of 100 individual simulation runs. This synthesis is presented in the form of a surface contour map plotting  $x$ ,  $y$ , and  $z$  as, respectively, maximum density of agents in a single territory, most dominant meme in the noosphere and total number of such couples (i.e. maximum density / dominant meme) accumulated over all simulations. Each simulation represents a point on the  $x$ ,  $y$  plane. The cumulative distribution of these points is used to give a  $z$  component. This gives a contour map of the relative frequencies of stable noosphere compositions (based on the dominant meme) against an optimality measure (maximum density)<sup>[8]</sup>. The contour maps are shown in Figures 1, 3, 4, 5, 9 and 10. In order to illustrate the dynamics of individual runs, graphs are given showing the changing composition of the noosphere and the movement of agents between locations over time (Figures 2, 6, 7 and 8 ).

## Experiments without Meta-Memes

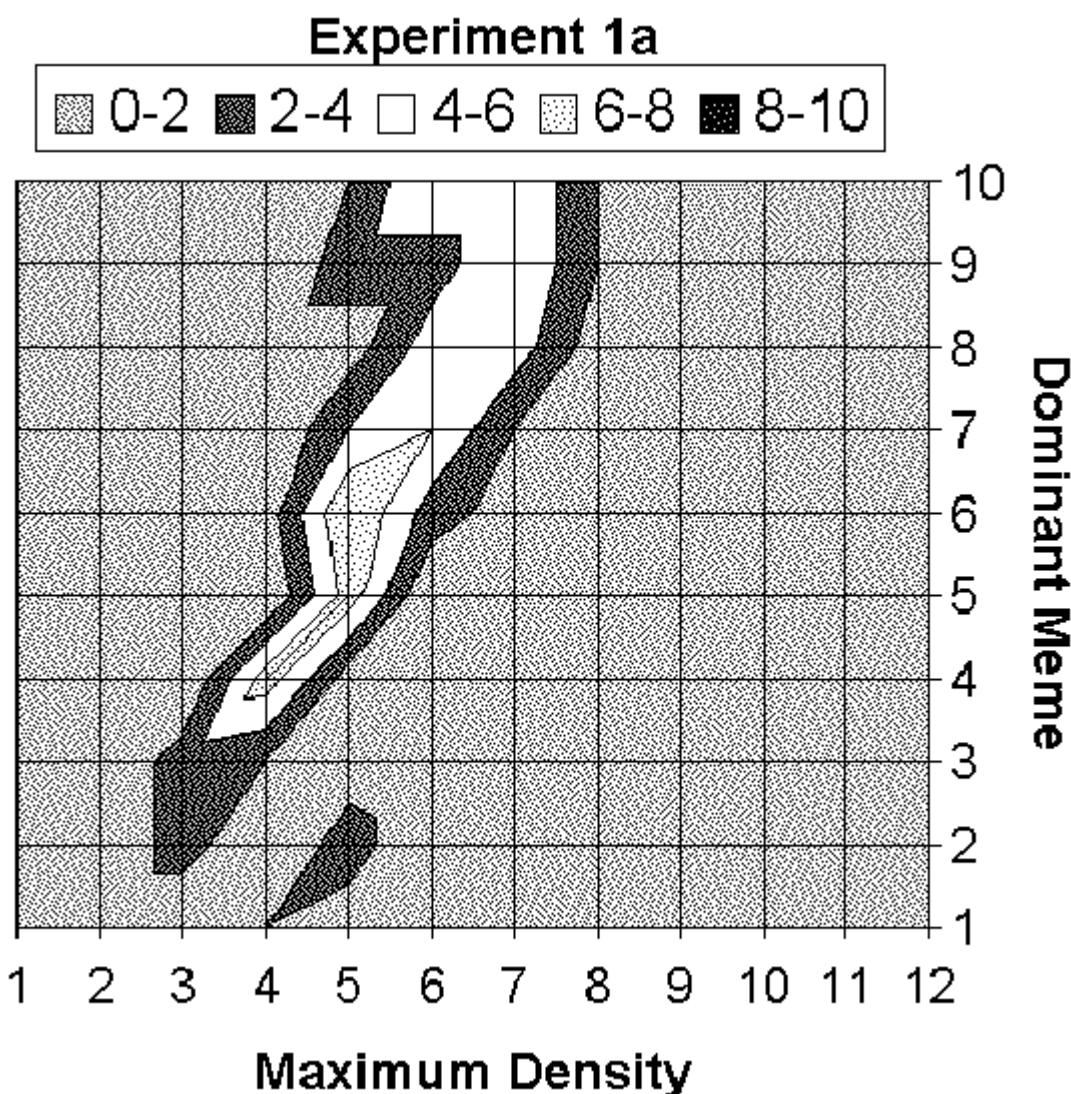
### 6.1

Each of the grazer simulation scenarios were initially executed without meta-memes.

### Experiment 1a - "Just Enough Food"

6.2

By the 1,000th cycle 76 per cent of the simulation runs had reached equilibrium. By the 3,000th cycle 97 per cent had done so. Most of the runs (94%) did not result in an optimal population distribution but the results are more optimal than would be expected from a totally random distribution (see Figure 1). The "self-catalytic" [9] process is strongly evident.

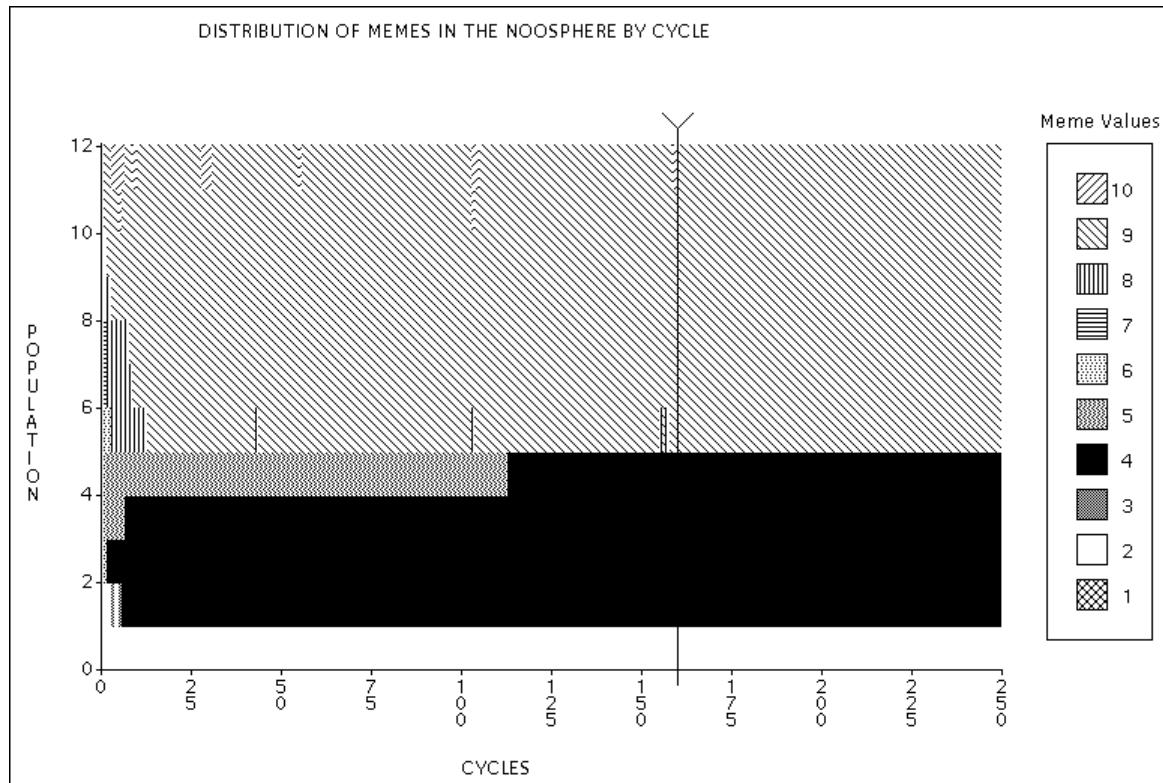


**Figure 1. Dominant meme / Maximum density synthesis**

6.3

Figure 2 shows the evolution of the noosphere in a typical run. There is a speedy domination of the noosphere by the "9" meme. This takes place via the self-catalytic process in a single overpopulated territory. Notice that the single "5" meme (cycle 15 to cycle 115) lasts for about

100 cycles before succumbing to the "4" meme (which becomes dominant within another territory). The death rate is high before and after equilibrium.



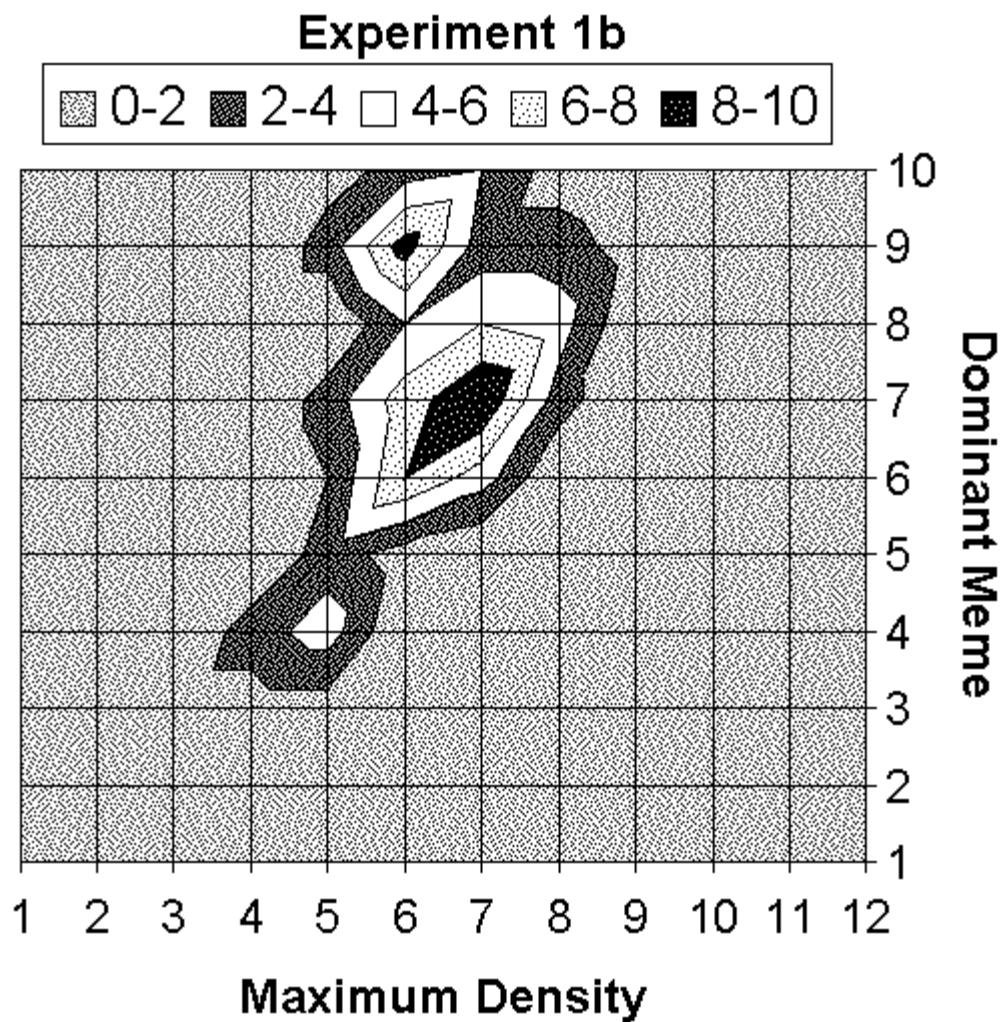
**Figure 2. Distribution of memes in the noosphere. Experiment 1a - Just enough food.**

Equilibrium is reached at cycle 160. The single black line through the graph indicates this point (just after the elimination of the single "10" meme). At the start of the run the "9" meme quickly takes over the whole of a territory. The "5" meme manages to hold out for over 100 cycles before it is replaced by the "4" meme (dominant in its territory).

### Experiment 1b - "Too Much Food"

#### 6.4

By the 300th cycle 92 per cent of the simulation runs had found an equilibrium. By the 800th cycle all (100%) had reached equilibrium. As illustrated in Figure 3, most of these are far from optimal. Relaxation of the environment constraint significantly speeds up the self-catalytic process due to the reduced death rate.

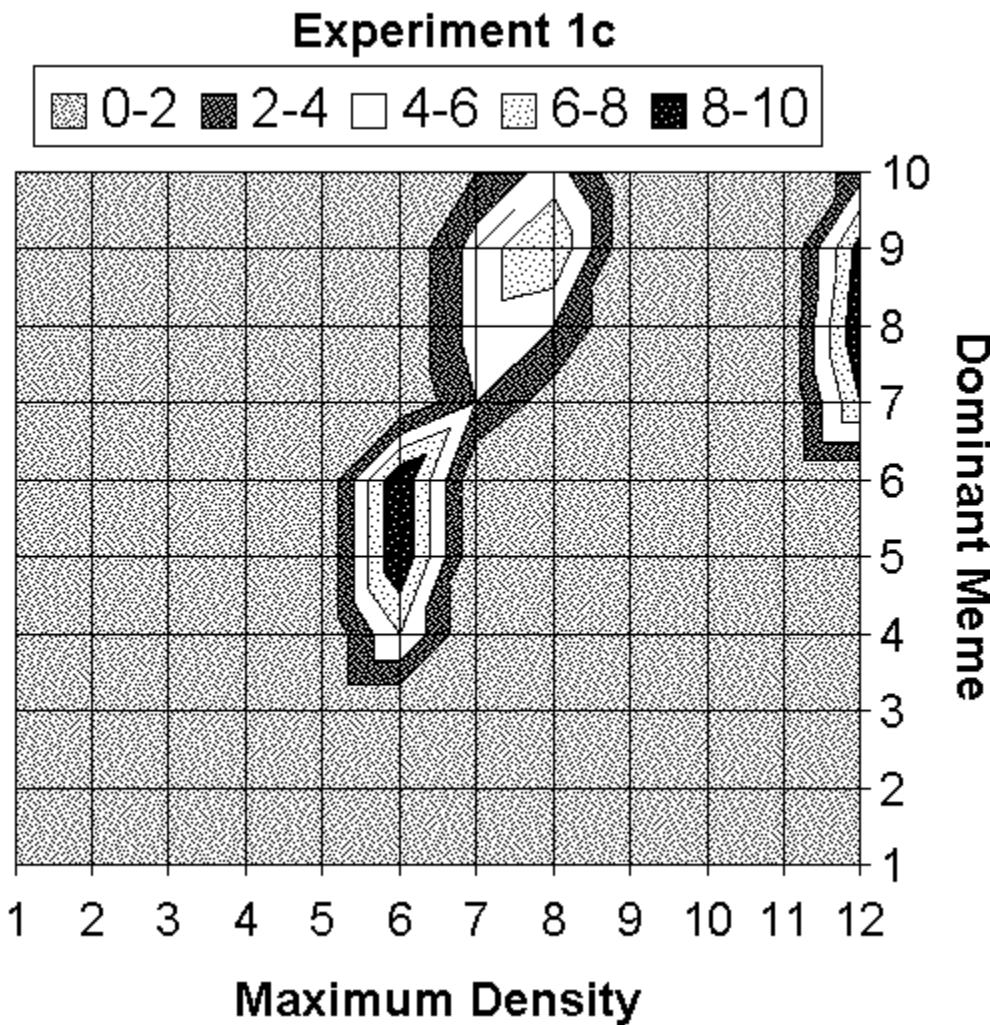


**Figure 3. Dominant meme / Maximum density synthesis**

#### **Experiment 1c - "Too Much Food with Predators"**

6.5

By the 300th cycle 67 per cent of the simulation runs had found an equilibrium. By the 1000th cycle it was 92 per cent. Only 2 per cent of the simulation runs resulted in optimal population distributions (see Figure 4). The attacks of predators increased the effects of the self-catalytic process by forcing grazers into overpopulated territories. They also increased the time taken to attain equilibrium due to the increased death rate.



**Figure 4. Dominant meme / Maximum density synthesis**

6.6

Notice the far right grouping in Figure 4, indicating that in a significant number (32%) of simulation runs, high value memes formed an equilibrium even when all grazers were in the same territory. Consequently the average optimality of the population distribution is low (see the CAE measure given in Table 2).

#### **Observations from the experiments**

6.7

Consideration of the experiments without meta-memes leads to the following observations:

*Many stable noosphere states*

Many distributions of memes produce a stable noosphere. The model therefore, produces many viable "cultures" given the same conditions. One consequence of this is that disbelief is high (in

the sense of the mismatch between actual carrying capacities and the memes which predominate in the noosphere).

### *Optimal distributions in the minority*

Most of the simulation runs produce non-optimal stabilities. This means that the death rate can be high and constant but the noosphere stays stable. This indicates that a viable "culture" is not based on the optimality of the population as a whole. In this sense memes do not need to keep agents alive to prosper.

### *Dynamic equilibria of population distribution*

A stable noosphere does not necessarily indicate a stable population distribution. Oscillations or seemingly random movements are sometimes observed. This is interesting since it suggests that certain stable noosphere compositions accommodate complex dynamical behaviours of populations.

### *Killing memes can prosper*

The "self-catalytic" effect of the production of aggressive "killing memes" is well described by Bura and easy enough to understand. Abstracting the observation from the specifics of the simulation we might say that: any meme that can influence an agent's behaviour in so as to reinforce and spread itself can continue to exist regardless of its side-effects. It may become dominant even if this is dysfunctional to agents individually or as a population. In the context of the model this works by mutual reinforcement. In the context of the specifics of the grazer simulations this involves getting lots of agents into one territory. This experimental evidence throws doubt on Bonner's(1980) statement concerning the possibility of successful "killing memes":

The instinct for survival is important to culture because a meme, in order to be invented or acquired must pass a severe test: If it in any way endangers the lives of the animals concerned, it will automatically be rejected. (Bonner 1980, p197)

## 6.8

Without some perfect evaluation function to "screen-out" killing memes, how can an animal avoid the traps that these agents have fallen into? Could meta-memes help to dampen such a process?

## **The Introduction of Meta-Memes into the Model**

### 7.1

In the grazer simulations a simple unit of behaviour (herding) is represented by a meme. The meme takes the form of different varieties of herding. These memes are simply varieties of the same behaviour. We can say they are part of the same "meme family" (Bura 1994). Of course it is quite possible to have memes which influence different sorts of agent behaviour. In the context of the grazer simulations the agents are simple; they move, feed and communicate memes. In the

simulations so far, movement was determined by the herding memes held by the agents. But the grazers' handling of memes is a behaviour that can itself be mediated by memes. This is what meta-memes are. They are a subset of all possible memes which directly affect an agent's meme handling abilities. In a sense they are *ideas about ideas* [10].

#### 7.2

In human society the "meme" concept is itself a meta-meme. The "scientific method" could be considered to be a high-level meta-meme (its primary function to filter other ideas, theories, beliefs etc.). But statements such as: "Do not believe what agent B believes" can also be viewed as meta-memes.

#### 7.3

Meta-memes require high-level cognitive and communication behaviours such as language. It is hard to imagine how meta-memes could replicate via simple imitation (apart from indirectly through some side-effect of an imitated behaviour [11]). In these experiments meta-memes have been introduced to the model. They do not "emerge" from the model.

#### 7.4

Meta-memes effectively allow a constraint to be turned into a variable and adapted differentially across individual agents. In the previous experiments it was found that the noosphere tends to stabilise with a variety of different memes. It is very rare for the whole noosphere to be dominated by a single meme. The introduction of meta-memes should therefore increase diversity of behaviour, specifically, behaviour related to the way agents handle memes. The population as a whole might therefore be more flexible.

#### 7.5

The potential interaction of memes, meta-memes, agents and environment could be very complex. In order to take a step back from this complexity, analysis is best kept to population level indicators. The introduction of meta-memes could increase or decrease: (a) stability in the noosphere; (b) the death rate [12]; and (c) the optimality of population distribution.

#### Meta-Memes and the Model

#### 7.6

In order to implement meta-memes the memory capacity of the agents is increased so they can hold one standard herding meme and one meta-meme. Standard memes and meta-memes do not compete with each other directly (i.e. a standard meme can not replace or repel a meta-meme). The standard meme is "executed" during the action phase of the system cycle whereas the meta-meme is "executed" during the meme phase. For each agent the meme phase is carried out once for the standard meme and once for the meta-meme. The size of the noosphere is therefore doubled but partitioned. One half of the noosphere is occupied by meta-memes, the other by standard memes.

## The "Open-Mindedness" Meta-Meme

7.7

Intuitively one can envisage certain kinds of meta-memes that might completely "kill" noosphere dynamics. Consider a meta-meme that stops agents attempting to replicate memes or allows them to repel all memes from other agents. Such meta-memes would appear to have the potential to halt noosphere evolution. The self-referential nature of meta-memes seems to indicate that self-catalytic phenomena at a meta-level could be swifter and more disastrous to the population than those already observed.

7.8

In the context of the experiments this intuition is put to the test. An "open-mindedness" meta-meme was selected for the simulations. It has the capacity to increase an agent's probability of repelling or accepting attempted meme infections from other agents.

7.9

The "open-mindedness" meta-meme refers to a family of memes represented by the integers 1 to 10. The higher the meta-meme, the more "open-minded" the grazer possessing the meme becomes. A high value predisposes a grazer to accept memes from other grazers during an attempted infection. A low value predisposes a grazer to repel memes. A grazer with an open-mindedness meme of "6" will behave identically to a grazer without meta-memes.

7.10

In Minimeme each copy of each meme has two values associated with it: aggression and change. The change value [0..1] determines the probability that the meme will succumb to infection and be replaced by a different meme. The open-mindedness meta-meme is added to the change value (see Table 1) for the purposes of deciding if infection takes place. The value of the meta-meme can therefore be seen as a bias which is added to the change values of all memes held by the agent during an attempted infection from another agent<sup>[13]</sup>. In this way the probability of infection or repelling is modified by the meta-meme value. A grazer that posses a low value meta-meme reduces the change value of all its memes<sup>[14]</sup>. Conversely a grazer with a high value meta-meme increases the change value of all its memes<sup>[15]</sup>.

---

**Table 1. Meta-meme values and biases**

---

Meta-meme	1	2	3	4	5	6	7	8	9	10
Bias value	-1.0	-0.8	-0.6	-0.4	-0.2	0.0	+0.2	+0.4	+0.6	+0.8

---

Each meta-meme value is shown along with the bias it produces. This value is added to the change parameter of each meme held by the agent when an attempted infection is taking place. Note that meta-meme value "6" has a zero bias. This means that a grazer possessing a "6" meta-meme will behave in exactly the same way as a grazer with no meta-meme at all.

### 7.11

For a simulation with meta-memes we have two noosphere graphs (one for standard memes and one for meta-memes). The two together represent the entire noosphere. We can analyse them separately because the meta-memes and standard memes do not compete for space in the noosphere<sup>[16]</sup>.

### 7.12

Since the noosphere is now twice the size, we might expect it to take longer to settle down to an equilibrium. We also might expect low value (closed-minded) meta-memes to take over the noosphere because they are self-reinforcing, whereas higher value meta-memes are self-destroying<sup>[17]</sup>.

## Experiments with Meta-Memes

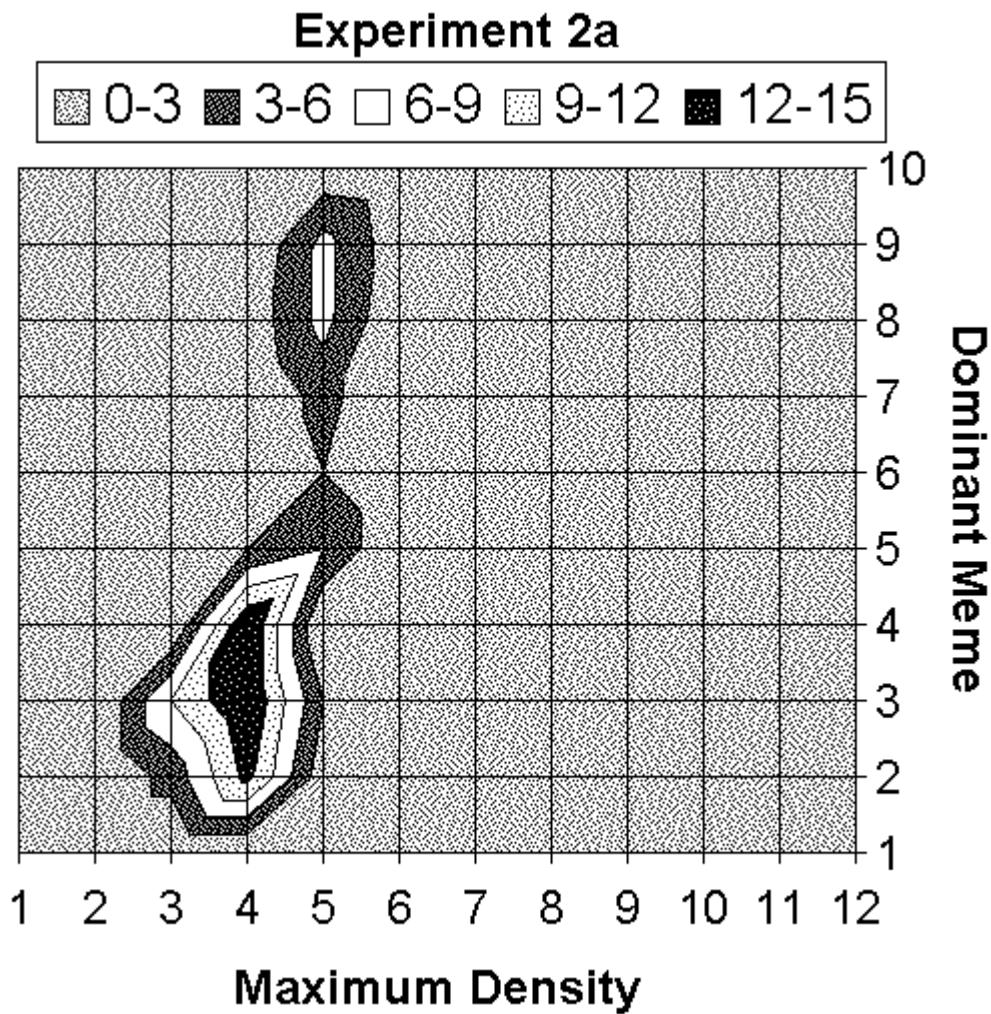
### 8.1

The three previous experiments were performed again but with the open-mindedness meta-meme turned on.

### Experiment 2a - "Just Enough Food" With Meta-Memes

### 8.2

By the 1,000th cycle 76 per cent (same as experiment 1a) of the simulation runs had found an equilibrium. By the 2,000th cycle it was 99 per cent (2 per cent more than experiment 1a -- an insignificant difference given the inherent randomness in the simulations). Most of the runs (87%) did not result in an optimal population distribution at equilibrium (this is only slightly better than experiment 1a). However, the non-optimal equilibria obtained were substantially more optimal (on average) than those obtained for experiment 1a (compare Figures 1 and 5). An optimality measure which takes account of dynamic population distributions can be obtained by calculating the average total cumulative deaths after equilibrium has been attained. For each of the 100 simulation runs, the number of agent deaths is measured during the last 100 cycles (during which the noosphere does not change). This value is then averaged (divided by 100). For experiment 1a this value was 18, and for experiment 2a it was 2 (see Table 2). The introduction of meta-memes has therefore increased average optimality while keeping stability about the same (i.e. time taken to attain noosphere equilibrium).

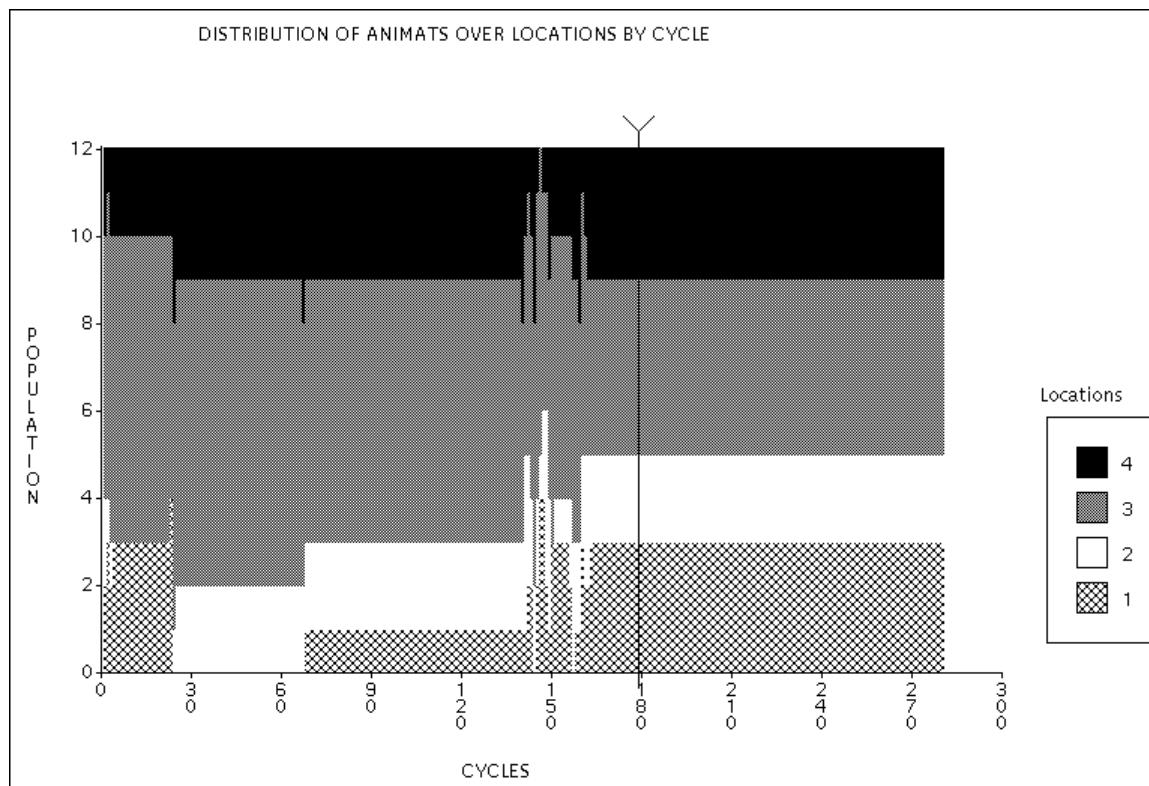


**Figure 5. Dominant meme / Maximum density synthesis**

### 8.3

Figures 6, 7 and 8 show the results of a typical simulation run. At the start of the simulation Figure 6 shows an overpopulated territory (location 3) occupied by 6 grazers. Within a few cycles this increases to 7 grazers. Figure 7 shows the self-catalytic process occurring in this territory (the high meme values). This is as observed for experiment 1a. With the introduction of meta-memes however this process is less stable and breaks-down completely around cycle 150. During the self-catalytic process the death rate is high. Notice the oscillations of high and low meta-meme values (Figure 8) up to cycle 160. These oscillations correspond to the period during which the self-catalytic process is occurring. Between cycles 30 and 50 there is an increase of low value meta-memes. By about cycle 60 however, higher value meta-memes have begun to dominate. After about cycle 100 low value meta-memes make a come back reaching a plateau around cycle 150. This meta-meme oscillation process appears to destabilise the part of the noosphere containing the standard memes. The oscillations appear to correlate with periods of high death rates. High death rates are a natural consequence of overpopulated territories and it is the self-catalytic process that forces territories to become overpopulated. The meta-meme

noosphere is stable between about cycle 140 and cycle 160. In this period movement occurs (Figure 6) and the death rate is low.



**Figure 6. Distribution of the population over the four territories. Experiment 2a - Just enough food.**

Notice the redistribution of the population between cycles 140 and 160.

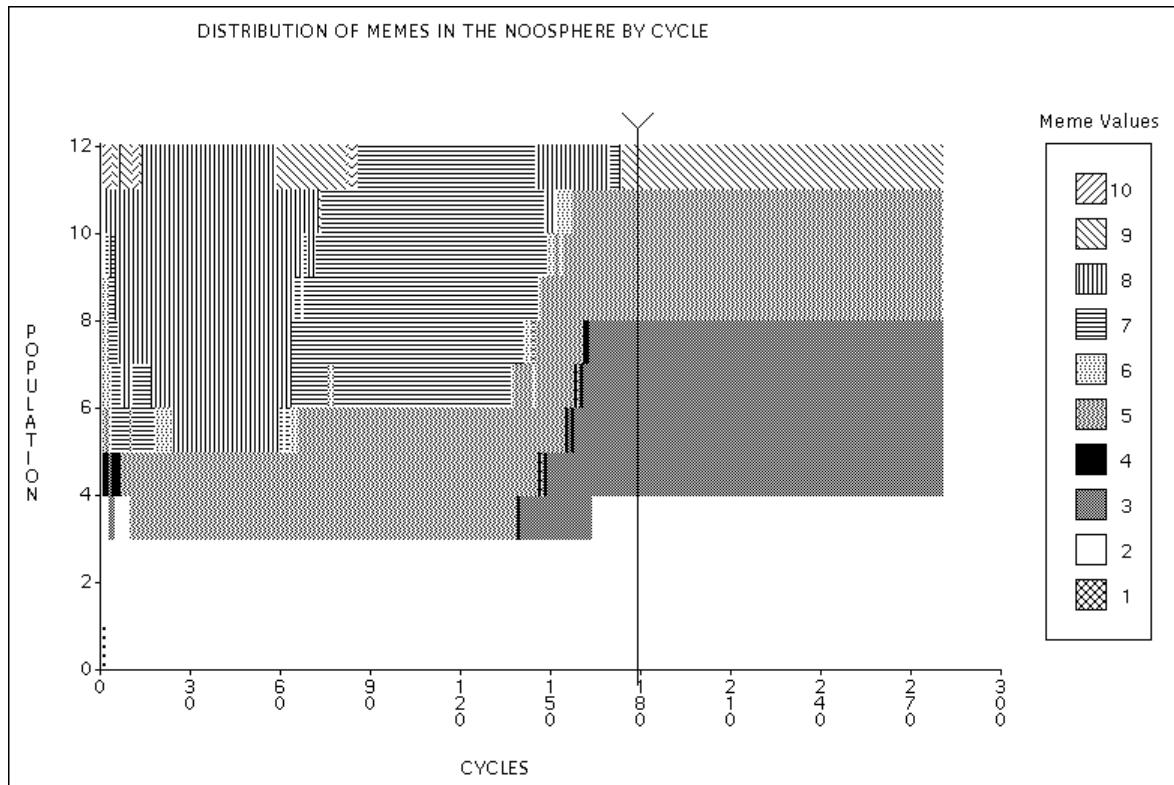


Figure 7. Distribution of memes in the noosphere. Experiment 2a - Just enough food.  
Notice the self-catalytic process and when it breaks down.

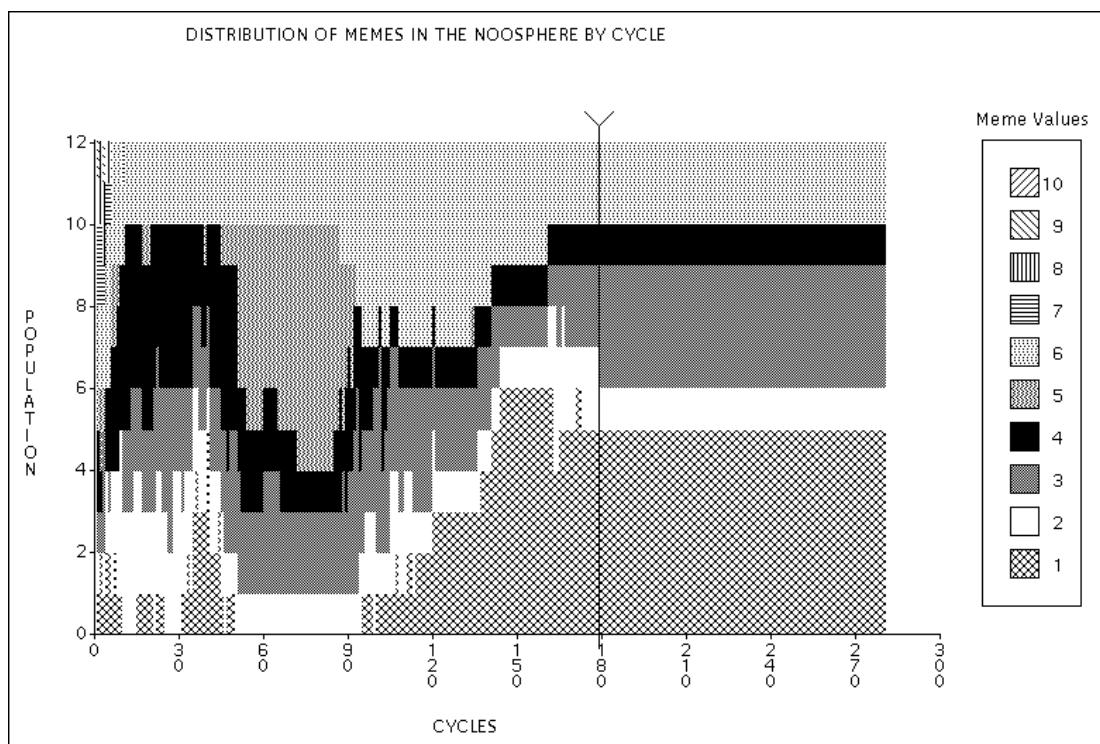
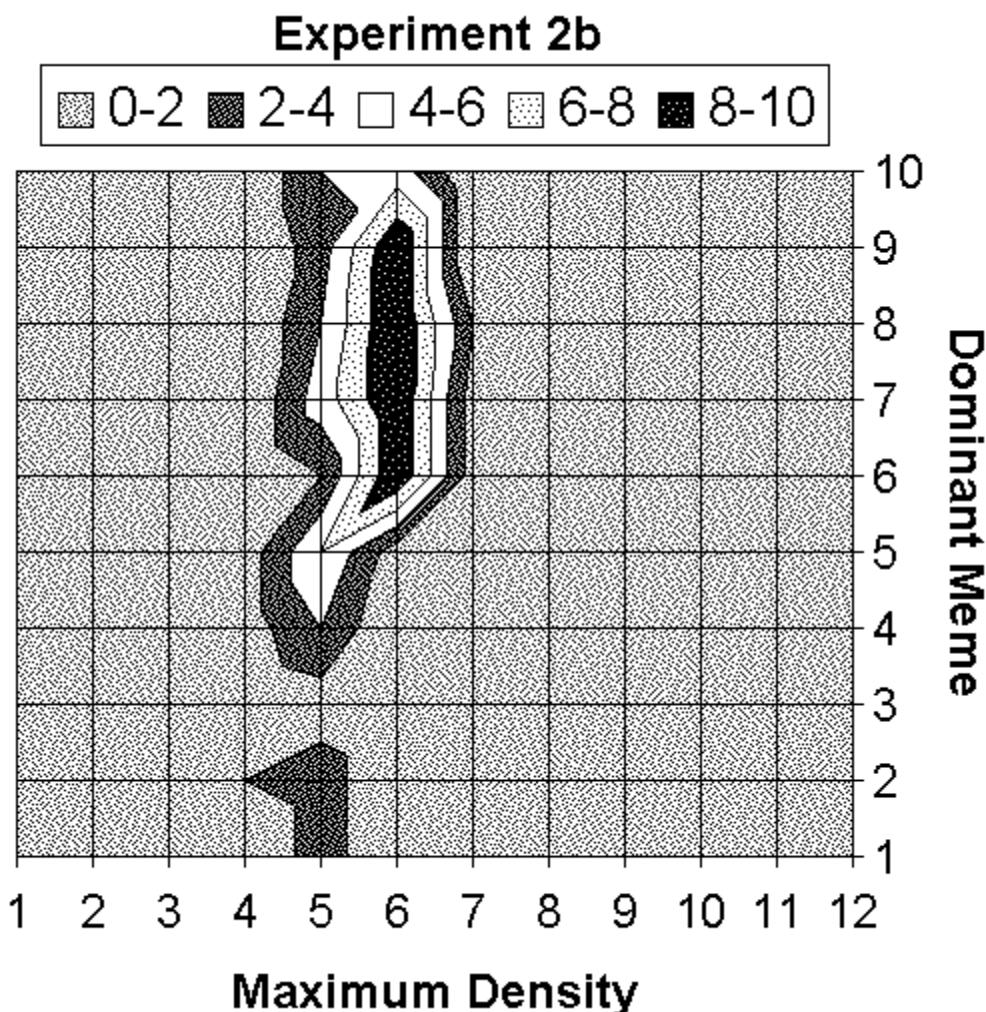


Figure 8. Distribution of memes in the noosphere. Experiment 2a - Just enough food.  
Notice the oscillations and then the minor stability between cycles 140 and 160.

### **Experiment 2b - "Too Much Food" With Meta-Memes**

8.4

By the 300th cycle 62 per cent (30 per cent less than experiment 1b) of the simulation runs had found an equilibrium. By the 1000th cycle it was 97 per cent (3 per cent less than experiment 1b). The equilibria obtained were substantially more optimal than those obtained for experiment 1b (compare Figures 3 and 9). The average total cumulative deaths after equilibrium was 2 compared to 9 for experiment 1b. The introduction of meta-memes has substantially increased the time taken to attain equilibrium but increased the optimality of those equilibria.



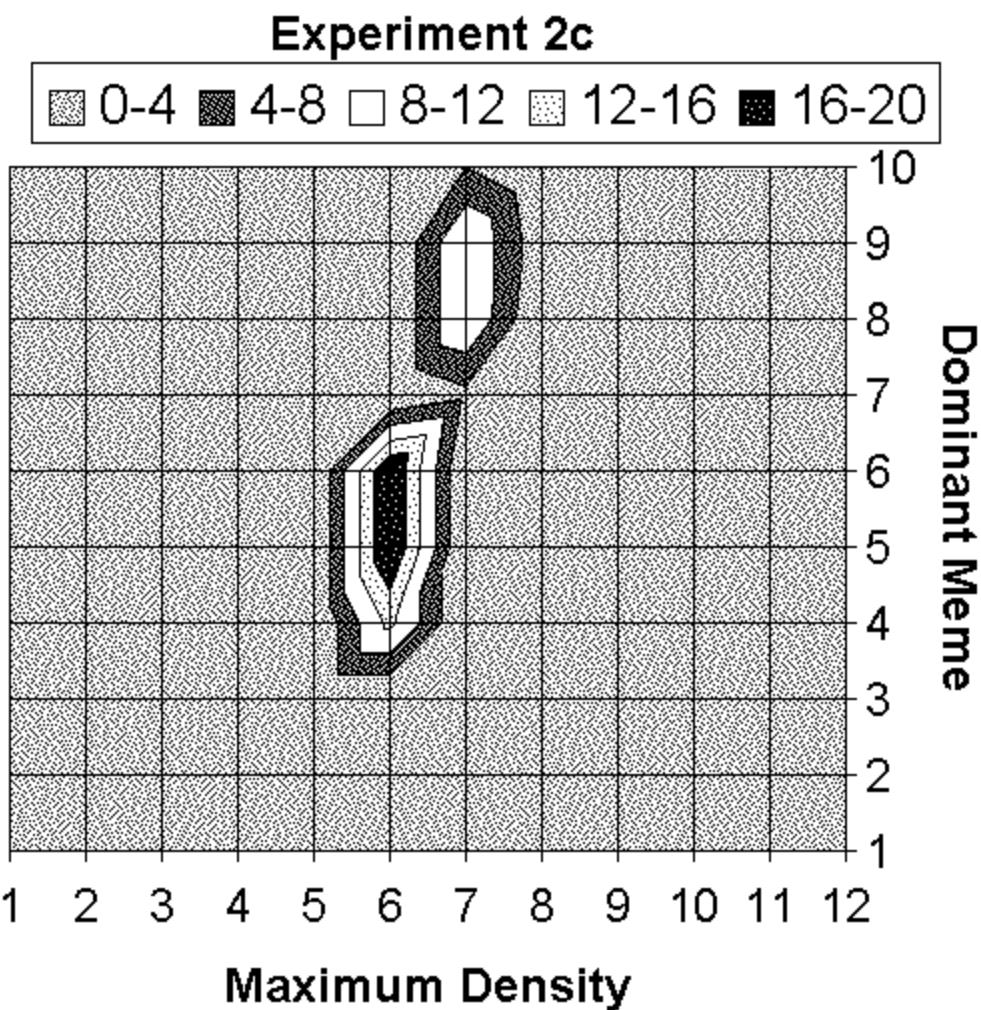
**Figure 9. Dominant meme / Maximum density synthesis**

### **Experiment 2c - "Too Much Food with Predators" With Meta-Memes**

8.5

By the 300th cycle 19 per cent (47 per cent less than experiment 1c) of the simulation runs had found an equilibrium. By 5,000 it was 95 per cent. Only 3 per cent of the simulation runs found a

noosphere equilibrium with an optimal population distribution compared to 2 per cent for experiment 1c. The population distributions were more optimal than experiment 1c (compare Figs. 4 and 10). Notice that the far right grouping (in Figure 4) has been totally removed. This is a result of the meta-memes breaking down the most extreme manifestation of the self-catalytic process (when all grazers occupy a single territory). The mechanisms by which this process occurs are discussed in detail later but note that in such a situation (all grazers in one territory) the death rate will be very high and as such will de-stabilise the meta-meme noosphere.



**Figure 10. Dominant meme / Maximum density synthesis**

#### 8.6

The introduction of meta-memes has substantially increased the time taken to attain equilibrium but increased the optimality of those equilibria significantly.

## Observations and Findings

### 9.1

Table 2 gives a summary of the results of all the experiments. Taking averages across all three experiments, the introduction of the open-mindedness meta-meme resulted in the following (compared to experiments 1a, 1b and 1c) [18]:

- A decrease in the stability of the noosphere (i.e. equilibrium is reached more slowly).
- A slight increase in the number of agent deaths before equilibrium [19].
- An increase in the optimality of the equilibriums found.
- An increase in the "accuracy" of the dominant meme (more closely reflects the actual carrying capacities of the territories; see Figures 11a, 11b and 11c).

---

**Table 2. A summary of results**

---

Simulation Description	0.3	1	3	5	10	CBE	CAE
.....Experiments without meta-memes							
1a) Just enough food	49	76	97	100	100	296	18
1b) Too much food	92	100	100	100	100	22	9
1c) Too much food & predators	67	92	100	100	100	349	41
Averages	69	89	99	100	100	222	23
.....Experiments with meta-memes							
2a) Just enough food	26	76	99	99	100	173	2
2b) Too much food	62	97	100	100	100	32	2
2c) Too much food & predators	19	49	85	95	100	690	7
Averages	33	73	95	98	100	298	4

---

The numbered columns represent cycles (in thousands). The numbers in those columns represent the percentage of simulation runs that had reached an equilibrium by the given number of cycles. The CBE column shows the average Cumulative deaths Before Equilibrium. The CAE column shows the average Cumulative deaths After Equilibrium. After each set of three experiments the average of the columns is given.

### 9.2

Table 2 shows the optimality in terms of average "Cumulative deaths After Equilibrium" in the "CAE" column. The CAE value is the number of agents that died for the 100 cycles at the end of each run (during which the noosphere is stable) averaged over the 100 runs. Notice that experiment 2c (too much food with predators) has a much lower CAE at the expense of a much higher "CBE" (average Cumulative deaths Before Equilibrium). The CBE value is the number of

agents that died up to the point that equilibrium was attained averaged over the 100 runs. Comparison of Figures 4 and 10 show that the removal of the far right grouping, where all grazers stay in one territory, is responsible for the bulk of this reduction in the CAE.

## What's Going On In The Meta-Noosphere?

9.3

An explanation for all the above effects which is consistent with the experimental results can be summarised as: *Two opposing processes create oscillations in the meta-meme noosphere during periods of high death rates. This causes instability and population migrations. The self-catalytic process in which killing memes prosper in overpopulated territories is generally broken after a few hundred cycles.*

9.4

The meta-meme noosphere tends to oscillate <sup>[21]</sup> during periods with a high death rate. This will occur whenever there is a self-catalytic process (overpopulation of a territory with high standard memes dominating it). This oscillation in the meta-meme noosphere has two effects. Firstly, a noosphere stability is prevented. This stops an equilibrium from being achieved at a point which is highly non-optimal. Secondly, the standard meme noosphere is affected when the oscillations become extreme. If the meta-meme noosphere becomes dominated by either extreme of meta-meme (highly open-minded or highly closed minded) the standard meme noosphere becomes unstable and vulnerable to dramatic changes based on mutation. This tends to push the population out of a territory where a self-catalytic process is occurring.

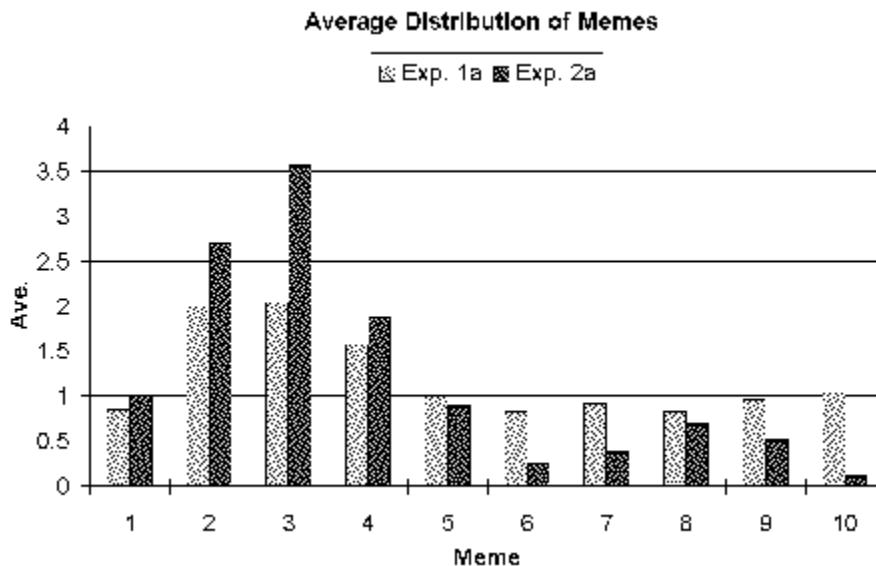
9.5

Oscillations in the meta-meme noosphere are caused by the interaction of two opposing processes. There are two ways in which meta-memes can increase stability in the noosphere: *High value meta-memes predominate producing increased "homogenisation".* If a population in a given territory are strongly open-minded (high meta-meme values) then any new grazer entering the territory has a high probability of being infected with a high value meta-meme thus "converting" a potentially closed-minded grazer into an open-minded one. Such a newly converted open-minded grazer subsequently has a high probability of infection by the dominant standard meme within the territory. A deviant mutated standard meme generated from within the territory is easily suppressed due to the already open-minded nature of the grazer. The grazer tends to get quickly re-infected with the dominant standard meme for the given territory. However, if a closed-minded grazer manages to infect another host in the territory before being infected itself the territory can quickly become closed-minded.

*Low value meta-memes predominate producing a closed-minded population.* If low value meta-memes predominate in a given territory then all attempted infections are strongly resisted. Such a territory is however vulnerable because any new arrival into the territory (or mutation) has a high probability of being infected by a low value meta-meme making a potentially deviant standard meme resistant against infection.

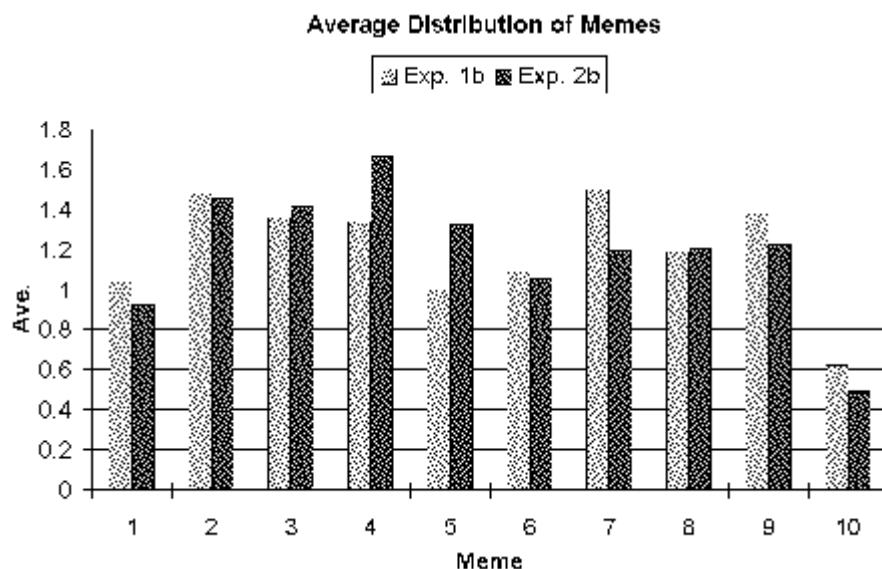
## 9.6

Mutation is high when the death rate of a territory is high. During times of high mutation it is more likely that the vulnerabilities (outlined above) of the two stability producing processes will be exploited. When this occurs, the meta-meme noosphere tends to oscillate between the two. In effect, the open-mindedness meta-meme can stabilise the noosphere in either of the above ways but will tend to oscillate between the two when the death rate is high. This oscillation prevents equilibrium and tends to produce instability in the standard meme noosphere.



**Figure 11a. Distribution of memes averaged over all the simulation runs for scenario A: "Just enough food".**

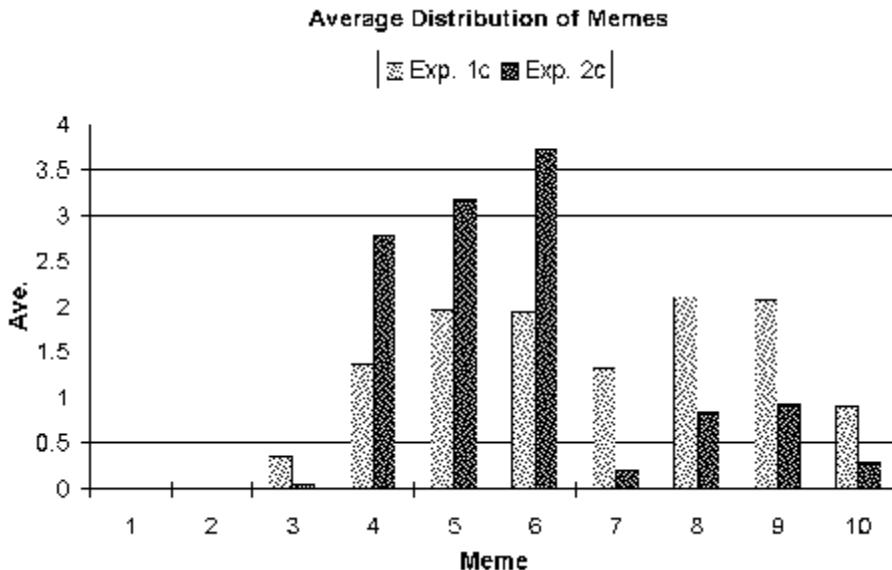
The darker bars show the results of the meta-meme experiments. After the introduction of meta-memes, meme "3" representing the actual carrying capacity is favoured. Consequently the average accuracy of the memes is improved.



**Figure 11b. Distribution of memes averaged over all the simulation runs**

### for scenario B: "Too much food".

The darker bars show the results of the meta-meme experiments. After the introduction of meta-memes, meme "4" representing the actual carrying capacity is favoured. Consequently the average accuracy of the memes is improved.



**Figure 11c. Distribution of memes averaged over all the simulation runs for scenario C: "Too much food with predators".**

The darker bars show the results of the meta-meme experiments. After the introduction of meta-memes, lower value memes are favoured. Consequently the average accuracy of the memes is improved.

### 9.7

The open-mindedness meta-meme will therefore increase the stability of the population when the death rate is low and decrease it when the death rate is high. This extra instability is functional since it produces population instabilities (movement) which generally change the death rate. If this new death rate is low enough then an equilibrium may be achieved otherwise the whole cycle will repeat. The system therefore lurches quite blindly but tends over time to increase stability and lower the death rate.

### 9.8

Experiment 2c (too much food with predators) has a substantially higher CBE and takes longer to find an equilibrium than did 1c. This result is due to the high death rates that such a scenario produces. This results in a constantly oscillating meta-meme noosphere which causes constant population shifts. It is therefore more difficult for the system to achieve an equilibrium. When equilibrium is attained however, it is much more optimal (the CAE drops from 41 to 7).

### 9.9

It was previously speculated that closed-minded (i.e. low value) meta-memes would prosper since they are self-selecting because any host that carries them is less likely to change memes through infection. Indeed the meta-meme noosphere quickly reduces to low values. However results show the "1" meta-meme is no more successful than the "2", "3" or "4" meta-memes.

## Conclusions

### 10.1

Starting from a set of assumptions which specified a co-ordination problem based on resource harvesting and agents which individually sacrifice, a memetic process was investigated. The process involved the replication, reinforcement and repelling of memes held by agents. The memes held by an agent determined its behaviour. They were exchanged via social contact between agents located within the same territory.

### 10.2

The original results found by Bura (1994) were reproduced indicating that memes which lead to rapid agent death can prosper by causing agents to participate in a "self-catalytic" process in which they gathered in a single territory exhausting its resources and reinforcing the very meme which brought them there. This result is at odds with Bonner's (1980) intuition.

### 10.3

With the addition of an open-mindedness / closed-mindedness meta-meme to the same scenario the "self-catalytic" process was destabilised resulting in less agent deaths. More optimal population distributions were found. Also the intuition that complete closed-mindedness would predominate is shown to be incorrect. These results are determined by the complex interplay of several feedback processes involving agents, resources and memes.

### 10.4

Much human social phenomena results from consciously planned and co-ordinated interactions. Conversely, there is much which, it is claimed, is not. It is this latter class of phenomena (which challenge rational action theory) that can benefit from a memetic treatment. The "self-catalytic" process demonstrates that human and artificial societies which consist of agents who practice "boundedly rational" and "socially docile" strategies suggested by Simon (1990) can become trapped into highly non-optimal behaviours without any knowledge of what is occurring. Also the results here demonstrate that (at least some) such traps can be avoided without necessarily requiring agents to have a high level of social knowledge and planned co-ordination.

### 10.5

One major result that emerged was the vast diversity of meme distributions that produced optimal stabilities and hence the easy coexistence of different views (memes) of reality. The co-evolution of several "incorrect" memes can produce fairly optimal behaviour patterns: each meme tied together through mutual correction based on incorrect views of reality. This finding can be compared to the conception of functional disbelief (Doran 1994, Doran 1998).

### 10.6

Why produce computational models of meme spread? In many areas of anthropology and sociology theories are constructed and described using natural language. However, the complexity of the phenomena to be explained (namely social reality) poses two fundamental problems. Firstly, natural language is often ambiguous and theories may "gloss over" important areas. Secondly, it may be impossible to test or falsify such theories because it is very difficult to collect "clean" data from real societies. This is of particular relevance to the memetics

community since the hypothesised objects of interest (memes) are difficult to track or measure. If a computational artificial society is constructed then this forces the explicit statement of assumptions in the form of a computer program. The society can then be studied (conducting experiments and collecting clean data) which hopefully can feedback or at least inform theory construction for the real social world. Experimentation in the artificial realm is neither purely inductive nor deductive but involves both (Axelrod 1997). I use the term "ceduction" (Computer Experimental Induction/Deduction) to describe this mode of investigation (Hales 1998a, Hales 1998d). This kind of speculative ceductive mode of enquiry using artificial societies is a growing area and has been embraced by sociologists, philosophers, psychologists, economists, political scientists and computer scientists in order to test and build new social theory (Conte et al, Conte & Gilbert 1995, Doran & Gilbert 1994). However, existing disciplines possess their own traditions resulting in cultural inertia and consequently slow up-take of new methodological tools. Memetics as a young and (for the most part) computationally aware discipline can utilise these factors to its benefit if it embraces this new methodology as a testing ground for new theory construction and testing. On-going work using memetic simulations to investigate altruism, group formation and stereotyping are detailed in Hales (1998b) and Hales (1998c).

### 10.7

Memetic models do not automatically provide any new understanding. What they do provide is a fresh perspective on an old debate and a flexible experimental and objective method to test assumptions and intuitions. Of course, the assumptions we start with are based on our own ideas of social reality. It is notoriously difficult to get hard facts or principles concerning real social systems. The social sciences are dominated by many different views of social reality. This should come as no surprise. As the meme simulations presented in this paper suggest, social reality is constructed from competing ideas which do more than just coexist, they have often co-evolved to produce global actions which are functional even though the memes (beliefs) that predominate may be objectively incorrect.

## Acknowledgements

This work would not have been possible without the help, discussions and suggestions of Prof. Jim Doran, Department of Computer Science, University of Essex, Colchester, UK, and Stephane Bura, 4 Place Jussieu 75252 Paris cedex 05, France. Thanks also go to the JASSS reviewers and editors for comments and corrections and Jan Neil for lucid questioning which helped to keep me (at least partially) within the real (as opposed to the artificial) world. This work was supported by an EPSRC research studentship award. A significant proportion of this paper is based on an earlier paper (Hales 1997).

## Notes

<sup>1</sup>

In the context of Sugarscape this involves counting the number of "1" digits in the string.

<sup>2</sup>

This is an apt word to use. It was first introduced in an essay written in 1949 by Teilhard de Chardin. It is used to signify the realm in which mind is exercised. Teilhard's contention was that

in the ordinary course of the evolution of living things the biosphere is being supplanted by the noosphere (Bullock 1988).

3

A grazer can move into any territory from any other.

4

If more than one territory is equally desirable then one is chosen at random.

5

Within the scenario of the grazer simulations, "birth control" is applied. In such a scheme the population is kept at a fixed size (12 in this case). A new grazer is "born" every time an old grazer dies. The new born grazer is placed within the same territory. This maintains a fixed size noosphere for the purposes of analysis.

6

The order in which agents take turns in the Action, Feeding and Meme phases of the system cycle has been randomised. This is a modification of the original Minimeme model as presented by Bura. This change was made to avoid any artefacts that might result from synchronous turn-taking (Hagselmann 1996).

7

A stable noosphere is defined as one in which no change has occurred in the last 100 cycles. Experiments have shown that after such a state is reached the noosphere generally resists attacks from mutant memes and does not evolve anymore.

8

It is important to note that such a map does not show the full picture. It deals only with the dominant meme and the territory with the maximum density. It does however, give a general picture of the major trends across all simulation runs in a digestible form.

9

This is the term used by Bura to describe the following: If a grazer holds a high value meme it will move to an overpopulated territory. By definition, the grazers in this territory will also hold high value memes. When a large enough group is formed mutual reinforcement of this meme will continue although deaths may be high if the carrying capacity of the territory is exceeded.

10

The words "idea", "belief" and "meme" are used interchangeably throughout this paper.

11

For example: A preference for a certain location could isolate an agent from exposure to memes from agents who have a preference for a different location. But this is a side-effect of the meme. A true meta-meme operates directly on the meme process.

12

There are two death rates: the initial death rate which is the number of deaths up to the point at which equilibrium is attained (these will usually be sporadic and unpredictable); and the generally constant death rate after equilibrium. Each death rate is considered separately.

13

If the result of this addition is greater than 1 or less than 0 then the result is taken to be 1 or 0 respectively.

14

This includes the open-mindedness meta-meme itself.

15

This modification process is only performed when deciding if infection should occur in the meme phase of the system cycle. The open-mindedness meta-meme has no effect on the change value during other phases (i.e. mutation). The bias values given in Table 1 are the result of applying the following formula to the meta-meme value (m):

$$\text{bias} = (m - 6) * 0.2$$

This formula was used because it delivers uniform increments, a neutral bias (meta-meme "6" = zero bias) and a fully closed minded meta-meme (meta-meme "1" = -1 bias).

16

This is a simulation specific constraint, since a grazer can not hold two standard memes or two meta-memes because they are "concurrent". Other scenarios might not have such a constraint.

17

Any low value meta-meme will make the grazer less likely to accept an attempted replication over any of its memes and this includes the meta-meme itself. A low value meta-meme (once taken by an agent) will therefore be very resistant to being replaced by replication.

18

These observations are also consistent with the results of additional experiments using a scenario based on a 2-2-2-2 carrying capacity (not enough food). Interestingly, the CAE values for these experiments were roughly equivalent to experiments 1b and 2b. The harsh environment seemed to improve the accuracy of the standard memes which compensated for the lower level of resources available.

19

This is almost entirely due to the results of experiment 2c (too much food with predators). Indeed experiment 2a (just enough food) resulted in a substantial reduction of the CBE figure.

20

All of the experiments were reproduced using a different pseudo random number generation algorithm to avoid artefacts. All the conclusions drawn are consistent with both sets of results.

For these additional results see Hales (1995). For a discussion of random number generators see Mathews (1995) and Press et al. (1992).

21

Oscillations tend to take place within the lower half of the meta-meme range ("1" to "5"). In this sense the system does select closed-minded grazer behaviour. Interestingly it does not favour the "1" meta-meme above all others.

## References

- AXELROD, R. (1995). The Convergence and Stability of Cultures: Local Convergence and Global Polarization. Working Paper 95-03-028. Santa Fe, N.M.: Santa Fe Institute.
- AXELROD, R. (1997). Advancing the Art of Simulation in the Social Sciences. In Conte, R., Hegselmann, R. & Terna P. (Eds.) *Simulating Social Phenomena*. Berlin: Springer.
- BONNER, J.T. (1980). *The Evolution of Culture in Animals*. New Jersey: Princeton University Press.
- BOYD, R. & Richerson P. (1985). *Culture and the Evolutionary Process*. Chicago: University of Chicago Press.
- BULLOCK, A. (ed.) (1988). *Fontana Dictionary of Modern Thought*. London: Fontana Press.
- BURA, S. (1994). Minimeme: Of Life and Death in the Noosphere. In J. Mayer & S. Wilson (eds.) *From Animals to Animats 3: Proceedings of the 3rd International Conference on the Simulation of Adaptive Behaviour* (SAB94). London: MIT Press.
- CONTE, R. & Gilbert, N. (1995). Introduction: Computer Simulation for Social Theory. In Gilbert, N. & Conte, R. (Eds.) *Artificial Societies: the Computer Simulation of Social Life*. London: UCL Press.
- CONTE, R. et al (1997). Social Simulation - A New Disciplinary Synthesis. In Conte, R., Hegselmann, R. & Terna P. (Eds.) *Simulating Social Phenomena*. Berlin: Springer.
- DAWKINS, R. (1976). *The Selfish Gene*. New York: Oxford University Press.
- DAWKINS, R. (1982). *The Extended Phenotype*. Oxford: Freeman. pp97 - 117.
- DAWKINS, R. (1993). Viruses of the Mind. In B. Dahlbom. (ed.) *Dennett and His Critics*. Cambridge: Blackwell.
- DENNETT, D. (1995). *Darwin's Dangerous Idea*. New York: Simon & Schuster.

DORAN, J. (1994), Modelling Collective Belief and Misbelief. In M. Keane, P. Cunningham, M. Brady, & R. Byrne, (eds.) *Proceedings of the Seventh Annual Conference on AI and Cognitive Science* (AICS94). Dublin: Dublin University Press.

DORAN, J. (1998), Simulating Collective Belief and Misbelief. In the *Journal of Artificial Societies and Social Simulation* vol. 1, no.1, <http://www.soc.surrey.ac.uk/JASSS/1/1/3.html>

DORAN, J. & Gilbert, N. (1994), Simulating Societies: an Introduction In Gilbert, N. & Doran, J. (Eds.) *Simulating Societies: The Computer Simulation of Social Phenomena*. London: UCL Press.

EPSTEIN, J.M. and Axtell, R. (1996). *Growing Artificial Societies: Social Science From The Bottom Up*. London: MIT Press.

GABORA, L.M. (1995) Meme and Variation: A Computational Model of Cultural Evolution. In Nadel, L. & Stein, D. (Eds.) *1993 Lectures in Complex Systems*. New York: Addison-Wessley.

GABORA, L.M. (1997) The Origin and Evolution of Culture and Creativity. In *Journal of Memetics - Evolutionary Models of Information Transmission* vol.1, no.1, [http://www.cpm.mmu.ac.uk/jom-emit/1997/vol1/gabora\\_1.html](http://www.cpm.mmu.ac.uk/jom-emit/1997/vol1/gabora_1.html)

HALES, D. (1995). *Modelling Meta-Memes*. Unpublished MSc Dissertation. Department Of Computer Science, University Of Essex, UK.

HALES, D. (1997). Modelling Meta-Memes. In Conte, R., Hegselmann, R. & Terna P. (Eds.) *Simulating Social Phenomena*. Berlin: Springer.

HALES, D. (1998a). Artificial Societies, Theory Building and "Ceduction". Unpublished paper, Presented at the CRESS workshop "The Potential of Computer Simulation in the Social Sciences", January 1988, University of Surrey, UK.

HALES, D. (1998b). Selfish Memes and Selfless Agents: Altruism in the Swap Shop. In *The Proceedings of ICMAS 1998* (ICMAS'98). California: IEEE Computer Society.

HALES, D. (1998c). Stereotyping, Groups and Cultural Evolution. In *Proceedings of the Multi-Agent Based Simulation Workshop 1998* (MABS'98) To be published by Springer in the LNAI series late 1998.

HALES, D. (1998d). Artificial Societies, Theory Building and Memetics. In *Proceedings of the 15th International Congress on Cybernetics*. Namur: Int. Association for Cybernetics.

HEGSELMANN, R. (1996). Cellular Automata in the Social Sciences: Perspectives, Restrictions and Artefacts. In R. Hegselmann, U. Mueller & K.G. Troitzsch (eds.) *Modelling and simulation in the social sciences from a philosophy of science point of view* [Theory and Decision Library]. Dordrecht: Kluver.

MATHEWS, R. (1995). It's a Lottery. *New Scientist* 1987, 38-42.

PRESS, H., Flannery, B., Teukolsky, S., Vetterling, W. (1992). *Numerical Recipes in C : The Art Of Scientific Computing*. Cambridge: Cambridge University Press.

REYNOLDS, R. (1994). Learning to Co-operate using Cultural Algorithms. In Gilbert, N. & Doran, J. (Eds.) *Simulating Societies: The Computer Simulation of Social Phenomena*. London: UCL Press.

SIMON, H. (1990). A Mechanism for Social Selection and Successful Altruism. *Science* 250: 1665-1668

**BELIEF HAS UTILITY – AN INTENTIONAL STANCE**  
**A COMMENTARY ON GATHER’S PAPER: “WHY THE *THOUGHT CONTAGION* METAPHOR IS RETARDING THE PROGRESS OF MEMETICS”**

By  
**David Hales**  
Department of Computer Science  
University of Essex  
daphal@essex.ac.uk

Gatherer (1998) makes a clear distinction between two definitions of the meme. As such it is useful in clarifying potential terminological confusion between what might be called the "internalist" or "cognitive" memeticists (using the Dawkins-B meme definition) and the "externalists" or "behavioural" memeticists (using the Dawkins-A meme definition).

The behaviourists want to operationalise memes in observable and measurable behaviours, acts and artefacts etc. Certainly anything that can't be observed or measured can not be a basis for direct empirical investigation. This, I think, is incontestable.

Gatherer attacks unconstrained theorising about "memes" that become vaguely and imprecisely equated with unobservable and unproved "neural entities". Memes then become so under-specified as to be meaningless, representing anything from sensory qualia to belief in the existence of god. Promises of a future golden age, when the neural structure of such "memes" will be revealed much like the genetic code are, as Gatherer rightly points out, no excuse for such imprecise use of terminology.

I do not agree however that from acceptance of this position we should necessarily infer that memes be seen as necessarily and always equivalent with observable and measurable phenomena. For sure, in empirical studies where memes are to be identified and tracked an operational definition based on observable and measurable phenomena must be made. Importantly however, I contend that this does not necessarily and always require us to suppose that the unit of cultural reproduction under investigation is equivalent to these observable or measurable phenomena.

It is often necessary for us to take an "intentional stance" (Dennett 1989) when observing human agents in order to make sense of their behaviour. I do not understand how we can talk about human society and individuals at a meaningful level of abstraction without reference to the "beliefs" that individuals hold. I can not make sense of why I am writing this without utilising some kind of intentional stance towards the memetics community who will read it. Presumably Gatherer believes that the Dawkins-A definition puts memetics on a more sensible footing and presumably I do not. From this abstraction a lot of things make sense. The conception of a "belief" is a powerful abstraction. Why should it be rejected a priori? Assuming that individuals have beliefs does not imply that individuals share similar neural structures and the "intentional stance" does not imply an ontological commitment to the concept of belief as such. In its weakest form we take such a "stance" to be no more than an instrumentalist device which gives the user

enhanced predictive power or understanding of some particular observable or measurable phenomena. This "stance" currently gives you some handle on why you and others are bothering to read and write all this text about memes - or does it?

Given that at some level of abstraction and from an intentional stance we can observe behaviours consistent with the assumption that individuals can communicate and spread beliefs, I fail to see why memetic theory can not be meaningfully and soundly pitched at the intentional level. The prize is to generate theory that can enhance our understanding of observable social processes and phenomena. The challenge facing memetics is that of developing such theory and empirically verifying it. Both is lacking at present. One approach to theory construction is to use computational models (see Doran 1998; Hales 1997, 1998a, 1998b, 1998c, 1998d).

In the panel discussion at the end of the first symposium on memetics at the 1<sup>st</sup> International Conference of Cybernetics in Namur (Hales 1998e) my impression was of general agreement that precise definitions were less important than examples of good memetic work. Participants also expressed that as an emerging and highly inter-disciplinary area we should be tolerant of different approaches and terminology.

Gatherer concludes his paper, arguing that memeticists should adopt the Dawkins-A definition "for the sake of intellectual survival". It is still early days and we should be more optimistic. However, in the long run if memetics cannot deliver new theory for the solution of problems which cannot be solved in other paradigms then we should not mourn but welcome its demise and move on to better frameworks.

## References

- Dennett, D. (1989). *The Intentional Stance*. Cambridge MA: MIT Press.
- Doran, J. (1998), Simulating Collective Belief and Misbelief. *Journal of Artificial Societies and Social Simulation*, 1(1). <http://www.soc.surrey.ac.uk/JASSS/1/1/3.html>
- Gatherer, D. (1998) Why the "Thought Contagion" Metaphor is Retarding the Progress of memetics. *Journal of Memetics - Evolutionary Models of Information Transmission*, 2. [http://jom-emit.cfpm.org/1998/vol2/gatherer\\_d.html](http://jom-emit.cfpm.org/1998/vol2/gatherer_d.html)
- Hales, D. (1997). Modelling Meta-Memes. In Conte, R., Hegselmann, R. & Terna P. (Eds.). *Simulating Social Phenomena*. Berlin: Springer-Verlag.
- Hales, D. (1998a), Selfish Memes and Selfless Agents - Altruism in the Swap Shop. In the Proceedings of the 3rd International Conference on Multi-Agent Systems (ICMAS'98). IEEE Computer Society, California.
- Hales, D. (1998b), Stereotyping, Groups and Cultural Evolution. In Sichman, J., Conte, R., & Gilbert, N. (Eds.) *Multi-Agent Systems and Agent-Based Simulation. Lecture Notes in Artificial Intelligence*, 1534.

Hales, D. (1998c), Artificial Societies, Theory Building and Memetics. Proceedings of the 15th International Congress on Cybernetics. Namur: International Association for Cybernetics.

Hales, D. (1998d), An Open-Mind is Not an Empty Mind. Journal of Artificial Societies and Social Simulation, 1(4). <http://www.soc.surrey.ac.uk/JASSS/1/4/2.html>

Hales, D. (1998e). Report on the Panel Discussion that occurred at the Symposium on Memetics. Journal of Memetics - Evolutionary Models of Information Transmission, 2. [http://jom-emit.cfpm.org/1998/vol2/panel\\_discussion.html](http://jom-emit.cfpm.org/1998/vol2/panel_discussion.html)

# MEMETICS

By  
**F. Heylighen,**  
Aug 18, 1994  
Principia Cybernetica

- **Meme:** an information pattern, held in an individual's memory, which is capable of being copied to another individual's memory.
- **Memetics:** the theoretical and empirical science that studies the replication, spread and evolution of memes

Cultural evolution, including the evolution of knowledge, can be modelled through the same basic principles of variation and selection that underly biological evolution. This implies a shift from genes as units of biological information to a new type of units of cultural information: *memes*.

A meme is a cognitive or behavioral pattern that can be transmitted from one individual to another one. Since the individual who transmitted the meme will continue to carry it, the transmission can be interpreted as a *replication*: a copy of the meme is made in the memory of another individual, making him or her into a *carrier* of the meme. This process of self-reproduction (the memetic life-cycle), leading to spreading over a growing group of individuals, defines the meme as a replicator, similar in that respect to the gene (Dawkins, 1976; Moritz, 1991).

Dawkins listed the following three characteristics for any successful replicator:

**Copying-fidelity:** The more faithful the copy, the more will remain of the initial pattern after several rounds of copying. If a painting is reproduced by making photocopies from photocopies, the underlying pattern will quickly become unrecognizable.

**Fecundity:** The faster the rate of copying, the more the replicator will spread. An industrial printing press can churn out many more copies of a text than an office copying machine.

**Longevity:** The longer any instance of the replicating pattern survives, the more copies can be made of it. A drawing made by etching lines in the sand is likely to be erased before anybody could have photographed or otherwise reproduced it.

## Memes versus genes

In these general characteristics, memes are similar to genes and to other replicators, such as computer viruses or crystals. The genetic metaphor for cultural transmission is limited, though. Genes can only be transmitted from parent to child ("vertical transmission"). Memes can be transmitted between any two individuals ("horizontal transmission" or "multiple parenting"). In that sense they are more similar to parasites or infections (cf. Cullen, 1998).

For genes to be transmitted, you need a generation. Memes only take minutes to replicate, and thus have potentially much higher fecundity (see Competition between Memes and Genes). On the other hand, the copying-fidelity of memes is in general much lower. If a story is spread by being told from person to person, the final version will be very different from the original one. It is this variability or fuzziness that perhaps distinguishes cultural patterns most strikingly from DNA structures: every individual's version of an idea or belief will be in some respect different from the others'. That makes it difficult to analyse or delimit memes. This does not imply that meme evolution cannot be accurately modeled, though. After all, genetics was a well-established science long before the precise DNA structure of genes was discovered.

Examples of memes in the animal world are most bird songs, and certain techniques for hunting or using tools that are passed from parents or the social group to the youngsters (Bonner, 1980). In human society, almost any cultural entity can be seen as a meme: religions, language, fashions, songs, techniques, scientific theories and concepts, conventions, traditions, etc. The defining characteristic of memes as informational patterns is that they can be replicated in unlimited amounts by communication between individuals, independently of any replication at the level of the genes.

## Modelling memes

Memetics can be defined as an approach trying to model the evolution of memes. Memes undergo processes of variation (mutation, recombination) of their internal structure. Different variants will compete for the limited memory space available in different individuals. The most fit variants will win this competition, and spread most extensively. This spreading can in principle be modelled mathematically (see e.g. Boyd & Richerson, 1985; Cavalli-Sforza & Feldman, 1981; Lumsden & Wilson, 1981; Csanyi, 1991; Lynch, 1998), although in practice it will be very difficult to determine the exact values of the parameters of the model. A more practical, qualitative approach is to formulate specific criteria for the fitness of a meme, relative to other memes, taking into account the subsequent stages of the memetic life-cycle.

As is the case with genes, it is not necessary to know the exact coding or even the exact size or boundaries of a meme in order to discuss its fitness, and thus to make predictions about its further spreading, survival or extinction within the population of competing memes. Such predictions can be empirically tested. For example, a memetic hypothesis might state that simpler memes will spread more quickly. This can be tested by observing the spread (perhaps in a controlled environment) of two memes that are similar in all respects, except that the one is simpler. Theories can also be induced from empirical observation of meme behavior "in the wild" (see e.g. Best, 1998). Given the differences in variation and selection mechanisms, it is also possible to make predictions about the competition between memes and genes.

Variation, replication and selection on the basis of meme fitness determine a complex dynamics. This dynamics will be influenced by the medium through which memes are communicated, and the copying-fidelity, fecundity and longevity it allows. Perhaps the most powerful medium for meme transmission is the computer network, and this implies some specific characteristics for memes on the net.

## References:

- Aunger R. (ed.) (2000): Darwinizing Culture: The Status of Memetics as a Science, Oxford University Press.
- Best, M., L., 1997; Models for Interacting Populations of Memes: Competition and Niche Behavior. *Journal of Memetics - Evolutionary Models of Information Transmission*, 1.
- Blackmore S. (2000): The Meme Machine, (Oxford Univ. Press).
- Bonner J.T. (1980): The Evolution of Culture in Animals, (Princeton University Press, Princeton).
- Boyd R. & Richerson P.J. (1985): Culture and the Evolutionary Process, (Chicago University Press, Chicago).
- Cavalli-Sforza L.L. & Feldman M.W. (1981): Cultural Transmission and Evolution: a quantitative approach, (Princeton University Press, Princeton).
- Csanyi V. (1991): Evolutionary Systems and Society: a general theory, (Duke University Press, Durham, NC).
- Cullen B. (1998): "Parasite Ecology and the Evolution of Religion", in: Heylighen F., Bollen J. & Riegler A. (eds.) (1999): The Evolution of Complexity (Kluwer Academic, Dordrecht).
- Dawkins R. (1976): The Selfish Gene, (Oxford University Press, New York).
- Heylighen F. (1992): "Selfish Memes and the Evolution of Cooperation", *Journal of Ideas*, Vol. 2, #4, pp 77-84.
- Heylighen F. (1998): "What makes a meme successful? Selection criteria for cultural evolution", in: Proc. 16th Int. Congress on Cybernetics (Association Internat. de Cybernetique, Namur), p. 423-418.
- Lynch, Aaron (1997) "Units, Events, and Dynamics in Memetic Evolution", *Journal of Memetics*
- Lumsden, Charles, and Wilson, Edward (1981): Genes, Mind, and Culture: the Coevolutionary Process, (Harvard University Press, Cambridge).
- Moritz E. (1990): "Memetic Science: I - General Introduction", *Journal of Ideas* 1, p. 1-23
- Moritz E. (1995): Metasystems, Memes and Cybernetic Immortality, in: Heylighen F., Joslyn C. & Turchin V. (eds.), *The Quantum of Evolution. Toward a theory of metasystem transitions*, (Gordon and Breach Science Publishers, New York) (special issue of "World Futures: the journal of general evolution, vol. 45, p. 155-171).

# **COMPETITION BETWEEN MEMES AND GENES**

By  
**F. Heylighen**  
Aug 18, 1994  
1994 Principia Cybernetica

## **Different mechanisms of evolution**

Though memetic and genetic evolution are subjected to the same basic principles of blind variation and natural selection on the basis of fitness, memetic evolution is basically a much more flexible mechanism. Genes can only be transmitted from parents (or parent in the case of asexual reproduction) to offspring. Memes can in principle be transmitted between any two individuals (though it will become more difficult the larger the differences in cognitive mechanisms and language are). (This is sometimes called "multiple parenting").

For genes to be transmitted, you typically need one generation, which for higher organism means several years. Memes can be transmitted in the space of hours. Meme spreading is also much faster than gene spreading, because gene replication is restricted by the rather small number of offspring a single parent can have, whereas the amount of individuals that can take over a meme from a single individual is almost unlimited. Moreover, it seems much easier for memes to undergo variation, since the information in the nervous system is more plastic than that in the DNA, and since individuals can come into contact with much more different sources of novel memes. On the other hand, selection processes can be more efficient because of "vicarious" selection (Campbell, 1974): the meme carrier himself does not need to be killed in order to eliminate an inadequate meme; it can suffice that he witnesses or hears about the troubles of another individual due to that same meme.

The conclusion is that memetic evolution will be several orders of magnitude faster and more efficient than genetic evolution. It should not surprise us then that during the last ten thousand years humans have almost not changed on the genetic level, whereas their culture (i.e. the total set of memes) has undergone the most radical developments. In practice the superior "evolvability" of memes would also mean that in cases where genetic and memetic replicators are in competition, we would expect the memes to win in the long term, even though the genes would start with the advantage of a well-established, stable structure. This explains why sociobiological models of human behavior can only be partially correct, as they neglect memetic factors.

## **Different selection criteria**

When memetic and genetic fitness criteria are inconsistent, the different implicit objectives of memes and genes will lead to a direct competition for control of the carrier's behavior. Both replicators have similar aims to the degree that they use the same vehicles: individual organisms. Everything that strengthens the vehicles should in general be good for the replicators, and hence both genes and memes should be selected on the basis of their support for increased survivability and reproducability of their carriers. However, the implicit goals of genes and memes are

different to the degree that they use different mechanisms for spreading from one vehicle to another one. Memes will be positively selected mainly for increased *communicability*. Genes will be selected mainly for *sexual reproducability*. These different emphases may lead to direct conflicts.

For example, priests in many religions are prohibited to marry and to have children, in striking disagreement with genetic injunctions. Yet we can easily imagine that the religious meme of celibacy would have been selected because unmarried priests can spend more time and energy on "spreading the word", and hence replicating the meme.

An even more vivid example of countergenic behavior, closely related to the issue of altruism, is that of martyrs, suicide teams, or kamikaze pilots, who are willing to give up their life in order to promote the spread of a meme: a religion, an ideology or a nation (i.e. a group defined by a common culture or ideology). In that case, the loss of one or a few carriers is compensated by the increased chances of survival for the other carriers or for the meme itself. For example, the suicide of an individual may attract the attention of other individuals to the meme he is carrying, and thus facilitate its spreading. A well-known example is Jan Palach, the Czech student who put himself to fire in order to protest the Soviet suppression of the "Prague Spring". In this case the meme would be the Czech version of "socialism with a human face".

Reference: Heylighen F. (1992): "Selfish Memes and the Evolution of Cooperation", Journal of Ideas, Vol. 2, #4, pp 77-84.

# **STRUCTURE OF MEMES**

By  
**F. Heylighen,**  
1994 Principia Cybernetica

## **Modelling meme units**

The main criticism that can be raised against the memetic approach is that memes are difficult to define. What are the elements or units that make up a meme? Does a meme correspond to a complete symphony, or to a symphonic movement, a melody, a musical phrase, or even a single note?

In order to model meme structure, we may use some concepts from cognitive science. Perhaps the most popular unit used to represent knowledge in artificial intelligence is the production rule. It has the form "if condition, then action". In symbols:

**If A, then B or  $A \rightarrow B$**

A represents a condition that is distinguished, B represents an action that is executed or another condition that is activated. The action leads in general to the activation of another condition. In fact a production rule can be analysed as a combination of even more primitive elements: two distinctions (which discriminate between presence and absence of the condition and the action respectively) and a connection (the "then" part, which makes the first distinction entail the second one) (Heylighen, 1991d; see also Heylighen, 1990). For example, a meme like "God is omnipotent" can be modelled as "if a phenomenon is God (distinction of God from non-God), then that phenomenon is omnipotent".

Production rules are connected when the output condition (action) of the one matches the input condition of the other. E.g.  $A \rightarrow B$ ,  $B \rightarrow C$ . This makes it possible to construct complex cognitive systems on the basis of elementary rules. Even remembered melodies might be modelled in such a way, as concatenations of production rules of the type "if C (musical note distinguished), then E (note produced and subsequently distinguished)", "if E, then A", and so on.

A similar model applies to genes. A gene corresponds to a string of DNA codons, which respond to the presence of certain activating proteins or the absence of certain inhibiting proteins (condition) by manufacturing new proteins (action). This may in turn activate further genes, depending on the present of specific chemicals in the cell, and so on. This leads to complex networks of "if... then" productions (Kauffmann, 1992).

## **Variation of memetic units**

It has been shown that production rules (or at least a simplified, binary representation of them, called "classifiers") can be used to build quite impressive computer simulations of cognitive

evolution, using mutations, recombinations, and selection on the basis of "fitness" (Holland et al., 1986).

Distinctions can be represented as combinations (strings) of elementary *yes-no* (1-0) observables. Mutation or recombination of distinctions can then be modelled by either randomly changing certain binary digits in a string, or by concatenating the first part of one string (A) with the second part of another string (B), like in the following example:

A = 1001|001 mutation: A' = 1001|000

B = 0010|011 recombination ("crossing-over") of A and B:

A~B = 1000|011

Although these models do not as yet take into account distinct carriers, this looks like a very promising road to study memes formally and computationally.

## Meme complexes

Even if we would model memes as connected sets of production rules, we still have the problem of how many production rules define a single meme. If we call a religion or a scientific theory a meme, it is clear that this will encompass a very large number of interconnected rules. In practice it will be impossible to enumerate all rules, or to define sharp boundaries between the rules that belong to the meme and those that do not. However, that should not detract us from using memetic mechanisms in analysing evolution.

Indeed, Darwinian models of genetic evolution have certainly proven their usefulness, even though it is in practice impossible to specify the exact DNA codons that determine the gene for, say, blue eyes or altruism towards siblings. As Dawkins (1976) notes, it is not necessary to be explicit about what are the constitutive elements of a gene, postulated to explain a particular characteristic or type of behavior. It is sufficient that we can distinguish the phenotypical effects of that gene from the effects of its rival genes (alleles). If we can determine the fitness resulting from these effects, taking into account the environment and the context of different, non-rival genes present in the genome, then we can make predictions about evolution.

The same applies to memes. If, for example, we observe that one meme (say Catholicism) induces its carriers to have more children than its competitors (say Calvinism and Anglicanism), and that the children tend to take over their memes from their parents, then, all other things being equal, we can predict that after sufficient time that meme will dominate in the population. Of course, in practice it is never the case that all other things are equal, but that is the predicament of all scientific modelling: we must always simplify, and ignore potentially important influences. The question is to do that as wisely as possible, and to maximally include relevant variables without making the model too complex.

## References

Dawkins R. (1976): The Selfish Gene, (Oxford University Press, New York).

Heylighen F. (1992): "Selfish Memes and the Evolution of Cooperation", Journal of Ideas , Vol. 2, #4, pp 77-84.

Holland J.H., Holyoak K.J., Nisbett R.E. & Thagard P.R. (1986): Induction: processes of inference, learning and discovery, (MIT Press, Massachusetts).

Kauffman S.A. (1992): Origins of Order: self-organization and selection in evolution, (Oxford University Press, Oxford).

# **MEMETIC SELECTION CRITERIA**

By  
**F. Heylighen**  
1994 Principia Cybernetica

During the different stages of their life-cycle, memes are subjected to objective, subjective intersubjective and meme-centered selection criteria

These are the criteria that determine the overall fitness of a meme, whether it will maintain within an individual's memory and spread to other individuals, or be eliminated. As meme spreading depends on different objective, subjective and intersubjective mechanisms, the criteria are sometimes contradictory. See the general selection criteria for knowledge for more details on individual criteria.

## **Objective Criteria**

Objective criteria denote selection by phenomena or objects independent of the hosts and memes involved in the process. The distinctiveness criterion functions mainly during the assimilation stage. It states that phenomena that are distinct, detailed or contrasted are more likely to be noticed and understood, and therefore assimilated. The invariance and controllability criteria, on the other hand, apply mainly to the retention stage. According to the invariance criterion, phenomena that recur, independently of the way in which they are perceived, are more likely to be maintained in memory. Controllability notes that phenomena which react differentially to the subject's actions are also more likely to leave a permanent memory trace.

## **Subjective Criteria**

Subjective criteria represent selection by the subject who assimilates the meme. The main criteria at the assimilation stage are novelty (facilitates assimilation by attracting the subject's attention) and simplicity (requires less processing for the meme to be understood). The criterion of coherence (connection, consistency and support between new perception and existing memory trace) facilitates the understanding and acceptance parts of the assimilation stage, since it represents the ease with which the new meme can "fit in" with the memory that is already there. It also facilitates the retention stage since memories that cohere are more easy to retrieve and use and are therefore less likely to be forgotten. The criterion of utility, like controllability, functions mainly at the retention stage, since useful memes are more likely to be effectively used and thus reinforced, although it will also help assimilation, by making it more worthwhile for the host to do the effort to assimilate.

## **Intersubjective Criteria**

Intersubjective criteria represent selection through the interactions between different subjects. Group utility is an emergent criterion that is implicit in all four stages: a memes that is useful to the group of all its hosts is more likely to survive because it helps the group itself to survive and

grow, and thus to absorb other individuals. Authority functions mainly at the assimilation stage: memes from authoritative sources, i.e. hosts or vehicles that are held in high regard or considered to represent expertise in the domain, will be more easily noticed and accepted. Formality (i.e. precise, unambiguous expression) too helps assimilation, at least of the original memetic content of the expression. It will contribute basically to what Dawkins (1976) calls copying-fidelity. (On the other hand, informal expression, because it tends to be simpler, may facilitate assimilation, but of an idea different from the one initially expressed). Conformity, the reinforcement of the same meme by different hosts belonging to the same group, will boost acceptance and retention (cf. Boyd & Richerson, 1985). Expressivity, the ease with which the meme can be expressed in an intersubjective medium, will obviously contribute to the expression stage. Publicity, finally, the effort put by the host(s) into the broad distribution of the message, will maximize transmission.

## Meme-centered Criteria

Finally, the meme-centered criteria represent selection on the level of the meme itself. They depend only on the internal structure of the meme, not on its "fit" to external selectors, such as subjects, objects, or groups. These criteria will typically select for "selfish" (cf. Heylighen, 1992) or "parasitic" (cf. Cullen, 1998) memes, whose only goal is to spread themselves, "infecting" a maximum of hosts without regard for their hosts' well-being. This does not imply that the same meme cannot satisfy both selfish and non-selfish criteria. Religions often have this mixture of parasitic and beneficial traits (cf. Cullen, 1998)

Self-justification, the degree to which the components of a meme mutually support each other, will facilitate understanding and acceptance. Self-reinforcement, the degree to which the meme stimulates its host to rehearse itself, e.g. by repetition, meditation, prayer, etc., will strengthen retention. Intolerance, the degree to which a meme excludes rival memes from being assimilated or retained, will also help the meme to retain a stable position in memory. Proselytism, the degree to which the meme urges its host to maximally spread the meme to other hosts, will increase the rates of expression and transmission.

<b>Stage</b>	<b>Objective</b>	<b>Subjective</b>	<b>Inter-subjective</b>	<b>Meme-centered</b>
<b>Assimilation</b>	distinctiveness	novelty simplicity coherence	authority formality	self-justification
<b>Retention</b>	invariance controllability	coherence utility	conformity	self-reinforcement intolerance
<b>Expression</b>			expressivity	proselytism
<b>Transmission</b>			publicity	proselytism

**Table: a summary of the main selection criteria for memes, classified according to the stage during which they are most active, and the system responsible for the selection.**

## **References:**

- Boyd R. & Richerson P.J. (1985): Culture and the Evolutionary Process, (Chicago University Press, Chicago).
- Cullen B. (1998): "Parasite Ecology and the Evolution of Religion", in: Heylighen F. (ed.) (1998): The Evolution of Complexity (Kluwer Academic, Dordrecht)
- Dawkins R. (1976): The Selfish Gene, (Oxford University Press, New York).
- Heylighen F. (1992): "Selfish Memes and the Evolution of Cooperation", Journal of Ideas , Vol. 2, #4, pp 77-84.
- Heylighen F. (1998): "What makes a meme successful? Selection criteria for cultural evolution", in: Proc. 16th Int. Congress on Cybernetics (Association Internat. de Cybernetique, Namur), p. 423-418.
- Heylighen F. (1997): "Objective, subjective and intersubjective selectors of knowledge", Evolution and Cognition 3:1, p. 63-67.

# **EVOLUTION OF MEMES ON THE NETWORK: FROM CHAIN LETTERS TO THE GLOBAL BRAIN**

**By**

**Francis Heylighen**

Center "Leo Apostel", Free University of Brussels,  
Pleinlaan 2, B-1050 Brussels, Belgium  
[fheylich@vnet3.vub.ac.be](mailto:fheylich@vnet3.vub.ac.be)

To be published in: *Ars Electronica Catalogue*, Ingrid Fischer (ed.) (Springer, Vienna/New York, 1996)

## **Memes as Replicators**

An essential characteristic of genes, the units of biological information, is that they *replicate*: they produce copies of themselves, and thereby spread and increase in numbers. Sometimes mutations or copying errors are introduced, producing different variants. Only the best or "fittest" will manage to spread widely. This is the process of *natural selection*, which weeds out inadequate genes. Variation and selection together produce evolution, the perpetual creation of new, better adapted genes.

Richard Dawkins (1976) suggested that a similar mechanism applies to cultural information. Ideas, habits and traditions are communicated from individual to individual. This can be interpreted as a replication: a copy (possibly with errors) of the information is made in the memory of a second individual. Dawkins called the units of this cultural replication process "memes". Practically all cultural entities are memes: images, books, poems, theories, religions, language, melodies, rumours, etc. It suffices that the underlying informational or behavioral pattern is copied, e.g. when people imitate other people's habits or styles, when they learn other's ideas, or reproduce works of art. Like genes, cultural variants can be more or less successful in spreading through the population. They therefore undergo natural selection, and thus evolution.

Dawkins listed the following three characteristics for any successful replicator: 1) *copying-fidelity*: the more faithful the copy, the more will remain of the initial pattern after several rounds of copying. If a painting is reproduced by making photocopies from photocopies, the underlying pattern will quickly become unrecognizable. 2) *Fecundity*: the faster the rate of copying, the more the replicator will spread. An industrial printing press can churn out many more copies of a text than an office copying machine. 3) *Longevity*: the longer any instance of the replicating pattern survives, the more copies can be made of it. A drawing made by etching lines in the sand is likely to be erased before anybody could have photographed or otherwise reproduced it.

In these general respects, memes are similar to genes and to other replicators, like computer viruses or crystals. The genetic metaphor for cultural transmission is limited, though. Genes can only be transmitted from parent to child. Memes can be transmitted between any two individuals ("multiple parenting"). For genes to be transmitted, you need a generation. Memes only take minutes to replicate. On the other hand, the copying-fidelity of memes is in general much lower. If a story is spread by being told from person to person, the final version will be very different

from the original one. It is this variability or fuzziness that perhaps distinguishes cultural patterns most strikingly from DNA structures: every individual's version of an idea or belief will be in some respect different from the others'. That makes it difficult to define or delimit memes.

There are several selection criteria which determine in how far a particular meme will be successful. The more of these criteria a meme satisfies, the more likely it is that it will maintain and spread (Heylighen, 1993). Objective criteria determine whether the knowledge conveyed by a meme can reliably predict events in the outside world. Subjective criteria determine in how far an individual is willing to assimilate a particular meme. They include: 1) *coherence*: the meme is internally consistent, and does not contradict other beliefs the individual already has; 2) *novelty*: the meme adds something new, something remarkable, that attracts the person's attention; 3) *simplicity*: it is easy to grasp and to remember; 4) *individual utility*: the meme helps the individual to further his or her personal goals.

Intersubjective criteria determine how easily memes travel from subject to subject. They include: 5) *salience*: the meme is easily noticed by others, e.g. because it is shouted out loud, or printed on big posters; 6) *expressivity*: the meme is easily expressed in language or other codes of communication; 7) *formality*: the interpretation of the meme's expression depends little on person or context; 8) *infectiveness*: the individuals who carry the meme are inclined to "spread the word", to teach it to other people or to convert them to the belief; 9) *conformism*: the meme is supported by what the majority believe; 10) *collective utility*: the meme is useful for the group, without necessarily being useful for an individual (e.g. the traffic code).

The information in the genes of an organism constitutes its genotype. Its body, developed through the interaction of genotype with environment, constitutes its phenotype. The equivalent of a phenotype for memes is the *sociotype*, the concrete organization of the group of people carrying a collection of memes or *memotype*. For example, if the memotype is the whole of Mormon beliefs, then the sociotype is the group of all Mormons. As a meme is more fuzzy than a gene, a sociotype is likewise more fuzzy than a phenotype: it is in general not clear how a group of meme carriers can be delimited (Heylighen & Campbell, 1995).

## Memes on the Net

The above review sketches the properties of memes in general, be they scientific theories, religions, musical styles or the use of gadgets. It is obvious, though, that the media by which a meme is communicated, such as scientific journals, church preachings, or radio stations, will greatly influence its eventual spread. The most important medium at present is the emerging global computer network, which can transmit any type of information to practically any place on the planet, in a negligible time.

This highly increased efficiency of transmission directly affects the dynamics of replication. Meme transmission over the network has a much higher copying-fidelity than communication through image, sound or word. Digitalisation allows the transfer of information without loss, unlike the analog mechanisms of photocopying, filming or tape recording. Fecundity too is greatly increased, since computers can produce thousands of copies of a message in very little time. Longevity, finally, becomes potentially larger, since information can be stored indefinitely

on disks or in archives. Together, these three properties ensure that memes can replicate much more efficiently via the networks. This makes the corresponding memotypes and sociotypes potentially less fuzzy.

In addition, the network transcends geographical and cultural boundaries. This means that a new development does not need to diffuse gradually from a center outwards, as, e.g., fashions or rumours do. Such diffusion can easily be stopped by different kinds of physical or linguistic barriers. On the net, an idea can appear virtually simultaneously in different parts of the world, and spread independently of the distance or proximity between senders and receivers.

The simplest example of a meme that takes advantage of these network features is a chain-letter: a message sent to different people with the express request to copy it and distribute it further. This is motivated by anticipated rewards for those who do (and punishment for those who don't). Paper chain-letters are often poorly readable photocopies, or manuscripts retranscribed numerous times by hand or by typewriter, with the insertion of plenty of spelling and semantic errors. The effort and cost of copying and distribution moreover limit the number of copies per generation to about 20. Chain-letters distributed by electronic mail, on the other hand, can be sent to hundreds or thousands of people at once, at virtually no efforts or costs, and without information degradation.

Though I have received more chain-letters by email than by post, chain-letters on the net are still a minor phenomenon. Although their spread is very much facilitated by the net, the same applies to all other types of messages. That means that there is increased competition between all these different memes for a limited resource: the attention a user pays to the information he or she receives. Because chain-letters fulfil relatively few of the criteria that distinguish successful memes from unsuccessful ones, they are unlikely to win this competition.

The recent development from the net as carrier of email messages to the World-Wide Web as repository of interconnected documents has greatly changed the dynamics of meme replication.

On the Web, information is no longer distributed by sending copies of files to different recipients. The information is rather stored in one particular location, the "server", where everyone can consult it. "Consultation" means that a temporary copy of the file is downloaded to the RAM memory of the user's computer, so that it can be viewed on the screen. That copy is erased as soon the user moves on to other documents. There is no need to store a permanent copy since the original will always be available. That does not mean that replicator dynamics no longer apply: the interested user will normally create a "bookmark" or "link", i.e. a pointer with the address of the original file, so that it can be easily retrieved later. A link functions as a *virtual copy* (also called an "alias" file), which produces real, but temporary, copies the moment it is activated.

The success of a web document can then be measured by the number of virtual copies or links pointing to it: the documents with most pointers will be used most extensively. There are already web robots, i.e. programs which automatically scan the Web, that make "hit parades" of the documents which are linked to most often. For example, it is likely that a reproduction of the

works of Van Gogh on the Web will be much more popular in number of pointers than the work of some unknown 20th century painter.

## Cooperating Memes: towards a Global Brain?

Discussions about memes, and about evolutionary systems in general, usually emphasize *competition*, i.e. the "survival of the fittest" at the expense of the less fit. However, in biological evolution, *cooperation* between evolving systems, in the sense of symbiosis and mutual support, is at least as important. Multicellular organisms, in which the individual cells cooperate for the collective good, are a prime example.

For genes, competition is limited to "alleles": alternative versions of a gene which compete for the same position on an organism's chromosome. Genes residing in different parts of the chromosome, on the other hand, do no compete but cooperate in steering the development of the organism. Each gene produces a particular type of protein, in reaction to the presence or absence of other proteins in the cell. Together, these proteins build up the cell, digest food molecules, eliminate poisonous molecules, and generally restore the cell equilibrium after different perturbations. Genes cooperate directly through their arrangement in networks: the product of one gene can activate or deactivate a number of other genes, which in turn may activate further genes, and so on.

Similarly, memes can be said to cooperate if they are coherent or support each other. For example, the belief that the Earth is round and the belief that the Earth circles around the sun are mutually reinforcing. The "roundness" meme makes it easier to visualize the "circling" meme, and vice versa. On the other hand, the roundness meme contradicts the meme which says that the Earth is flat. Roundness and flatness behave like alleles, which compete for the same position in a person's view of the world. Though flatness does not directly contradict circling, it is clear that they fit less well together than roundness and circling. Similar examples can be found in the domains of art and fashion. Debussy's music seems to fit in much better with impressionism in painting than with expressionism. Heavy metal music seems to go together with riding motor bikes, but not so much with bicycles.

Mutually supporting memes will tend to group together in larger cooperating ensembles, like ideologies, theories or religions. Mutually exclusive ensembles, like Catholicism and Protestantism, the Copernican and Ptolemaic views of the solar system, or the hippie and punk movements, will compete for converts. It is here that the "multiple parenting" issue, distinguishing memes and genes, comes into play. Since many different people ("parents") can try to convert the same individual, that individual will need to make a choice between the different memes presented to her. (In contrast, you cannot choose from which parents you inherit your genes). For two otherwise equivalent memes, the determining characteristic will be the number of already existing converts: the more people try to convince you of something, the more likely it is that you will follow the lead. If slightly more people believe in one meme, that meme will make more new converts, and thus increase its lead over the competition. This is a self-reinforcing evolution, where success breeds success. As confirmed by a mathematical model of meme transmission (see Heylighen & Campbell, 1995), the result will be that everyone in a group ends up believing the same things. This is what I called earlier "conformist" selection.

However, different groups, with little communication between them, will generally believe different things, since conformist transmission tends to amplify small differences in initial distribution of beliefs.

Let us now see how these mechanisms of cooperation and competition are affected by the global network. Like genes, memes on the web are arranged in networks, where one document points to a number of supporting documents, which in turn link to further supporting documents. Linked documents cooperate, in the sense that they support, confirm or extend each other's ideas. Competing documents, such as announcements of commercial competitors, will not link to each other, or only refer to each other with a phrase like "you should certainly not believe what is said there".

Assuming that two competing documents are equally convincing otherwise, the competition will be won or lost by the number of links that point to each of them. The more pointers to a document can be found, the more people will consult it, and the more further pointers will be made. This is the same kind of self-reinforcing process that leads to conformism, to all members of a group settling on the same meme ensemble. The difference is that now there are no separate groups: on the global network, everyone can communicate with everyone, and every document can link to every other document. The end result is likely to be the emergence of a globally shared ideology, or "world culture", transcending the old geographical, political and religious boundaries. (Note that such homogenization of memes only results for memes that are otherwise equivalent, such as conventions, standards or codes. Beliefs differing on the other dimensions of meme selection will be much less influenced by conformist selection.) Such a networked ideology would play a role similar to that of the genome, the network of interconnected genes that stores the blueprint, and controls the physiology, of a multicellular organism. The corresponding "organism" or sociotype for this meme network would be the whole of humanity, together with its supporting technology. Individual humans would play a role similar to the organism's cells, which in principle have access to the whole of the genome, but which in practice only use that part of it necessary to fulfil their specific function.

There is a better metaphor for the emerging global network. Rather than comparing it to an organism's genome, which is normally static and evolves only because of random copying errors, it can be likened to the organism's brain, which learns and develops in a non-random way. The network functions like a nervous system for the social *superorganism* (Stock, 1993), transmitting signals between its different "organs", memorizing its experiences, making them available for retrieval when needed, and generally steering and coordinating its different functions. Thus, it might be viewed as a *global brain* (Russell, 1995).

The learning of new associations can be implemented by automating the evolutionary process which creates new links. We have set up an experiment in which a hypertext network adapts its links to the pattern of its usage by "learning" the implicit semantics of its users. Such a learning web can be made more intelligent by implementing the equivalent of "thoughts": software agents, which search the net by spreading out while following associative links, collecting information that answers the user's questions (Heylighen & Bollen, 1996). Such brain-like networks may seem far removed from our initial meme model, but they are still based on the same dynamics of variation and selection of (real or virtual) copies of information.

## References

- Dawkins R. (1976): *The Selfish Gene* (Oxford University Press, New York).
- Heylighen F. (1993): "Selection Criteria for the Evolution of Knowledge", *Proc. 13th Int. Cong. on Cybernetics* (Int. Ass. of Cybernetics, Namur), p. 524-528.
- Heylighen F. & Bollen J. (1996): "The World-Wide Web as a Super-Brain", in: *Cybernetics and System '96* (Austrian Society for Cybernetic Studies), p. 917-922.
- Heylighen F. & Campbell D.T. (1995): "Selection of Organization at the Social Level", *World Futures*: 45, p. 181-212.
- Russell P. (1995): *The Global Brain Awakens: Our Next Evolutionary Leap* (Miles River Press).
- Stock G. (1993): *Metaman: the merging of humans and machines into a global superorganism* (Simon & Schuster, New York).

# WHAT MAKES A MEME SUCCESSFUL? SELECTION CRITERIA FOR CULTURAL EVOLUTION

By

**Francis Heylighen**

CLEA, Free University of Brussels, Pleinlaan 2, B-1050 Brussels, Belgium

E-mail: fheylich@vub.ac.be,

To be published in: Proc. 15th Int. Congress on Cybernetics (Association Internat. de Cybernétique, Namur).

**ABSTRACT.** Meme replication is described as a 4-stage process, consisting of assimilation, retention, expression and transmission. The effect of different objective, subjective, intersubjective and meme-centered selection criteria on these different stages is discussed.

## Introduction

Cultural evolution, including the evolution of knowledge, can be modelled through the same basic principles of variation and selection that underlie biological evolution (Boyd & Richerson, 1985; Cavalli-Sforza & Feldman, 1981). This implies a shift from genes as (replicating) units of biological information to a new type of (replicating) units of cultural information: *memes* (Dawkins, 1976). A meme can be defined as an information pattern, held in an individual's memory, which is capable of being copied to another individual's memory. This includes anything that can be learned or remembered: ideas, knowledge, habits, beliefs, skills, images, etc. Memetics can then be defined as the theoretical and empirical science that studies the replication, spread and evolution of memes (Moritz, 1990).

To be replicated, a meme must pass successfully through four subsequent stages: 1) assimilation by an individual, who thereby becomes a *host* of the meme; 2) retention in that individual's memory; 3) expression by the individual in language, behavior or another form that can be perceived by others; 4) transmission of the thus created message or *meme vehicle* to one or more other individuals. This last stage is followed again by stage 1, thus closing the replication loop. At each stage there is selection, meaning that some memes will be eliminated. The present paper will look in more detail at the mechanisms governing these four stages, and present a list of selection criteria that allow us to estimate the fitness of a meme relative to its competitors.

## The four stages of meme replication

### Assimilation

A successful meme must be able to "infect" a new host, that is, enter into its memory. Let us assume that a meme is presented to a potential new host. "Presented" means either that the individual encounters a meme vehicle, or that he or she independently discovers it, by observation of outside phenomena or by thought, i.e. recombination of existing cognitive elements. To be assimilated, the presented meme must be respectively *noticed*, *understood* and *accepted* by the host. Noticing requires that the meme vehicle be sufficiently salient to attract the host's attention. Understanding means that the host recognizes the meme as something that can

be represented in his or her cognitive system. The mind is not a blank slate on which any idea can be impressed. To be understood, a new idea or phenomenon must connect to cognitive structures that are already available to the individual. Finally, a host that has understood a new idea must also be willing to believe it or to take it serious. For example, although you are likely to understand the proposition that your car was built by little green men from Mars, you are unlikely to accept that proposition without very strong evidence. Therefore, you will in general not memorize it, and the meme will not manage to infect you.

## **Retention**

The second stage of memetic replication is the retention of the meme in memory. By definition, memes must remain some time in memory, otherwise they cannot be called memes. The longer the meme stays, the more opportunities it will have to spread further by infecting other hosts. This is Dawkins's (1976) *longevity* characteristic for replicators.

Just like assimilation, retention is characterized by strong selection, which few memes will survive. Indeed, most of the things we hear, see or understand during the day are not stored in memory for longer than a few hours. Although you may have very clearly assimilated the news that the progressive liberal party won the Swaziland elections with 54% of the votes, you are unlikely to remember anything of this a week later--unless you live in Swaziland, perhaps. Retention will depend on how important the idea is to you, and how often it is repeated, either by recurrent perception or by internal rehearsal. All learning paradigms agree that experiences are encoded more strongly into memory by frequent reinforcement.

## **Expression**

To be communicated to other individuals, a meme must emerge from its storage as memory pattern and enter into a physical shape that can be perceived by others. This process may be called "expression". The most obvious means of expression is speech. Other common means for meme expression are text, pictures, and behavior. Expression does not require the conscious decision of the host to communicate the meme. A meme can be expressed simply by the way somebody walks or manipulates an object, or by what he or she wears.

Some retained memes will never be expressed, for example because the host does not consider the meme interesting enough for others to know, uses it unconsciously without it showing up in his or her behavior, does not know how to express it, or wants to keep it secret. On the other hand, the host may be convinced that the meme is so important that it must be expressed again and again to everybody he or she meets.

## **Transmission**

To reach another individual, an expression needs a physical carrier or medium which is sufficiently stable to transmit the expression without too much loss or deformation. Speech, for example, uses sound to transmit an expression, while text will be transmitted through ink on paper or electrical impulses in a wire. The expression will take the form of a physical signal, modulating the carrier into a specific shape from which the original meme can be re-derived.

This physical shape may be called the meme *vehicle*. For example, meme vehicles can be books, photographs, artefacts or CD-ROMs.

Selection at the transmission stage happens through either elimination of certain memes, when the vehicle is destroyed or gets corrupted before it is perceived by another individual, or through differential multiplication, when the vehicle is reproduced into many copies. For example, a manuscript may be put into the shredder or it may be turned into a book which is printed in thousands of copies. A radio communication may get lost because of noise, or it may be broadcasted to millions of listeners. Especially since the emergence of mass media, the transmission stage is the one where the contrast between successful and unsuccessful memes is largest, and where selection may have the largest impact.

## Meme fitness

The overall survival rate of a meme  $m$  can be expressed as the meme *fitness*  $F(m)$ , which measures the average number of memes at moment  $t$  divided by the average number of memes at the previous time step or "generation"  $t - 1$ . This fitness can be expressed in a simplified model as the product of the fitnesses or survival rates for each of the four stages, respectively assimilation  $A$ , retention  $R$ , expression  $E$  and transmission  $T$ :

$$F(m) = A(m) \cdot R(m) \cdot E(m) \cdot T(m)$$

$A$  denotes the proportion of memes vehicles encountered (or memes independently discovered) by the host that are assimilated.  $R$  represents the proportion of these assimilated memes that are retained in memory. Therefore,  $A \leq 1$ ,  $R \leq 1$ .  $E$  is the number of times a retained meme is expressed by the host.  $T$  is the number of copies of an expression that is transmitted to a potential new host. Unlike  $A$  and  $R$ ,  $E$  and  $T$  do not have an upper bound, although  $E$  is likely to be more restricted than  $T$ . Note that  $F$  is zero as soon as one of its components ( $A, R, E, T$ ) is zero. This expresses the fact that a meme must successfully pass through *all* four stages in order to replicate. Also note that for a meme to spread ( $F > 1$ ), you must have  $E > 1$  or  $T > 1$ .

## General Selection Criteria for Memes

Which memes will most successfully pass all these stages can be modelled by a series of selection criteria. These criteria are discussed in more detail in earlier papers (Heylighen, 1993, 1997). I will here basically situate them with respect to the four replication stages. The criteria can be grouped into different families, distinguished by the system responsible for the selection. At present, we have no method to derive the value of the fitness components from the degree to which a meme fulfils the different criteria. This does not mean that no predictions can be made, though. All other things being equal, *a meme that scores better on one of these criteria is predicted to become more numerous* in the population than a meme that scores worse.

This is a falsifiable hypothesis, which can be tested through experiments or observations. For that, it suffices to operationalize the tested criterion. This has already been done for criteria such as invariance (Van Overwalle & Heylighen, 1995), formality (Heylighen & Dewaele, 1998) or

conformity (cf. Boyd & Richerson, 1985), and seems relatively easy to do for the others as well by using standard social science methodologies, e.g. for developing test for personality traits.

## Objective Criteria

Objective criteria denote selection by phenomena or objects independent of the hosts and memes involved in the process. The *distinctiveness* criterion functions mainly during the assimilation stage. It states that phenomena that are distinct, detailed or contrasted are more likely to be noticed and understood, and therefore assimilated. The *invariance* and *controllability* criteria, on the other hand, apply mainly to the retention stage. According to the invariance criterion, phenomena that recur, independently of the way in which they are perceived, are more likely to be maintained in memory. Controllability notes that phenomena which react differentially to the subject's actions are also more likely to leave a permanent memory trace.

## Subjective Criteria

Subjective criteria represent selection by the subject who assimilates the meme. The main criteria at the assimilation stage are *novelty* (facilitates assimilation by attracting the subject's attention) and *simplicity* (requires less processing for the meme to be understood). The criterion of *coherence* (connection, consistency and support between new perception and existing memory trace) facilitates the understanding and acceptance parts of the assimilation stage, since it represents the ease with which the new meme can "fit in" with the memory that is already there. It also facilitates the retention stage since memories that cohere are more easy to retrieve and use and are therefore less likely to be forgotten. The criterion of *utility*, like controllability, functions mainly at the retention stage, since useful memes are more likely to be effectively used and thus reinforced, although it will also help assimilation, by making it more worthwhile for the host to do the effort to assimilate.

## Intersubjective Criteria

Intersubjective criteria represent selection through the interactions between different subjects. *Group utility* is an emergent criterion that is implicit in all four stages: a memes that is useful to the group of all its hosts is more likely to survive because it helps the group itself to survive and grow, and thus to absorb other individuals. *Authority* functions mainly at the assimilation stage: memes from authoritative sources, i.e. hosts or vehicles that are held in high regard or considered to represent expertise in the domain, will be more easily noticed and accepted. *Formality* (i.e. precise, unambiguous expression) too helps assimilation, at least of the original memetic content of the expression. It will contribute basically to what Dawkins (1976) calls copying-fidelity. (On the other hand, informal expression, because it tends to be simpler, may facilitate assimilation, but of an idea different from the one initially expressed). *Conformity*, the reinforcement of the same meme by different hosts belonging to the same group, will boost acceptance and retention (cf. Boyd & Richerson, 1985). *Expressivity*, the ease with which the meme can be expressed in an intersubjective medium, will obviously contribute to the expression stage. *Publicity*, finally, the effort put by the host(s) into the broad distribution of the message, will maximize transmission.

## Meme-centered Criteria

Finally, the meme-centered criteria represent selection on the level of the meme itself. They depend only on the internal structure of the meme, not on its "fit" to external selectors, such as subjects, objects, or groups. These criteria will typically select for "selfish" (cf. Heylighen, 1992) or "parasitic" (cf. Cullen, 1998) memes, whose only goal is to spread themselves, "infecting" a maximum of hosts without regard for their hosts' well-being. This does not imply that the same meme cannot satisfy both selfish and non-selfish criteria. Religions often have this mixture of parasitic and beneficial traits (cf. Cullen, 1998)

*Self-justification*, the degree to which the components of a meme mutually support each other, will facilitate understanding and acceptance. *Self-reinforcement*, the degree to which the meme stimulates its host to rehearse itself, e.g. by repetition, meditation, prayer, etc., will strengthen retention. *Intolerance*, the degree to which a meme excludes rival memes from being assimilated or retained, will also help the meme to retain a stable position in memory. *Proselytism*, the degree to which the meme urges its host to maximally spread the meme to other hosts, will increase the rates of expression and transmission.

stages\selectors	Objective	Subjective	Inter-subjective	Meme-centered
<b>Assimilation</b>	distinctiveness	novelty simplicity coherence	authority formality	self-justification
<b>Retention</b>	invariance controllability	coherence utility	conformity	self-reinforcement intolerance
<b>Expression</b>			expressivity	proselytism
<b>Transmission</b>			publicity	proselytism

**Table 1: a summary of the main selection criteria for memes, classified according to the stage during which they are most active, and the system responsible for the selection.**

## Conclusion

This simple four stage model helps us to analyse the mechanics of meme replication, and the different requirements a meme must satisfy to spread successfully. It moreover helps us to situate and to systematize a more intuitively developed list of objective, subjective, intersubjective and meme-centered selection criteria. Although the four stage model suggests a formula for calculating memetic fitness, the theory is as yet insufficiently developed to unambiguously determine the parameters of the equation. However, the list of selection criteria does produce a range of qualitative predictions, which can be empirically tested.

## Acknowledgments:

Francis Heylighen has been supported during these investigations by the Fund for Scientific Research - Flanders (FWO) as a Research Associate.

## References

- Boyd R. & Richerson P.J. (1985): *Culture and the Evolutionary Process*, (Chicago University Press, Chicago).
- Cavalli-Sforza L.L. & Feldman M.W. (1981): *Cultural Transmission and Evolution: a quantitative approach*, (Princeton University Press, Princeton).
- Cullen B. (1998): "Parasite Ecology and the Evolution of Religion", in: Heylighen F. & Aerts D. (ed.) (1998): *The Evolution of Complexity* (Kluwer Academic, Dordrecht)
- Dawkins R. (1976): *The Selfish Gene*, (Oxford University Press, New York).
- Heylighen F. (1992): "Selfish Memes and the Evolution of Cooperation", *Journal of Ideas*, Vol. 2, #4, pp 77-84.
- Heylighen F. (1993): "Selection Criteria for the Evolution of Knowledge", in: *Proc. 13th Int. Congress on Cybernetics* (Association Internat. de Cybérnétique, Namur), p. 524-528.
- Heylighen F. (1997): "Objective, subjective and intersubjective selectors of knowledge", *Evolution and Cognition* 3:1, p. 63-67.
- Heylighen F. & Dewaele J-M. (1998): "Formality of Language: definition, measurement and behavioral determinants", submitted to *Behavioral and Brain Sciences*
- Moritz E. (1990): "Memetic Science: I - General Introduction", *Journal of Ideas* 1, p. 1-23
- Van Overwalle, F., & Heylighen, F. (1995): "Relating covariation information to causal dimensions through principles of contrast and invariance", *European Journal of Social Psychology*, 25, p. 435-455.

**THE NECESSITY OF THEORETICAL CONSTRUCTS: A REBUTTAL OF  
THE BEHAVIORIST APPROACH TO MEMETICS**

**A COMMENTARY ON GATHERER'S PAPER: *WHY THE THOUGH  
CONTAGION METAPHOR IS RETARDING THE PROGRESS OF  
MEMETICS***

**By**

**Francis Heylighen**

CLEA, Free University of Brussels, Pleinlaan 2, B-1050 Brussels, Belgium

*fheylich@vub.ac.be*

Derek Gatherer's (1998) thesis that the "thought contagion" metaphor has slowed down progress in memetics is clearly exposed, well-argued and based on what is obviously a broad knowledge of the most recent literature relevant to memetics. However, the epistemology underlying his argumentation is totally outdated. Although Gatherer bases his thesis on Watson's behaviourism, which was developed during the 1920's and 1930's, his philosophy that only observable entities should be part of a scientific theory goes back much further, to Comte's 19<sup>th</sup> century positivism. It was elaborated in the beginning of the 20<sup>th</sup> century by the logical positivists (or logical empiricists) of the Vienna Circle.

His view that "scientists are generally constrained by the requirement that their models should *mirror* nature" (my emphasis) is older still. It was already criticized by some of the positivists themselves, and is now completely outdated (cf. Turchin 1993). At least, Gatherer seems to be aware that this mirroring conception is problematic when he notes that quantum physicists were forced to abandon it because of "the overwhelming weight of empirical evidence".

But positivism and behaviourism too have by now been completely abandoned and discredited, the first as a methodology for science in general, the second as a methodology for psychology in particular. Popper (1959), Wittgenstein and other philosophers of science have irrefutably shown why a positivist approach cannot work. Cognitive psychology has irreversibly taken over from behaviourist psychology (although it must be said that in psychology the behaviourist position still has some influence). There is no space here to review all arguments against positivism.

Let me just mention two facts that should be obvious:

- (1) All scientific models have unobservable, theoretical constructs at their core;
- (2) All empirical research, to be in any way meaningful or useful, presupposes - explicitly or implicitly - a theoretical construct.

As evidence for proposition (1), it suffices to look at the fundamental theories of physics, the most established and empirically best supported of all sciences. Essential concepts such as quarks, electromagnetic fields, wave functions, black holes or space-time geometries are all in principle unobservable. You do not even need to consider weird, non-classical theories such as quantum physics to find unobservable entities: Newtonian mechanics, the most basic of all scientific theories, is built on the concept of "force", a mysterious, non-material entity that can only be observed indirectly through its effects on the movement of objects.

This does not mean that such theories lack empirical support. Popper (1959) has made it clear that no theory or law can ever be verified or proven by observations. However, observations are necessary to refute theories that make incorrect predictions. Thus, a theory can be considered reliable if its implications have been repeatedly tested by experiments, without being refuted. As Turchin (1993) notes, models or theories are recursive generators of predictions. The hypothesis that nucleons consist of quarks does not generate any immediately testable predictions, since quarks can by definition never be observed outside of a nucleon. However, indirectly the quark hypothesis entails other hypotheses, which entail further hypotheses, and so on, until we come to the level where a hypothesis can be tested by observation. Of course, we would like to make the path from theoretical construct to empirical test as short as possible, but if we insist on too short paths, we will be stuck with theories that have a very restricted domain of application (Turchin, 1993).

As to proposition (2), all scientists will agree that observations are only useful if they can be repeated, and if repeated observations produce similar results. But this presupposes that the different observations and their respective results are all recognized as instances of the same phenomenon. If each observation is unique, there is no regularity and thus no scientific law to be inferred. However, the only way we can decide that different observations confirm the same law is because we have a theoretical conception of what such a law could be.

Let me illustrate this principle with an example that is mentioned by Gatherer as representative of the more 'behaviourist' field of social contagion: Phillips' (1974) test of the hypothesis that subjects who hear about others committing suicide are more likely to commit suicide themselves. At first sight, this seems a good illustration of Gatherer's view of a meme as a replicating behaviour: suicide in one person leads to suicide in another person. No need to postulate unobservable constructs, such as a "mnemon" for suicide, you might think.

The first difficulty with this view is that there is no simple, unique "suicide behaviour". You can commit suicide in millions of different ways: by drinking poison, inhaling gas, shooting yourself with a gun, jumping from a cliff, or driving at high speed into a wall. All these types of behaviour look utterly different to the observer. Yet, we would all classify them as "suicide". You might argue that they all lead to a person's heart stopping to beat, an unambiguous observation. But so do millions of other events which we do not classify as suicide. What distinguishes suicide from accidental death is the *intention* of the subject to make an end to his or her life - another theoretical construct which should be rejected outright by any genuine behaviourist.

But things get even worse for the behaviourist interpretation of suicide contagion. The hypothesis which Phillips confirmed is called the "Werther effect". The main observation is that people who were exposed to *stories* of suicide (originally, Goethe's novel "The suffering of the young Werther") are more inclined to commit suicide. Most of the victims of the Werther effect never actually witnessed somebody committing suicide. So, what was replicating was not actual suicidal *behaviour*, but the *idea* that suicide is a good way out of an apparently unsolvable problem. The only way to explain the variety of observations gathered by Phillips is to assume the existence and replication of an unobservable, mental entity: the intention to end one's life.

Of course, as Gatherer might reply, we can never actually observe the replication of the suicide idea. But as I argued earlier, this is a general problem with all scientific concepts. We can merely observe certain indirect effects of the postulated phenomenon, just like astronomers can only observe the presence of a black hole through its effect on nearby matter, not by seeing it directly. Such indirect effects can be misleading, as Gatherer illustrates with his example of the man making Windsor knots for his colleagues: our observations may make us think that these people each know how to make a Windsor knot. But Gatherer is wrong when he concludes from this example that it would be more fruitful to study the replication of the Windsor knot itself, rather than of the ability to make Windsor knots. This is similar to the argument that astronomers should better concern themselves with the study of the light rays that enter their telescopes rather than with the study of stars, galaxies and planets.

Let me suggest an alternative example of a copying machine containing a particular page. The machine may produce 1, 10, 1000, or one million copies of that page. Let us assume that the page contains information that could revolutionize society. For example, it could describe a method to produce unlimited amounts of energy through cold fusion. However, in how far the information on that page will affect the evolution of culture and society does not depend on the actual number of copies that come out of the machine. It depends on whether other people will read that information, whether they will understand it, believe it, communicate it to others, be willing to apply it, etc. These are all mental processes, which are not directly observable. The actual number of physical copies ("artefacts") is irrelevant if those copies are merely stored in a drawer, dropped in a trashbin, or blown away by the wind. It is the cognitive interpretation that determines whether the meme will affect other people's behaviour.

Gatherer also uses his Windsor knot example to argue that the situation in memetics is essentially different from the one in Mendelian genetics. Since one "mnemon" can produce hundred observable instantiations ("phenotypes"), there is no one-to-one relation between memotype and phenotype. But the same problem exists in genetics: one gene can produce dozens of feathers, antlers, branches, leaves or flowers. If you find different pieces of wood, shells or feathers in a geological deposit, you may be misled to think that they were produced by different organisms. In both cases, memetic and genetic, we have to be very careful how we infer the presence of an unobserved, but hypothesized, entity from observable phenomena. Instead of counting the number of Windsor knots, we should better test people on their ability to make a Windsor knot.

This ambiguous correspondence between observed and hypothesized phenomena is a general problem for all scientific modelling, not a specific difficulty for memetics. The problem is particularly acute in psychology, where the hypothesized, "mental" entities tend to be especially complex and variable. This explains why positivism, in the guise of behaviourism, has survived so long in that domain. Yet, by now psychology has developed a wide range of methods to empirically determine the presence of theoretical constructs. The different tests used in personality psychology to assess traits such as intelligence, introversion, dominance, etc. can serve as examples. There is no direct way to observe intelligence. Yet, the scores on different standardized IQ tests provide reliable predictions of a subject's chances for success in other intelligence-related activities, such as academic achievement. The simplest way to explain these

observed regularities is to assume that "intelligence", although a purely mental, unobservable entity objectively exists.

In a similar way, we could devise tests for different memes or features of memes, and then make predictions that people who "have" these memes, as measured by the test, will behave differently from people who score negatively on the test. For example, you could test subjects' knowledge of a hypothetical suicide meme, and then predict that the suicide rate in people with such knowledge will be higher than in those without such knowledge.

The basis of most psychological tests is natural language: you can simply *ask* people whether they consider suicide as an option. Of course, we know that such questions and their answers can be very subjective, and that there are plenty of spurious factors that may influence the results. However, that is insufficient reason to dismiss linguistic communication as a means to detect the presence of memes. *A priori* rejecting this method simply means rejecting the possibility of any kind of communication or transfer of information via language. And that would mean rejecting the idea that reading and writing papers, the activity we are engaged in at this moment, can contribute to the development of scientific knowledge.

To cope with such problems, psychologists have devised numerous techniques that minimize the subjectivity inherent in asking questions, for example: repeating the same question with different words, asking questions to which the answer is already known in order to test the subject's sincerity, extracting hidden, but reliable patterns from the list of responses, correlating answers to questions with independent observations, etc. All these techniques can be useful for memeticists who wish to operationalize their concept of a meme as a mental entity that influences behaviour.

To conclude my criticism of Gatherer's paper, I would like to ask him why the existing fields of social contagion and diffusion studies, which he holds up as an example of how memetics should develop, have produced so little understanding of cultural evolution. He suggests that this is merely due to an anti-evolutionary prejudice among social psychologists. To me it seems that the problem is more fundamental: it is the behaviourist approach itself which precludes the development of evolutionary models.

Indeed, if the only thing you study is observable behaviour and the way it replicates, then how can you ever account for variation and selection, the basic components of any evolutionary process? Variation in cultural evolution is not normally caused by the copying errors made by a printing press or photocopying machine: it is produced by the people who introduce certain changes in the behaviour they copy. Similarly, behaviours are not generally selected by observable transmission mechanisms such as printing or broadcasting. Behaviours or artefacts are replicated because people for some personal reason consider these behaviours to be worth copying. Thus, we can only get a real understanding of memetic variation and selection if we understand what goes on inside people's head when they listen to a song, read a book, or hear a story.

I have proposed a number of selection criteria that would allow us to predict which memes will survive or be eliminated (Heylighen, 1993, 1997, 1998). These criteria are mostly (though not

exclusively) internal to the person or to the meme. I have also suggested that these criteria, although not directly observable in the behaviourist sense, can be operationalized and empirically tested. That seems to me the way to go if we want to create a real science of memetics: to develop clear theoretical constructs, with a broad explanatory power, which are then operationalized so that they become empirically testable. I agree with Gatherer that many of the theoretical constructs used by memeticists, such as Lynch's (1998) distinction between "belief" and "awareness", are still problematic. But that does not mean that we should stop theorizing. On the contrary, we should reflect more profoundly on what a meme is and how it can be modelled in the most general and yet most precise and concrete way possible.

In conclusion, Gatherer's paper is useful because it reminds us of the importance of operationalization, and the danger of remaining stuck in vagueness and abstraction. However, it is downright misleading in its suggestion that in order to achieve operationalization we should throw out the essential theoretical construct of meme as memory unit, only because it is not directly observable.

## References

Gatherer, Derek (1998). Why the 'Thought Contagion' Metaphor is Retarding the Progress of Memetics, *Journal of Memetics - Evolutionary Models of Information Transmission*, **2**.  
[http://jom-emit.cfpm.org/1998/vol2/gatherer\\_d.html](http://jom-emit.cfpm.org/1998/vol2/gatherer_d.html)

Heylighen F. (1993). Selection Criteria for the Evolution of Knowledge. In: Proceedings 13<sup>th</sup> International Congress on Cybernetics (Association International de Cybernétique, Namur), 524-528.  
[ftp://ftp.vub.ac.be/pub/projects/Principia\\_Cybernetica/Papers\\_Heylighen/Knowledge\\_Selection\\_Criteria.txt](ftp://ftp.vub.ac.be/pub/projects/Principia_Cybernetica/Papers_Heylighen/Knowledge_Selection_Criteria.txt)

Heylighen F. (1997). Objective, subjective and intersubjective selectors of knowledge. *Evolution and Cognition*, **3**:63-67. <http://pespmc1.vub.ac.be/papers/knowledgeselectors.html>

Heylighen F. (1998). What makes a meme successful? In: Proceedings. 16<sup>th</sup> International Congress on Cybernetics (Association International de Cybernétique, Namur), in press.  
<http://pespmc1.vub.ac.be/Papers/MemeticsNamur.html>

Lynch, Aaron (1998). Units, Events and Dynamics in Memetic Evolution, *Journal of Memetics - Evolutionary Models of Information Transmission* **2**. [http://jom-emit.cfpm.org/1998/vol2/lynch\\_a.html](http://jom-emit.cfpm.org/1998/vol2/lynch_a.html)

Phillips, D.P. (1974) The influence of suggestion on suicide: substantive and theoretical implications of the Werther Effect. *American Sociological Review* 39:340-354.

Popper K. (1959). *The Logic of Scientific Discovery*. London: Hutchinson.

Turchin V. (1993). On Cybernetic Epistemology. *Systems Research*, 10:3-28.  
[ftp://ftp.vub.ac.be/pub/projects/Principia\\_Cybernetica/Papers\\_Turchin/Cybernetic\\_Epistemology.tex](ftp://ftp.vub.ac.be/pub/projects/Principia_Cybernetica/Papers_Turchin/Cybernetic_Epistemology.tex)

# **STRATEGIES IN MEME THEORY**

## **A COMMENTARY ON ROSE'S PAPER:**

## **CONTROVERSIES IN MEME THEORY**

**By**

**David L. Hull**

Department of Philosophy

Evanston, IL 60208

d-hull@nwu.edu

Rose does not care very much for the 'self', in particular the self doing anything, what he terms self-centered selectionism. He finds fault with those who "*believe that consciousness has the power to select memes in order to fulfil some life goal.*" He finds the claim that "*we can intentionally design and choose memes*" is "*entirely at odds with the proposal that culture is an evolutionary process.*"

My immediate response is to ask whether Rose objects to agency and intentionality only in memetic evolution or whether he also objects to it in genetic evolution as well. In most gene-based evolution, intentions and goals play no role in the selection process, but the same can be said for much of meme-based evolution. However, in both gene-based and meme-based evolution, selection in terms of conscious agents doing things for a purpose at least seems to play a role. Rose limits his discussion to meme-based evolution, but the problem arises just as seriously in gene-based evolution.

Darwin began his 'Origin of Species' with a long discussion of artificial selection. In Darwin's day and before, plant and animal breeders selected "sports" that they found useful and bred them selectively so that these variations became more common; e.g., ancon sheep. Darwin reasoned that if breeders could do so much with so little, one can only imagine how powerful natural selection must be. Apparently, Rose rejects Darwin's argument because the goals of conscious agents play a crucial role in artificial selection. Breeders did not produce the variations, but they did select them. Nowadays, breeders are in even a stronger position. They can also design the variations that they want. They can intentionally design and choose genes. Is artificial selection an "*unnatural*" process? Can it be explained without reference to selves, goals and intentions?

Rose seems to think that intentional behavior is Lamarckian if not downright miraculous. I think that it is neither. Even so, I agree with Rose that including it in theories of genetic or memetic evolution right now is a bad strategy because it introduces whole series of issues that are currently highly problematic. Perhaps someday in the future we can handle the roles that conscious agents play in both sorts of evolution, but not right now. The strategy that I am suggesting is the same one adopted by evolutionary biologists with respect to development. In Darwin's day, development looked promising. Some of the biggest names in science were working on it, but at the time evolutionary biologists could not find a way to integrate these findings in their theory. Instead, they worked their way around development, not to mention genetics. Their contemporaries were happy to point out the 'holes' in their theoretical structure, but the Darwinians could do nothing, until, finally, at the turn of the century, a theory of heredity

was formulated that (after some overly obtuse wrangling) fit nicely into evolutionary theory. We are still waiting for development. But right now things look promising!

Those of us working in memetics need not deride the self, agents, goals, intentions, and the like. Someday we may be able to work them into our theory, but not now. What is lost if we do restrict our attention? How important are intentions in organisms other than human beings? Organisms do strive to get away from predators as these predators strive to capture them. It is difficult not to see these phenomena in terms of goals. But even plants exhibit tropisms. They seem to strive toward sunlight and the like. But one thing that these organisms do not do is strive to evolve. Now, can all this 'striving' be interpreted solely in behaviorist terms? The history of behavioral psychology is spotty. Behavioral psychologists have made significant headway in their program, but even with respect to nonhuman animals, problems remain. However, I think that the safest bet, for now, is to adopt a behaviorist view of organisms other than humans. Little is lost, and quite a bit is gained. One might mention that operant learning is itself a selection process (Skinner 1974 and Glenn 1991).

When the issue is humans, the vast majority of people, scientists included, reject a purely behavioral account of behavior. We are conscious. We do have goals, and we do strive to realize these goals. I happen to be a bit cynical about human beings, me included. I think that a huge chunk of our behavior can be explained in purely behavioral terms without significant loss. For example, we constantly hear about all the damage that the human species is doing to the environment. I hasten to add that almost none of this degradation was intentional. We didn't mean to do it. To make matters worse, the vast majority of human beings have never heard of biological evolution. Of those who have, most do not understand it very well, and even fewer accept it. Only a very tiny percentage of the human race understands biological evolution and believes it. This small minority is in a position to introduce their intentions and goals into the evolutionary process. I shudder to think of the effects. In the past most of the modifications that we have consciously made in the environment on the basis of the best knowledge at the time did not turn out the way that we expected. It is still possible for people who understand evolutionary biology to produce the effects that they intend, but these effects are likely to be quite minor when compared to all the unconscious effects. If our intentions matter so little with respect to such phenomena, I fail to see why we have to mention them, except to dismiss them.

Perhaps the human race does not have a very good track record with respect to such huge phenomena, but how about other areas of human endeavor? On a day-to-day time scale, people behave intentionally. We get hungry and go to the local Wends to get a hamburger. Is there anything so wrong in saying as much? Most of our behavior can be explained entirely in behaviorist terms, but a substantial amount remains. Behavioral psychologists no longer claim that they can explain all human behavior in terms of operant conditioning, even of the most sophisticated sort. Maybe, some time in the future, they will be able to explain more, but all? That is highly questionable, especially when one turns to science itself. Science is as intentional a process as has ever appeared on Earth. When Rose urges that we reject references to 'selves', 'intentions', 'purposes', etc. in memetic evolution, is he not doing so for a purpose? He does not intend for us to change our focus of research? But all of this is just so much everyday talk. Once psychology gets sufficiently powerful, we can explain human behavior without reference, no matter how covert, to the entities and processes that Rose finds so objectionable. Perhaps so, but

I suggest a more cautious strategy. Just as evolutionary biologists have waited for the science of development to catch up before they attempted to incorporate it into evolutionary biology, I suggest that we memeticists exhibit similar caution in committing ourselves so totally with respect to a behaviorist stance for all of human behavior. Behavioral psychology is sufficiently well developed that we can use it and only it for the behavior of all other creatures without significant loss, but when we turn to us, I think that behavioral psychology is not currently up to the task. Let's hold these issues off to the side while we attack those problems that seem more soluble. Attacking the Big Issue right now and failing is certainly designed to harm our emerging research program. (OK, OK, maybe we are not doing it on purpose, but I suggest that it is the most likely result.)

To return to the opening paragraphs of this note, if Rose wants to make memetics as pure as he thinks it should be, then he must reject artificial selection as so much magic. To be sure in artificial selection the replicators are genes and not memes, but the selection process itself is riddled with intentional, goal-directed behavior, and we are currently in no position to eliminate reference to this behavior because we currently have nothing to put in its place. But if Rose's position is put in slightly different terms, I am all in favor of it. Because of all the problems surrounding 'selves' behaving 'intentionally' to fulfil 'goals', for now at least let's ignore it and continue to work on the vast majority of behavior in which such considerations do not arise.

Right now we have only scratched the surface with respect to these much less problematic phenomena. Once we get really good at handling them, then we can turn to the hardest parts. One of the faults that I find with philosophers is that we attack the big issues first. Scientists have had the good sense to attack at least some of the more soluble problems early on. As Medawar (1967) put it, science is the art of the soluble.

## References

Glenn, S.S., (1991). Contingencies and metacontingencies: relations among behavioral, cultural, and biological evolution, In P.A. Lamal (ed), *Behavioral Analysis of Societies and Cultural Practices*, Hemisphere Press, 39-73.

Medawar, P.B., (1967). *The Art of the Soluble*, Barnes & Noble.

Skinner, B.F., (1974). *About Behaviorism*, Knopf.

# **BRIDGING THE GAP: MEMETICS AS A METHODOLOGICAL TOOL TO CLOSE THE RANKS BETWEEN SOCIAL AND TRADITIONAL HISTORY**

**By**  
**Rogan Jacobson**

This paper will look at the evolutionary nature of information replication and transfer, and to use memetics as a methodological tool for historical analysis, and one which helps to rebuild the gap between traditional and sociological historical methods. Traditionally, modern scholars write histories using contemporary texts and analyzing them through the window of hindsight, causality and comparison to surrounding texts; social historians use various methods of analysis to contextualize the information and plot or map the surrounding action, yet these largely treat the content of the text as being of primary importance, more often than not leaving unassessed the structure of the information presentation.

Memetic science is in its infancy, and is currently being used in many and varied academic and non-academic fields; such as the work of Richard Dawkins in the field of evolutionary biology, Daniel Dennett in his research on the philosophy of mind, others in the fields of social sciences such as Francis Heylighen, and a host of others in such diverse fields as artificial intelligence research, corporate strategy planning, psychology, sociology and cultural evolution.

However, to the best of my knowledge memetics has not as yet been used extensively in the study of history. This is probably fairly easily explained. The science itself is, as I said, in its infancy. Therefore, a great deal of time and effort is currently being spent upon the method itself, seeking a more rigid analytical framework. These issues and many more are currently being fleshed out by memeticists, or interested academics, all over the globe. However, it cannot be denied that memetics does have something to offer us now, even in its youth, as an explanation of thought by analyzing the structures by which 'successful' or popular ideas transmit themselves from brain to brain. The science offers us the ability to structurally assign properties to certain types of ideas, empirically defend this formula, and then assess what this 'idea', this meme, can be expected to do or more importantly not do. For the historian working with, at best, often shaky accounts of actual events, and more often than not simply working with accounts expressing various successful ideas of the day, this science may prove invaluable.

Memetics can be used by the historian of ideas (from either camp) to determine the nature of any historical information transfer structure by mapping, as it were, the way in which the data propagates and the structural information/belief issues comprising the 'meme pool' of the time being investigated, thereby adding color to the historian's analysis and providing a fuller history.

To provide scope for an explication of memetics as methodology, albeit in its infancy, and as an academic discipline in its own right, I shall be looking specifically at the field of ancient religious history in the time of St Augustine. Ancient religious history, more so than more modern or contemporary history, suffers in research from a dearth of texts (due largely to lesser technologies and lower literacy rates) leaving traditional historians often forced to lean towards 'best guess' invention to complete a textual picture of the times. In cases where texts simply weren't written, and oral histories have been lost, historians are often faced with massive linear

breaks in their analysis, and resort to economic or geographical considerations to fill in the blanks. This concentration on the content of texts and the actions of contemporary actors, and not the structural nature of the information replication itself, presents a problem in modern historiography that memetic analysis should be used to redress.

In determining an interpretation of events, it is often practiced procedure for traditional historians to try and bleed their sources of the input of the author, to remove any personal, subjective comments in the interest of a Dragnet 'just the facts ma'am' objectivity. This history of 'heroes and monsters' is slowly being subsumed by social historical theory [1], such as that proposed by Roland Barthes in his essay on 'Historical Discourse'. Barthes argues that this practice of analysis is detrimental to the history itself. The historian, in writing, tries to remove his/her own persona from the work, to give the impression that we have before us nothing but statistical data. This attempt is seen by the structuralists as semantically impossible, given that not only must the historian interpret the data, but he/she must then write a history within the linguistic terms of his/her own contemporary society, their *langue*, and under the code-restrictions of their day.[2] . Traditional historians try to argue that the referent is speaking for itself, ontologically privileging the historical fact above and beyond the signs which make it known [3].

The sociological historians, however, approach the information in a radically different way. Structuralist history theories, such as those put forward by historians and social theorists such as Michel Foucault, stress that the analysis of historical texts, or our information packages, should look less toward the content of texts, and more toward the signification to be found concerning the wider society located therein, looking for ideological ciphers to extrapolate the world view surrounding them. As Foucault notes, the role of the 'new' history is not to interpret the document; instead it "organizes the document, divides it up, distributes it, orders, arranges it in levels, establishes series, distinguishes between what is relevant and what is not, discovers, elements, defines unities, describes relations"[4]. Both types of history have evolved towards each other, with traditional historians often using sociological method in their analysis[5], yet leaving the pragmatic 'traditional' assessment of replication mechanisms untouched.

In the case of St Augustine we see a prime example of this. In his lifetime, Augustine of Hippo was a prolific writer, seeking to promote a specific type of church and propagation of Christianity, namely one heavily steeped in the realities of the world around it, and one keen to acquire power and unity. This message, most clearly outlined textually in his work *De Doctrina Christiana*, centers around notions of education in Christian discourse and philosophical involvement with pagan writings and arguments as opposed to the discord among believers which arises when dealing with conflicting beliefs. "Nay, but let every good and true Christian understand that wherever truth may be found, it belongs to his Master; and while he recognizes and acknowledges the truth, even in their religious literature, let him reject the figments of superstition... And even when in the course of an historical narrative former institutions of men are described, the history itself is not to be reckoned among human institutions; because things that are past and gone and cannot be undone are to be reckoned as belonging to the course of time, of which God is the author and governor." [6].

Traditional historical analysis treats the text of the document as sacred, seeking objectivity in analysis by determining from the text the nature of the author's ideas on the structure of the church as a focal point for the changing nature of early church history. Structuralist approaches

use the text to depart from it to the audience and divine the nature of the society around St Augustine, arriving at assumptions about education, and learning etc. Both methods come together well, yet neither addresses the mechanical reasons behind the replication of such a text.

## Memetic fitness criteria

Memetic analysis proposes certain criteria for the survival and successful replication of an idea, any idea, without references to the empirical or theoretical 'truth' of the content. In this sense, memetics offers the researcher into the history of ideas an analytical tool for looking at text without fear of being subjectively swayed by the actual content of the texts being studied. The criteria by which memes are deemed fit or unfit in terms of their ability to replicate centre essentially, as noted by Dawkins [7], around their fecundity, copying fidelity and longevity. The analysis of these criteria provides a structural blueprint for the historian in that they allow a reassessment of texts in terms of the cultural structure, or meme pool, in which they are found to replicate.

The fecundity of any meme, i.e. its ability to be understood, presents many interesting possibilities for the historian. *De Doctrina Christiana*, for example, has been treated by modern scholars largely in light of its similarity to modern semiotic theory. The text is being treated and analyzed in terms of the effect such theories have had on modern intellectual discourse. However, by looking at the ability of contemporary sources to understand basic sign-theory, we see a smaller pool of possible replication. Such theories previously only existed in rhetorical schools of Greece and Rome, and held little sway over the general populous. To this end, we see St Augustine increasing the fecundity of the work by adding references and analyses that would easily be understood by a wider audience - music, the games, pagan epics and myths etc. Thus the historian is able to see, before analyzing the content, that the meme of the work is a fit one for the surrounding meme pool and ought to do well.

The copying fidelity of any meme can also be measured a great degree of accuracy. Again, our example of *De Doctrina Christiana* would replicate well. It consists of short paragraphs, easily sent in letter form, or read out in sermons. It refers directly to, and quotes heavily from, Scriptures which were already circulated and can be written or spoken at academic or basic levels of intellectual understanding.

And finally, the longevity of the work shows the historian how likely the meme is to propagate. In *De Doctrina Christiana*, we see a work which although specific in its reference to biblical texts (already strong memes, and already shown to be long-lived) other contemporary references are kept at a minimum and metaphor is brought into play. In this way the work is able to achieve an intellectually 'timeless' quality, which enables it to persist and replicate long after references to local conditions have passed from memory.

Fitness criteria can easily be adjusted constantly according to information on the available means of replication. That is to say, for example, a meme in fourth century Europe would not necessarily be deemed unfit for its inability to transfer well on television. All that is required for an idea to be treated as a meme is that it should behave like one; ie that it should be highly successful in replicating itself within a given community. Further and more specific memetic

conditions for replication fitness are contained in the handout of the work by Ron Hale-Evans[8]. These criteria, more specific in their inquiry, provide the historian a blueprint, as it were, for the strength of any meme in a given society, and a base from which to embark upon other methodologies of analysis. In terms of cultural evolution, therefore, the researcher is able to generate a blueprint of the meme pool of the day by looking at means of information dispensation and replication and the mechanisms by which the more successful memes of the time propagate; historians of any period can usually pinpoint, with a fair degree of accuracy, such factors as influence the replication of ideas:

- languages spoken, and who understood them,
- state of literacy
- most common means of general populous communication
- technological means of the society
- interaction with other surrounding societies and ease of transfer between them

For example, we are able to see that in 4th century Europe oral histories delivered in terms of narrative, sermon and parable, philosophical debate, letter writing and book production were the key machinery for the replication of ideas concerning the early church. Memetic analysis of these mechanisms, determining who had access to them and how they were used, provides the historian with a clear view of possible successful memes in terms of structure.

## **What's in a meme: increased objectivity in historical analysis**

Memes replicate without any overt responsibility for the nature of their content. For the historian this provides a clean arena for analysis; it cannot be argued that the meme of Christianity was passed on and replicated itself successfully in its chosen form because it was right, or because it was true, but merely because over the period of its birth and growth, the form in which the meme survived was the most successfully copied, more easily able to reach 'uninfected minds' (for reasons open to discussion) and more able to meld with the other beliefs stored in the mind of the follower. The structure of the information package was, and indeed still is, a memetically fit one.

A meme that successfully replicates itself is not necessarily *good* or *right*; rather it merely operates more effectively within the cultural climate it finds itself. We see then that St Augustine's meme of Christianity was able to replicate itself more frequently, among more minds and across more generations while maintaining a high degree of copying fidelity than the other religious memes competing with Christianity at the time. One meme supplanted the other for the predominant mind-space, yet both 'survived'. Memetics is a useful method for the historian seeking to paint a picture of a contemporary meme-pool in concert with more traditional assessments of ancient European culture, which as often as not take into account at a paramount level; such things as the local geography, ideological voids left with the passing of the strength of the Roman Empire, military issues concerning the 'friendship' of the Church with the emperor, the nature of the religion in a historical sense, perceived religious voids and other such issues. In other words, historians of the period can use memetics to look at the replication of two complementary and competing ideas of Christianity, not merely at paganism and Christianity. Further, memetic methodology can be used to more effectively assess the effect of the actions of the individuals who hold and replicate these ideas upon the replication and mutation of that idea

throughout the meme pool. Normal methods of historical inquiry, whether social or traditional, are inadequate in this field, as they concentrate too heavily upon the actions of the 'agents', and not enough upon the structure and force of the ideas themselves.

This brings us to an interesting criticism of memetic method, namely the nature and relevance of truth claims. Memetic historical methodology involves investigating what can be referred to as the mechanisms for a 'successful idea', and yet in the case of religion it happens to be an idea (or more correctly a complex series of ideas) which is packed to the rafters with truth statements. It may therefore be argued that since this particular idea (or indeed any other successful meme) most aptly succeeds in the human cultural sphere, it lends itself to notions of truth. This, however, is not a function of the science. All we are researching here is the strength (replicating ability/attractiveness/copying fidelity) of an idea, and how one meme can supplant another in the minds of believers, an area which in no way suggests that in our case Augustine was 'right', and orthodox pagan religions 'wrong', but does imply that Augustine's ideas were memetically strong in their phrasing and reproduction, giving them a far better chance of replication in the society of the time.

Memetic theory in history is a methodology for the assessment of the replication possibilities of ideas, not their 'truth'. We could just as easily use memetics to assess the 'flat earth theory' (still alive today in some ever decreasing circles) without ever mentioning the 'truth' of its claims. The only 'truth' claimed by the science is a mechanical one. That is, the prevalence and success of certain ideas to reach more people can be traced to certain memetics structural imperatives, and the growth of such ideas more accurately charted. Most methods of analysis when charting or explaining the growth and history of ideas have relied upon their perceived effect upon the 'readership', and their historical relationship to what has gone before. Memetics, however, provides a more objective standpoint. With it we can look at the replication of an idea in itself, then go on to attribute this idea with certain proactive 'abilities', or avenues for replication and spread. This, more fully explained and researched idea will then be submitted to regular methods of analysis, but from all sides. The language will come under scrutiny, also the subject matter, the changes in audience, the perspective of the writer, the perceived world-view of the laity and the goal of the work. This, I believe, will give us a new direction to ponder in terms of historical research and analysis.

## Select Bibliography

Dawkins, R, *The Selfish Gene*, Paladin 1976

Eco, U, *The Name of the Rose*, Picador, 1984 St Augustine, *De Doctrina Christiana*, book II, translation from Select Library of Nicene and Post-Nicene Fathers at URL  
<<http://ccat.sas.upenn.edu/jod/augustine/ddc.html>>.

Foucault, M, *The Archaeology of Knowledge*, trans. A.M. Sheridan Smith, London: Tavistock 1972

Winks, R.W., *The Historian as Detective: Essays on Evidence*, Harper & Row 1968  
*Critical and Effective histories: Foucault's Methods and Historical Sociology*, ed Mitchell Dean, Routledge 1994

*Cultural Selection*, Agner Fog, Electronic Book at URL <<http://announce.com/agner/cultsel/>>

*Discourse and Difference: Post-structuralism, feminism and the moment of history*, ed Andrew Milner and Chris Worth, Centre for general and comparative literature, Monash University 1990

*Educational theory and practice in St Augustine*, George Howie M.A. M.Ed PhD, Routledge and Kegan Paul, London 1969

*Historical Sociology*, Philip Abrams, The Pitman Press, 1982

*Justifying Historical Descriptions*, C Behan McCullagh, Cambridge University Press 1984

*Main Trends in Cultural History: Ten Essays*, ed Willem Melching & Wyger Velema, Rodopi, Amsterdam - Atlanta 1994

Barthes, R, "Historical Discourse" in *Structuralism: a Reader*, ed M.Lane, Jonathan Cape, Thirty Bedford Square, London 1970, pp145 - 69. Hale-Evans, R. "Memetics: A Systems Metabiology", Version 950220 at URL <<http://www.apocalypse.org/pub/u/rwhe/memetics.html>> Last, J "The nature of history" pp 142 - 157 in *Interpreting archaeology: finding meaning in the past*. Ian Hodden et al, Routledge, London 1995

Markus R.A., "Signs, Communication and Communities", pp.97 - 108 in *De Doctrina Christiana : a classic of western culture* edited by Duane W.H. Arnold and Pamela Bright. Notre Dame : University of Notre Dame Press, c1995. Meard, M. "Re-reading (Vestal) Virginity", ch 11 p171 in *Women in Antiquity*, eds R.Hawley and B Lerick 1995.

[1] See Jonathan Last's article "The nature of history" pp 142 - 157 in *Interpreting archaeology: finding meaning in the past*. Ian Hodden et al, Routledge, London 1995

[2] Barthes, R, "Historical Discourse", p149: "What really happens is that the author discards the human persona but replaces it with an 'objective' one; the authorial subject is as evident as ever, but it has become an objective subject. Structuralism: a Reader, ed M.Lane, Jonathan Cape, Thirty Bedford Square, London 1970, pp145 - 69.

[3] See Beard's reread of Beard in "Re-reading (Vestal) Virginity", ch 11 p171 in *Women in Antiquity*, eds R.Hawley and B Lerick 1995.

[4] Foucault, M *The Archaeology of Knowledge*, trans. A.M. Sheridan Smith, London: Tavistock 1972, 6-7

[5] For example see RA Markus, "Signs, Communication and Communities", where in his analysis of St Augustine's *De Doctrina Christiana* he stresses the need to use semiotic methodology, pp.97 - 108 in *De Doctrina Christiana : a classic of western culture* edited by Duane W.H. Arnold and Pamela Bright. Notre Dame : University of Notre Dame Press, c1995.

[6] St Augustine, *De Doctrina Christiana*, book II, translation from Select Library of *Nicene and Post-Nicene Fathers* chs 18 (28) and 28 (42) at URL  
<<http://ccat.sas.upenn.edu/jod/augustine/ddc.html>>.

[7] Dawkins, R *The Selfish Gene*, Paladin 1976 p208

[8] Ron Hale-Evans, Memetics: A Systems Metabiology, Version 950220 at URL  
<<http://www.apocalypse.org/pub/u/rwhe/memetics.html>>

## GREAT MEN, GREAT THOUGHTS, AND THE ENVIRONMENT

By  
William James

Lecture delivered before the Harvard Natural History Society.

Published in the *Atlantic Monthly*, October, 1880

A remarkable parallel, which I think has never been noticed, obtains between the facts of social evolution on the one hand, and of zoological evolution as expounded by Mr. Darwin on the other.

It will be best to prepare the ground for my thesis by a few very general remarks on the method of getting at scientific truth. It is a common platitude that a complete acquaintance with any one thing, however small, would require a knowledge of the entire universe. Not a sparrow falls to the ground but some of the remote conditions of his fall are to be found in the Milky Way, in our federal constitution, or in the early history of Europe. That is to say, alter the Milky Way, alter our federal constitution, alter the facts of our barbarian ancestry, and the universe would so far be a different universe from what it now is. One fact involved in the difference might be that the particular little street-boy who threw the stone which brought down the sparrow might not find himself opposite the sparrow at that particular moment; or, finding himself there, he might not be in that particular serene and disengaged mood of mind which expressed itself in throwing the stone. But, true as all this is, it would be very foolish for anyone who was inquiring the cause of the sparrow's fall to overlook the boy as too personal, proximate, and so to speak anthropomorphic an agent, and to say that the true cause is the federal constitution, the westward migration of the Celtic race, or the structure of the milky way. If we proceeded on that method, we might say with perfect legitimacy that a friend of ours, who had slipped on the ice upon his door-step and cracked his skull, some months after dining with thirteen at the table, died because of that ominous feast. I know, in fact, one such instance; and I might, if I chose, contend with perfect logical propriety that the slip on the ice was no real accident. "There are no accidents," I might say, "for science. The whole history of the world converged to produce that slip. If anything had been left out, the slip would not have occurred just there and then. To say it would is to deny the relations of cause and effect throughout the universe. The real cause of the death was not the slip, *but the conditions which engendered the slip*, -- and among them his having sat at a table, six months previous, one among thirteen. *That* is truly the reason why he died within the year."

It will soon be seen whose arguments I am, in form, reproducing here. I would fain lay down the truth without polemics or recrimination. But unfortunately we never fully grasp the import of any true statement until we have a clear notion of what the opposite untrue statement would be. The error is needed to set off the truth, much as a dark background is required for exhibiting the brightness of a picture. And the error which I am going to use as a foil to set off what seems to me the truth of my own statements is contained in the philosophy of Mr. Herbert Spencer and his disciples. Our problem is: What are the causes that make communities change from generation to generation, - that make the England of Queen Anne so different from the England of Elizabeth, the Harvard College of to-day so different from that of thirty years ago?

I shall reply to this problem: The difference is due to the accumulated influences of individuals, of their examples, their initiatives, and their decisions. The Spenserian school replies: The changes are irrespective of persons, and independent of individual control. They are due to the environment, to the circumstances, the physical geography, the ancestral conditions, the increasing experience of our relations; to everything, in fact, except the Grants and the Bismarcks, the Joneses and the Smiths.

Now, I say that these theorizers are guilty of precisely the same fallacy as he who should ascribe the death of his friend to the dinner with thirteen, or the fall of the sparrow to the Milky Way. Like the dog in the fable, who drops his real bone to snatch at its image, they drop the real causes to snatch at others, which from no possible human point of view are available or attainable. Their fallacy is a practical one. Let us see where it lies. Although I believe in free-will myself, I will waive that belief in this discussion, and assume with the Spenserians the predestination of all human actions. On that assumption I gladly allow that were the intelligence investigating the man's or the sparrow's death omniscient and omnipresent, able to take in the whole of time and space at a single glance, there would not be the slightest objection to the milky way or the fatal feast being invoked among the sought-for causes. Such a divine intelligence would see instantaneously all the infinite lines of convergence towards a given result, and it would, moreover, see impartially: it would see the fatal feast to be as much a condition of the sparrow's death as of the man's; it would see the boy with the stone to be as much a condition of the man's fall as of the sparrow's.

The human mind, however, is constituted on an entirely different plan. It has no such power of universal intuition. Its finiteness obliges it to see but two or three things at a time. If it wishes to take wider sweeps it has to use 'general ideas,' as they are called, and in so doing to drop all concrete truths. Thus, in the present case, if we as men wish to feel the connection between the Milky Way and the boy and the dinner and the sparrow and the man's death, we can do so only by falling back on the enormous emptiness of what is called an abstract proposition. We must say: All things in the world are fatally predetermined, and hang together in the adamantine fixity of a system of natural law. But in the vagueness of this vast proposition we have lost all the concrete facts and links; and in all practical matters the concrete links are the only things of importance. The human mind is essentially partial. It can be efficient at all only by **picking out** what to attend to, and ignoring everything else, - by narrowing its point of view. Otherwise, what little strength it has is dispersed, and it loses its way altogether. Man always wants his curiosity gratified for a particular purpose. If, in the case of the sparrow, the purpose is punishment, it would be idiotic to wander off from the cats, boys, and other possible agencies close by in the street, to survey the early Celts and the Milky Way: the boy would meanwhile escape. And if, in the case of the unfortunate man, we lose ourselves in contemplation of the thirteen-at-table mystery, and fail to notice the ice on the step and cover it with ashes, some other poor fellow, who never dined out in his life, many slip on it in coming to the door, and fall and break his head too.

It is, then, a necessity laid upon us as human beings to limit our view. In mathematics we know how this method of ignoring and neglecting quantities lying outside a certain range has been adopted in the differential calculus. The calculator throws out all the "infinitesimals of the quantities he is considering. He treats them (under certain rules) as if they did not exist. In

themselves they exist perfectly all the while; but they are as if they did not exist for the purposes of his calculation. Just so an astronomer, dealing with the tidal movements of the ocean, takes no account of the waves made by the wind, or by the pressure of all the steamers which day upon night are moving their thousands of tons upon its surface. Just so the marksman, in sighting his rifle, allows for the motion of the wind, but not for the equally real motion of the earth and solar system. Just so a business man's punctuality may overlook an error of five minutes, while a physicist, measuring the velocity of light, must count each thousandth of a second.

There are, in short, ***different cycles of operation*** in nature; different departments, so to speak, relatively independent of one another, so that what goes on at any moment in one may be compatible with almost any condition of things at the same moment in the next. The mould on the biscuit in the store-room of a man-of-war vegetates in absolute indifference to the nationality of the flag, the direction of the voyage, the weather, and the human dramas that may go on onboard; and a mycologist may study it in complete abstraction from all these larger details. Only by so studying it, in fact, is there any chance of the mental concentration by which alone he may hope to learn something of its nature. On the other hand, the captain, who in maneuvering the vessel through a naval fight, should think it necessary to bring the moldy biscuit into his calculations, would very likely lose the battle by reason of the excessive "thoroughness" of his mind.

The causes which operate in these incommensurable cycles are connected with one another only ***if we take the whole universe into account.*** For all lesser points of view it is lawful - nay, more, it is for human wisdom necessary - to regard them as disconnected and irrelevant to one another.

And this brings us nearer to our special topic. If we look at an animal or a human being, distinguished from the rest of his kind by the possession of some extraordinary peculiarity, good or bad, we shall be able to discriminate between the causes which originally ***produced*** the peculiarity in him and the causes that ***maintained*** it after it is produced; and we shall see, if the peculiarity be one that he was born with, that these two sets of causes belong to two such irrelevant cycles. It was the triumphant originality of Darwin to see this, and to act accordingly. Separating the causes of production under the title of 'tendencies to spontaneous variation,' and relegating them to a physiological cycles which he forthwith agreed to ignore altogether, he confined his attention to the causes of preservation, and under the names of natural selection and sexual selection studied them exclusively as functions of the cycle of the environment.

Pre-Darwinian philosophers had also tried to establish the doctrine of descent with modification; but they all committed the blunder of clumping the two cycles of causation into one. What preserves an animal with his peculiarity, if it be a useful one, they saw to be the nature of the environment to which the peculiarity was adjusted. The giraffe with his peculiar neck is preserved by the fact that there are, in his environment, tall trees whose leaves he can digest. But these philosophers went further, and said that the presence of the trees not only maintained an animal with a long neck to browse upon their branches, but also produced him. They ***made*** his neck long by the constant striving they aroused in him to reach up to them. The environment, in short, was supposed by these writers to mould the animal by a kind of direct pressure, very much as a seal presses the wax into harmony with itself. Numerous instances were given of the way in which this goes on under our eyes. The exercise of the forge makes the right arm strong, the

palm grows callous to the oar, the mountain air distends the chest, the chased fox grows cunning and the chased bird shy, the arctic cold stimulates the animal combustion, and so forth. Now these changes, of which many more examples might be adduced, are at present distinguished by the special name of *adaptive* changes. Their peculiarity is that that very feature in the environment to which the animal's nature grows adjusted, itself produces the adjustment. The 'inner relation,' to use Mr. Spencer's phrase, 'corresponds' with its own efficient cause.

Darwin's first achievement was to show the utter insignificance in amount of these changes produced by direct adaptation, the immensely greater mass of changes being produced by internal molecular accidents, of which we know nothing. His next achievement was to define the true problem with which we have to deal when we study the effects of the visible environment on the animal. That problem is simply this: Is the environment more likely to *preserve or to destroy him*, on account of this or that peculiarity with which he may be born? In giving the name of "accidental variations" to those peculiarities with which an animal is born, Darwin does not for a moment mean to suggest that they are not the fixed outcome of natural law. If the total system of the universe be taken into account, the causes of these variations and the visible environment which preserves or destroys them, undoubtedly do, in some remote and round-about way, hang together. What Darwin means is, that, since the environment is a perfectly known thing, and its relations to the organism in the way of destruction or preservation are tangible and distinct, it would utterly confuse our finite understandings and frustrate our hopes of science to mix in with it facts from such a disparate and incommensurable cycle as that in which the variations are produced. This last cycle is that of occurrences before the animal is born. It is the cycle of influences upon ova and embryos; in which lie the causes that tip them and tilt them towards masculinity or femininity, towards strength or weakness, towards health or disease, and towards divergence from the parent type. What are the causes there?

In the first place, they are molecular and invisible, - inaccessible, therefore, to direct observation of any kind. Secondly, their operations are compatible with any social, political and physical conditions of the environment. The same parents, living in the same environing conditions, may at one birth produce a genius, at the next an idiot or a monster. The visible external conditions are therefore not direct determinants of this cycle; and the more we consider the matter, the more we are forced to believe that two children of the same parents are made to differ from each other by causes as disproportionate to their ultimate effects as is that famous pebble on the Rocky Mountain crest, which separates two rain-drops, to the Gulf of St. Lawrence and the Pacific Ocean towards which it makes them severally flow.

The great mechanical distinction between transitive forces and discharging forces is nowhere illustrated on such a scale as in physiology. Almost all causes there are forces of *detent*, which operate by simply unlocking energy already stored up. They are upsetters of unstable equilibria, and the resultant effect depends infinitely more on the nature of the materials upset than on that of the particular stimulus which joggles them down. Galvanic work, equal to unity, done on a frog's nerve will discharge from the muscle to which the nerve belongs mechanical work equal to seventy thousand; and exactly the same muscular effect will emerge if other irritants than galvanism are employed. The irritant has merely started or provoked something which then went on of itself, - as a match may start a fire which consumes a whole town. And qualitatively as well as quantitatively the effect may be absolutely incommensurable with the cause. We find this

condition of things in all organic matter. Chemists are distracted by the difficulties which the instability of albuminoid compounds opposes to their study. Two specimens, treated in what outwardly seem scrupulously identical conditions, behave in quite different ways. You know about the invisible factors of fermentation, and how the fate of a jar of milk - whether it turn into a sour clot or a mass of koumiss - depends on whether the lactic acid ferment or the alcoholic is introduced first, and gets ahead of the other in starting the process. Now, when the result is the tendency of an ovum, itself invisible to the naked eye, to tip towards this direction or that in its further evolution, - to bring forth a genius or a dunce, even as the rain-drop passes east or west of the pebble, - is it not obvious that the deflecting cause must lie in a region so recondite and minute, must be such a ferment of a ferment, an infinitesimal of so high an order, that surmise itself may never succeed even in attempting to frame an image of it?

Such being the case, was not Darwin right to turn his back upon that region altogether, and to keep his own problem carefully free from all entanglement with matters such as these? The success of his work is a sufficient affirmative reply.

And this brings us at last to the heart of our subject. The causes of production of great men lie in a sphere wholly inaccessible to the social philosopher. He must simply accept geniuses as data, just as Darwin accepts his spontaneous variations. For him, as for Darwin, the only problem is, these date being given: How does the environment affect them, and how do they affect the environment? Now, I affirm that the relation of the visible environment to the great man is in the main exactly what it is to the "variation" in the Darwinian philosophy. It chiefly adopts or rejects, preserves or destroys, in short *selects* him. And whenever it adopts and preserves the great man, it becomes modified by his influence in an entirely original and peculiar way. He acts as a ferment and changes its constitution, just as the advent of a new zoological species changes the faunal and floral equilibrium of the region in which it appears. We all recollect Mr. Darwin's famous statement of the influence of cats on the growth of clover in their neighborhood. We all have read of the effects of the European rabbit in New Zealand, and we have many of us taken part in the controversy about the English sparrow here, - whether he kills more canker worms, or drives away most native birds. Just so the great man, whether he be an important from without like Clive in India or Agassiz here, or whether he spring from the soil like Mahomet or Franklin, brings about a rearrangement, on a large or a small scale, of the pre-existing social relations.

The mutations of societies, then, from generation to generation, are in the main due directly or indirectly to the acts or the examples of individuals whose genius was so adapted to the receptivities of the moment, or whose accidental position of authority was so critical that they became ferments, initiators of movements, setters of precedent or fashion, centers of corruption, or destroyers of other persons, whose gifts, had they had free play, would have led society in another direction.

We see this power of individual initiative exemplified on a small scale all about us, and on a large scale in the case of the leaders of history. It is only following the common-sense method of a Lyell, a Darwin and a Whitney to interpret the unknown by the known, and reckon up cumulatively the only causes of social change we can directly observe. Societies of men are just like individuals, in that both at any given moment offer ambiguous potentialities of development. Whether a young man enters business or the ministry may depend on a decision which has to be

made before a certain day. He takes the place offered in the counting-house, and is *committed*. Little by little, the habits, the knowledge of the other career, which once lay so near, cease to be reckoned even among his possibilities. At first, he may sometimes doubt whether the self he murdered in that decisive hour might not have been the better of the two; but with the years such questions themselves expire, and the old alternative *ego*, once so vivid, fades into something less substantial than a dream. It is no otherwise with nations. They may be committed by kings and ministers to peace or war, by generals to victory or defeat, by prophets to this religion or that, by various geniuses to fame in art, science or industry. A war is a true point of bifurcation of future possibilities. Whether it fail or succeed, its declaration must be the starting-point of new policies. Just so does a revolution, or any great civic precedent, become a deflecting influence, whose operations widen with the course of time. Communities obey their ideals; and an accidental success fixes an ideal, as an accidental failure blights it.

Would England have to-day the "imperial" ideal which she now has, if a certain boy named Bob Clive had shot himself, as he tried to do, at Madras? Would she be the drifting raft she is now in European affairs if a Frederic the Great had inherited her throne instead of a Victoria, and if Messrs. Bentham, Mill, Cobden, and Bright had all been born in Prussia? England has, no doubt, to-day precisely the same intrinsic value relatively to the other nations that she ever had. There is no such fine accumulation of human material upon the globe. But in England the material has lost effective form, while in Germany it has found it. Leaders give the form. Would England be crying forward and backward at once, as she does now, "letting I will not wait upon I would," wishing to conquer but not to fight, if her ideal had in all these years been fixed by a succession of statesmen of supremely commanding personality, working in one direction? Certainly not. She would have espoused, for better or worse, either one course or another. Had Bismarck died in his cradle, the Germans would still be satisfied with appearing to themselves as a race of spectacled *Gelehrten* and political herbivora, and to the French as *ces bons*, or *ces naïfs, Allemands*. Bismarck's will showed them, to their own great astonishment, that they could play a far livelier game. The lesson will not be forgotten. Germany may have many vicissitudes, but they –

*"will never do away, I ween  
The marks of that which once hath been"* -

of Bismarck's initiative, namely, from 1860 to 1873.

The fermentative influence of geniuses must be admitted as, at any rate, one factor in the changes that constitute social evolution. The community *may* evolve in many ways. The accidental presence of this or that ferment decides in which way it *shall* evolve. Why, the very birds of the forest, the parrot, the mino, have the power of human speech, but never develop it of themselves; someone must be there to teach them. So with us individuals. Rembrandt must teach us to enjoy the struggle of light with darkness, Wagner to enjoy peculiar musical effects; Dickens gives a twist to our sentimentality, Artemus Ward to our humor; Emerson kindles a new moral light within us. But it is like Columbus's egg. "All can raise the flowers now, for all have got the seed." But if this be true of individuals in the community, how can it be false of the community as a whole? If shown a certain way, a community may take it; if not, it will never find it. And the ways are to a large extent indeterminate in advance. A nation may obey either of many

alternative impulses given by different men of genius, and still live and be prosperous, just as a man may enter either of many businesses. Only, the prosperities may differ in their type. But the indeterminism is not absolute. Not every "man" fits every "hour." Some incompatibilities there are. A given genius may come either too early or too late. Peter the Hermit would now be sent to an insane asylum. John Mill in the tenth century would have lived and died unknown. Cromwell and Napoleon need their revolutions, Grant his civil war. An Ajax gets no fame in the day of telescopic-sighted rifles; and, to express differently an instance which Spencer uses, what could a Watt have affected in a tribe which no precursive genius had taught to smelt iron or to turn a lathe?

Now, the important thing to notice is that what makes a certain genius now incompatible with his surroundings is usually the fact that some previous genius of a different strain has warped the community away from the sphere of his possible effectiveness. After Voltaire, now Peter the hermit; after Charles IX and Louis XIV, no general protestantization of France; after a Manchester school, a Beaconsfield's success is transient; after a Philip II, a Castelar makes little headway; and so on. Each bifurcation cuts off certain sides of the field altogether, and limits the future possible angles of deflection. A community is a living thing, and in words which I can do no better than quote from Professor Clifford, "it is the peculiarity of living things not merely that they change under the influence of surrounding circumstances, but that any change which takes place in them is not lost but retained, and as it were built into the organism to serve as the foundation for future actions. If you cause any distortion in the growth of a tree and make it crooked, whatever you may do afterwards to make the tree straight the mark of your distortion is there; it is absolutely indelible; it has become part of the tree's nature. ... Suppose, however, that you take a lump of gold, melt it, and let it cool. ... No one can tell by examining a piece of gold how often it has melted and cooled in geologic ages, or even in the last year by the hand of man. Anyone who cuts down an oak can tell by the rings of its trunk how many times winter has frozen it into widowhood, and how many times summer has warmed it into life. A living being must always contain within itself the history, not merely of its own existence, but of all its ancestors."

Every painter can tell us how each added line deflects his picture in a certain sense. Whatever lines follow must be built on those first laid down. Every author who starts to rewrite a piece of work knows how impossible it becomes to use any of the first-written pages again. The new beginning has already excluded the possibility of those earlier phrases and transitions, while it has at the same time created the possibility of an indefinite set of new ones, no one of which, however, is completely determined in advance. Just so the social surroundings of the past and present hour exclude the possibility of accepting certain contributions from individuals; but they do not positively define what contributions shall be accepted, for in themselves they are powerless to fix what the nature of the individual offerings shall be.

Thus social evolution is a resultant of the interaction of two wholly distinct factors, - the individual, deriving his peculiar gifts from the play of physiological and infra-social forces, but bearing all the power of initiative and origination in his hands; and, second, the social environment, with its power of adopting or rejecting both him and his gifts. Both factors are essential to change. The community stagnates without the impulse of the individual. The impulse dies away without the sympathy of the community.

All this seems nothing more than common-sense. All who wish to see it developed by a man of genius should read that golden little work, Bagehot's *Physics and Politics*, in which (it seems to me) the complete sense of the way in which concrete things grow and change is as livingly present as the straining after a pseudo-philosophy of evolution is livingly absent. But there are never-wanting minds to whom such views seem personal and contracted, and allied to an anthropomorphism long exploded in other fields of knowledge. "The individual withers, and the world is more and more," to these writers; and in a Buckle, a Draper, and a Taine we all know how much the "world" has come to be almost synonymous with the **climate**. We all know, too, how the controversy has been kept up between the partisans of a "science of history" and those who deny the existence of anything like necessary "laws" where human societies are concerned. Mr. Spencer, at the opening of his *Study of Sociology*, makes an onslaught on the "great-man theory" of history, from which a few passages may be quoted:

"The genesis of societies by the action of great man may be comfortably believed so long as, resting in general notions, you do not ask for particulars. But now, if, dissatisfied with vagueness, we demand that our ideas shall be brought into focus and exactly defined, we discover the hypothesis to be utterly incoherent. If, not stopping at the explanation of social progress as due to the great man, we go back a step and ask, Whence comes the great man? We find that the theory breaks down completely. The question has two conceivable answers: his origin is supernatural, or it is natural. Is his origin supernatural? Then he is a deputy god, and we have theocracy once removed, - or, rather, not removed at all. ... Is this an unacceptable solution? Then the origin of the great man is natural; and immediately this is recognized, he must be classed with all other phenomena in the society that gave him birth as a product of hits antecedents. Along with the whole generation of which he forms a minute part, along with its institutions, language, knowledge, manners, and its multitudinous arts and appliances, he is a **resultant**. ... You must admit that the genesis of the great man depends on the long series of complex influences which has produced the race in which he appears, and the social state into which that race has slowly grown. ... Before he can remake his society, his society must remake him. All those changes of which he is the proximate initiator have their chief causes in the generations he descended from. If there is to be anything like a real explanation of those changes, it must be sought in that aggregate of conditions out of which both he and they have arisen."

Now, it seems to me that there is something which one might almost call impudent in the attempt which Mr. Spencer makes, in the first sentence of this extract, to pin the reproach of vagueness upon those who believe in the power of initiative of the great man.

Suppose I say that the singular moderation which now distinguishes social, political and religious discussion in England, and contrasts so strongly with the bigotry and dogmatism of sixty years ago, is largely due to J. S. Mill's example. I may possibly be wrong about the facts; but I am, at any rate, "asking for particulars," and not "resting in general notions." And if Mr. Spencer should tell me it started from no personal influence whatever, but from the "aggregate of conditions," the "generations," Mill and all his contemporaries "descended from," the whole past order of nature in short, surely he, not I, would be the person "satisfied with vagueness."

The fact is that Mr. Spencer's sociological method is identical with that of one who would invoke the zodiac to account for the fall of the sparrow, and the thirteen at table to explain the

gentleman's death. It is of little more scientific value than the Oriental method of replying to whatever question arises by the unimpeachable truism, "God is great." *Not* to fall back on the gods, where a proximate principle may be found, has with us Westerners long since become the sign of an efficient as distinguished from an inefficient intellect.

To believe that the cause of everything is to be found in its antecedents is the starting-point, the initial postulate, not the goal and consummation, of science. If she is simply to lead us out of the labyrinth by the same whole we went in by three or four thousand years ago, it seems hardly worthwhile to have followed her through the darkness at all. If anything is humanly certain it is that the great man's society, properly so called, does *not* make him before he can remake it. Physiological forces, with which the social, political, geographical, and to a great extent anthropological conditions have just as much and just as little do as conditions of the crater of Vesuvius has to do with the flickering of this gas by which I write, are what make him. Can it be that Mr. Spencer holds the convergence of sociological pressures to have so impinged on Stratford-upon-Avon about the 26th of April, 1564, that a W. Shakespeare, with all his mental peculiarities, had to be born there, - as the pressure of water outside a certain boat will cause a stream of a certain form to ooze into a particular leak? And does he mean to say that if the aforesaid W. Shakespeare had died of cholera infantum, another mother at Stratford-upon-Avon would needs have engendered a duplicate copy of him, to restore the sociologic equilibrium? Or might the substitute arise at "Stratford-atte-Bowe"? Here, as elsewhere, it is very hard, in the midst of Mr. Spencer's vagueness, to tell what he does mean at all.

We have, however, in his disciple, Mr. Grant Allen, one who leaves us in no doubt whatever of his precise meaning. This widely informed, suggestive, and brilliant writer published last year a couple of articles in the *Gentleman's Magazine*, in which he maintained that individuals have no initiative in determining social change.

"The differences between one nation and another, whether in intellect, commerce, art, morals, or general temperament, ultimately depend, not upon any mysterious properties of race, nationality, or any other unknown and unintelligible abstractions, but simply and solely upon the physical circumstances to which they are exposed. If it be a fact, as we know it to be, that the French nation differs recognizably from the Chinese, and the people of Hamburg differ recognizably from the people of Timbuktu, then the notorious and conspicuous differences between them are wholly due to the geographical position of the various races. If the people who went to Hamburg had gone to Timbuktu, they would now be indistinguishable from the semi-barbarian negroes who inhabit that central African metropolis; and if the people who went to Timbuktu had gone to Hamburg, they would now have been white-skinned merchants driving a roaring trade in imitation sherry and indigestible port. ... The differentiating agency must be sought in the great permanent geographical features of land and sea; ... these have necessarily and inevitably molded the characters and histories of every nation upon the earth. ... We cannot regard any nation as an active agent in differentiating itself. Only the surrounding circumstances can have any effect in such a direction. [These two sentences dogmatically deny the existence of the relatively independent physiological cycle of causation. WJ]. To suppose otherwise is to suppose that the mind of man is exempt from the universal law of causation. There is no caprice, no spontaneous impulse, in human endeavors. Even tastes and inclinations *must* themselves be the result of surrounding causes."

Elsewhere Mr. Allen, writing of the Greek culture, says:

"It was absolutely and unreservedly the product of the geographical Hellas, acting upon the given factor of the undifferentiated Aryan brain. ... To me it seems a self-evident proposition that nothing whatsoever can differentiate one body of men from another, except the physical conditions in which they are set, - including, of course, under the term **physical conditions** the relations of place and time in which they stand with regard to other bodies of men. To suppose otherwise is to deny the primordial law of causation. To imagine that the mind can differentiate itself is to imagine that it can be differentiated without a cause. {9}

This outcry about the law of universal causation being undone, the moment we refuse to invest in the kind of causation which is peddled round by a particular school, makes one impatient. These writers have no imagination of alternatives. With them there is no **tertium quid** between outward environment and miracle. *Aut Cæsar, aut nullus! Aut* Spencerism, *aut* catechism!

If by "physical conditions" Mr. Allen means what he does mean, the outward cycle of visible nature and man, his assertion is simply physiologically false. For a national mind differentiates "itself" whenever a genius is born in its midst by causes acting in the invisible and molecular cycle. But if Mr. Allen means by "physical conditions" the whole of nature, his assertion, though true, forms but the vague Asiatic profession of belief in an all-enveloping fate, which certainly need not plume itself on any specially advanced or scientific character.

And how can a thinker so clever as Mr. Allen fail to have distinguished in these matters between **necessary** conditions and **sufficient** conditions of a given result? The French say that to have an omelet we must break our eggs; that is, the breaking of eggs is a necessary condition of the omelet. But is it a sufficient condition? Does an omelet appear whenever three eggs are broken? So of the Greek mind. To get such versatile intelligence it may be that such commercial dealings with the world as the geographical Hellas afforded are a necessary condition. But if they are a sufficient condition, why did not the Phoenicians outstrip the Greeks in intelligence? No geographical environment can produce a given type of mind. It can only foster and further certain types fortuitously produced and thwart and frustrate others. Once again, its function is simply selective, and determines what shall actually be only by destroying what is positively incompatible. An Arctic environment is incompatible with improvident habits in its denizens; but whether the inhabitants of such a region shall unite with their thrift the peacefulness of the Eskimo or the pugnacity of the Norseman is, so far as the climate is concerned, an accident. Evolutionists should not forget that we all have five fingers not because four or six would not do just as well, but merely because the first vertebrate above the fishes **happened** to have that number. He owed his prodigious success in founding a line of descent to some entirely other quality, - we know not which, - but the inessential five fingers were taken in tow and preserved to the present day. So of most social peculiarities. Which of them shall be taken in tow by the few qualities which the environment necessarily exacts is a matter of what physiological accidents shall happen among individuals. Mr. Allen promises to prove his thesis in detail by the examples of China, India, England, Rome, etc. I have not the smallest hesitation in predicting that he will do no more with these examples than he has done with Hellas. He will appear upon the scene after the fact, and show that the quality developed by each race was, naturally enough,

not incompatible with its habitat. But he will utterly fail to show that the particular form of compatibility fallen into in each case was the one necessary and only possible form.

Naturalists know well enough how indeterminate the harmonies between a fauna and its environment are. An animal may better his chances of existence in either of many ways, - growing aquatic, arboreal, or subterranean; small and swift, or massive and bulky; spiny, horny, slimy, or venomous; more timid or more pugnacious; more cunning or more fertile of offspring; more gregarious or more solitary; or in other ways besides, - and any one of these ways may suit him to many widely different environments.

Readers of Mr. A. R. Wallace will well remember the striking illustration of this in his *Malay Archipelago*:

"Borneo closely resembles New Guinea not only in its vast size and its freedom from volcanoes, but in its variety of geological structure, its uniformity of climate, and the general aspect of the forest vegetation that clothes its surface; the Moluccas are the counterpart of the Philippines in their volcanic structure, their extreme fertility, their luxuriant forests, and their frequent earthquakes; and Bali, with the east end of Java, has a climate almost as dry and a soil almost as arid as that of Timor. Yet between these corresponding groups of islands, constructed, as it were, after the same pattern, subjected to the same climate, and bathed by the same oceans, there exists the greatest possible contrast when we compare their animal productions. Nowhere does the ancient doctrine that differences or similarities in the various forms of life that inhabit different countries are due to corresponding physical differences or similarities in the countries themselves, meet with so direct and palpable a contradiction. Borneo and New Guinea, as alike physically as two distinct countries can be, are zoologically wide as the poles asunder; while Australia, with its dry winds, its open plains, its stony deserts, and its temperate climate, yet produces birds and quadrupeds which are closely related to those inhabiting the hot, damp, luxuriant forests which everywhere clothe the plains and mountains of New Guinea."

Here we have similar physical-geographical environments harmonizing with widely differing animal lives, and similar animal lives harmonizing with widely differing geographical environments. A singularly accomplished writer, E. Gryzanowski, in the *North America Review*, uses the instance of Sardinia and Corsica in support of this thesis with great effect. He says: "These sister islands, lying in the very center of the Mediterranean, at almost equal distances from the centers of Latin and Neo-Latin civilization, within easy reach of the Phoenician, the Greek, and the Saracen, with a coast-line of more than a thousand miles, endowed with obvious and tempting advantages, and hiding untold sources of agricultural and mineral wealth, have nevertheless remained unknown, unheeded, and certainly uncared for during the thirty centuries of European history. ... These islands have dialects, but no language; records of battles, but no history. They have customs, but no laws; the *vendetta*, but no justice. They have wants and wealth, but no commerce; timber and ports, but no shipping. They have legends, but no poetry; beauty, but no art; and twenty years ago it could still be said that they had universities, but no students. ... That Sardinia, with all her emotional and picturesque barbarism, has never produced a single artist is almost as strange as her barbarism itself. ... Near the focus of European civilization, in the very spot which an *à priori* geographer would point out as the most favorable

place for material and intellectual, commercial and political development, these strange sister islands have slept their secular sleep, like *nodes* on the sounding-board of history."

This writer then goes on to compare Sardinia and Sicily with some detail. All the material advantages are in favor of Sardinia, "and the Sardinian population, being of an ancestry more mixed than [even] that of the English race, would justify far higher expectations than that of Sicily." Yet Sicily's past history has been brilliant in the extreme, and her commerce to-day is great. Dr. Gryzanowski [sic] has his own theory of the historic torpor of these favored isles. He thinks they stagnated because they never gained political autonomy, being always owned by some Continental power. I will not dispute the theory; but I will ask: Why did they not gain it? and answer immediately: Simply because no individuals were born there with patriotism and ability enough to inflame their countrymen with national pride, ambition, and thirst for independent life. Corsicans and Sardinians are probably as good stuff as any of their neighbors. But the best wood-pile will not blaze till a torch is applied, and appropriate torches seem to have been wanting.

Sporadic great men come everywhere. But for a community to get vibrating through and through with intensely active life, many geniuses coming together and in rapid succession are required. This is why great epochs are so rare, - why the sudden bloom of a Greece, an early Rome, a Renaissance, is such a mystery. Blow must follow blow so fast that no cooling can occur in the intervals. Then the mass of the nation glows incandescent, and may continue to glow by pure inertia long after the originators of its internal movement have passed away. We often hear surprise expressed that in these high tides of human affairs not only the people should be filled with stronger life, but that individual geniuses should seem so exceptionally abundant. This mystery is just about as deep as the time-honored conundrum as to why great rivers flow by great towns. It is true that great public fermentations awaken and adopt many geniuses, who in more torpid times would have had no chance to work. But over and above this there must be an exceptional concourse of genius about a time, to make the fermentation begin at all. The unlikeliness of the concourse is far greater than the unlikeliness of any particular genius; hence the rarity of these periods and the exceptional aspect which they always wear.

It is folly, then, to speak of the "laws of history" as of something inevitable, which science has only to discover, and whose consequences any one can then foretell but do nothing to alter or avert. Why, the very laws of physics are conditional, and deal with *ifs*. The physicist does not say, "The water will boil anyhow"; he only says it will boil if a fire is kindled beneath it. And so the utmost the student of sociology can ever predict is that *if* a genius of a certain sort show the way, society will be sure to follow. It might long ago have been predicted with great confidence that both Italy and Germany would reach a stable unity if someone could but succeed in starting the process. It could not have been predicted, however, that the *modus operandi* in each case would be subordination to a paramount state rather than federation, because no historian could have calculated the freaks of birth and fortune which gave at the same moment such positions of authority to three such peculiar individuals as Napoleon III, Bismarck, and Cavour. So of our own politics. It is certain now that the movement of the independents, reformers, or whatever one pleases to call them, will triumph. But whether to do so by converting the Republican party to its ends, or by rearing a new party on the ruins of both our present factions, the historian cannot say. There can be no doubt that the reform movement would make more progress in one year with an

adequate personal leader than as now in ten without one. Were there a great citizen, splendid with every civic gift, to be its candidate, who can doubt that he would lead us to victory? But, at present, we, his environment, who sigh for him and would so gladly preserve and adopt him if he came, can neither move without him, nor yet do anything to bring him forth.

To conclude: The evolutionary view of history, when it denies the vital importance of individual initiative, is, then, an utterly vague and unscientific conception, a lapse from modern scientific determinism into the most ancient oriental fatalism. The lesson of the analysis that we have made (even on the completely deterministic hypothesis with which we started) forms an appeal of the most stimulating sort to the energy of the individual. Even the dogged resistance of the reactionary conservative to changes which he cannot hope entirely to defeat is justified and shown to be effective. He retards the movement; deflects it a little by the concessions he extracts; gives it a resultant momentum, compounded of his inertia and his adversaries' speed; and keeps up, in short, a constant lateral pressure, which, to be sure, never heads it round about, but brings it up at last at a goal far to the right or left of that to which it would have drifted had he allowed it to drift alone.

I now pass to the last division of my subject, the function of the environment in *mental* evolution. After what I have already said, I may be quite concise. Here, if anywhere, it would seem at first sight as if that school must be right which makes the mind passively plastic, and the environment actively productive of the form and order of its conceptions; which, in a word, thinks that all mental progress must result from a series of adaptive changes, in the sense already defined of that word. We know what a vast part of our mental furniture consists of purely remembered, not reasoned, experience. The entire field of our habits and associations by contiguity belongs here. The entire field of those abstract conceptions which were taught us with the language into which we were born belongs here also. And, more than this, there is reason to think that the order of "outer relations" experienced by the individual may itself determine the order in which the general characters imbedded therein shall be noticed and extracted by his mind. The pleasures and benefits, moreover, which certain parts of the environment yield, and the pains and hurts which other parts inflict, determine the direction of our interest and our attention, and so decide at which points the accumulation of mental experiences shall begin. It might, accordingly, seem as if there were no room for any agency other than this; as if the distinction we have found so useful between "spontaneous variation," as the producer of changed forms, and the environment, as their preserver and destroyer, did not hold in the case of mental progress; as if, in a word, the parallel with Darwinism might no longer obtain, and Spencer might be quite right with his fundamental law of intelligence, which says, "The cohesion between psychical states is proportionate to the frequency with which the relation between the answering external phenomena has been repeated in experience."

But, in spite of all these facts, I have no hesitation whatever in holding firm to the Darwinian distinction even here. I maintain that the facts in question are all drawn from the lower strata of the mind, so to speak, - from the sphere of its least evolved functions, from the region of intelligence which man possesses in common with the brutes. And I can easily show that throughout the whole extent of those mental departments which are highest, which are most characteristically human, Spencer's law is violated at every step; and that as a matter of fact the new conceptions, emotions, and active tendencies which evolve are originally produced in the

shape of random images, fancies, accidental out-births of spontaneous variation in the functional activity of the excessively unstable human brain, which the outer environment simply confirms or refutes, adopts or rejects, preserves or destroys, - selects, in short, just as it selects morphological and social variations due to molecular accidents of an analogous sort.

It is one of the tritest truisms that human intelligences of a simple order are very literal. They are slaves of habit, doing what they have been taught without variation; dry, prosaic, and matter-of-fact in their remarks; devoid of humor, except of the coarse physical kind which rejoices in a practical joke; taking the world for granted; and possessing in their faithfulness and honesty the single gift by which they are sometimes able to warm us into admiration. But even this faithfulness seems to have a sort of inorganic ring, and to remind us more of the immutable properties of a piece of inanimate matter than of the steadfastness of a human will capable of alternative choice. When we descend to the brutes, all these peculiarities are intensified. No reader of Schopenhauer can forget his frequent allusions to the *trockener ernst* of dogs and horses, nor to their *ehrlichkeit*. And every noticer of their ways must receive a deep impression of the fatally literal character of the few, simple, and treadmill-like operations of their minds.

But turn to the highest order of minds, and what a change! Instead of thoughts of concrete things patiently following one another in a beaten track of habitual suggestion, we have the most abrupt cross-cuts and transitions from one idea to another, the most rarefied abstractions and discriminations, the most unheard-of combinations of elements, the subtlest associations of analogy; in a word, we seem suddenly introduced into a seething caldron of ideas, where everything is fizzling and bobbing about in a state of bewildering activity, where partnerships can be joined or loosened in an instant, treadmill routine is unknown, and the unexpected seems the only law. According to the idiosyncrasy of the individual, the scintillations will have one character or another. They will be sallies of wit and humor; they will be flashes of poetry and eloquence; they will be constructions of dramatic fiction or of mechanical devices, logical or philosophic abstractions, business projects, or scientific hypotheses, with trains of experimental consequences based thereon; they will be musical sounds, or images of plastic beauty or picturesqueness, or visions of moral harmony. But, whatever their differences may be, they will all agree in this, - that their genesis is sudden and, as it were, spontaneous. That is to say, the same premises would not, in the mind of another individual, have engendered just that conclusion; although, when the conclusion is offered to the other individual, he may thoroughly accept and enjoy it, and envy the brilliancy of him to whom it first occurred.

To Professor Jevons is due the great credit of having emphatically pointed out how the genius of discovery depends altogether on the number of these random notions and guesses which visit the investigator's mind. To be fertile in hypotheses is the first requisite, and to be willing to throw them away the moment experience contradicts them is the next. The Baconian method of collating tables of instance may be a useful aid at certain times. But one might as well expect a chemist's note-book to write down the name of the body analyzed, or a weather table to sum itself up into a prediction of probabilities of its own accord, as to hope that the mere fact of mental confrontation with a certain series of facts will be sufficient to make *any* brain conceive their law. The conceiving of the law is a spontaneous variation in the strictest sense of the term. It flashes out of one brain, and no other, because the instability of that brain is such as to tip and upset itself in just that particular direction. But the important thing to notice is that the good

flashes and the bad flashes, the triumphant hypotheses and the absurd conceits, are on an exact equality in respect of their origin. Aristotle's absurd Physics and his immortal Logic flow from one source: the forces that produce the one produce the other.

When walking along the street, thinking of the blue sky or the fine spring weather, I may either smile at some grotesque whim which occurs to me, or I may suddenly catch an intuition of the solution of a long-unsolved problem, which at that moment was far from my thoughts. Both notions are shaken out of the same reservoir, - the reservoir of a brain in which the reproduction of images in the relations of their outward persistence or frequency has long ceased to be the dominant law. But to the thought, when it is once engendered, the consecration of agreement with outward relations may come. The conceit perishes in a moment, and is forgotten. The scientific hypothesis arouses in me a fever of desire for verification. I read, write, experiment, and consult experts. Everything corroborates my notion, which being then published in a book spreads from review to review and from mouth to mouth, till at last there is no doubt I am enshrined in the Pantheon of the great diviners of nature's ways. The environment *preserves* the conception which it was unable to *produce* in any brain less idiosyncratic than my own.

Now, the spontaneous upsettings of brains this way and that at particular moments into particular ideas and combinations are matched by their equally spontaneous permanent tiltings or saggings towards determinate directions. The humorous bent is quite characteristic; the sentimental one equally so. And the personal tone of each mind, which makes it more alive to certain impressions, more open to certain reasons, is equally the result of that invisible and imaginable play of the forces of growth within the nervous system which, irresponsibly [sic; "irresponsive"?] to the environment, makes the brain peculiarly apt to function in a certain way. Here again the selection goes on. The products of the mind with the determined esthetic bent please or displease the community. We adopt Wordsworth, and grow unsentimental and serene. We are fascinated by Schopenhauer, and learn from him the true luxury of woe. The adopted bent becomes a ferment in the community, and alters its tone. The alteration may be a benefit or a misfortune, for it is (*pace* Mr. Allen) a differentiation from within, which has to run the gauntlet of the larger environment's selective power. Civilized Languedoc, taking the tone of its scholars, poets, princes, and theologians, fell a prey to its rude Catholic environment in the Albigensian crusade. France in 1792, taking the tone of its St. Justs and Marats, plunged into its long career of unstable outward relations. Prussia in 1806, taking the tone of its Humboldts and its Steins, proved itself in the most signal way "adjusted" to its environment in 1872.

Mr. Spencer, in one of the strangest chapters of his *Psychology*, tries to show the necessary order in which the development of conceptions in the human race occurs. No abstract conception can be developed, according to him, until the outward experiences have reached a certain degree of heterogeneity, definiteness, coherence, and so forth.

"Thus the belief in an unchanging order, the belief in *law*, is a belief of which the primitive man is absolutely incapable. ... Experiences such as he receives furnish but few data for the conception of uniformity, whether as displayed in things or in relations. ... The daily impressions which the savage gets yield the notion very imperfectly, and in but few cases. Of all the objects around - trees, stones, hills, pieces of water, clouds and so forth, - most differ widely, ... and few approach complete likeness so nearly as to make discrimination difficult. Even between animals

of the same species it rarely happens that, whether alive or dead, they are presented in just the same attitudes. ... It is only along with a gradual development of the arts ... that there come frequent experiences of perfectly straight lines admitting of complete apposition, bringing the perceptions of equality and inequality. Still more devoid is savage life of the experiences which generate the conception of the uniformity of succession. The sequences observed from hour to hour and day to day seems anything but uniform; difference is a far more conspicuous trait among them. ... So that if we contemplate primitive human life as a whole, we see that multiformity of sequence, rather than uniformity, is the notion which it tends to generate. ... Only as fast as the practice of the arts develops the idea of measure can the consciousness of uniformity simultaneously make possible the notion of *exactness*. ... Hence the primitive man has little experience which cultivates the consciousness of what we call *truth*. How closely allied this is to the consciousness which the practice of the arts cultivates is implied even in language. We speak of a true surface as well as a true statement. Exactness describes perfection in a mechanical fit, as well as perfect agreement between the results of calculations."

The whole burden of Mr. Spencer's book is to show the fatal way in which the mind, supposed passive, is molded by its experiences of "outer relations." In this chapter the yard-stick, the balance, the chronometer, and other machines and instruments come to figure among the "relations" external to the mind. Surely they are so, after they have been manufactured; but only because of the preservative power of the social environment. Originally all these things and all other institutions were flashes of genius in an individual head, of which the outer environment showed no sign. Adopted by the race and become its heritage, they then supply instigations to new geniuses whom they environ to make new inventions and discoveries; and so the ball of progress rolls. But take out the geniuses, or alter their idiosyncrasies, and what increasing uniformities will the environment show? We defy Mr. Spencer or any one else to reply.

The plain truth is that the "philosophy" of evolution (as distinguished from our special information about particular cases of change) is a metaphysical creed, and nothing else. It is a mood of contemplation, an emotional attitude, rather than a system of thought, - a mood which is as old as the world, and which no refutation of any one incarnation of it (such as the Spenserian philosophy) will dispel; the mood of fatalistic pantheism, with its intuition of the One and All, which was, and is, and ever shall be, and from whose womb each single thing proceeds. Far be it from us to speak slightly here of so hoary and mighty a style of looking on the world as this. What we at present call scientific discoveries had nothing to do with bringing it to birth, nor can one easily conceive that they should even give it its *quietus*, no matter how logically incompatible with its spirit the ultimate phenomenal distinctions which science accumulates should turn out to be. It can laugh at the phenomenal distinctions on which science is based, for it draws its vital breath from a region which - whether above or below - is at least altogether different from that in which science dwells. A critic, however, who cannot disprove the truth of the metaphysic creed, can at least raise his voice in protest against its disguising itself in "scientific" plumes. I think that all who have had the patience to follow me thus far will agree that the Spenserian "philosophy" of social and intellectual progress is an obsolete anachronism, reverting to a pre-Darwinian type of thought, just as the Spenserian philosophy of "Force," effacing all the previous distinctions between actual and potential energy, momentum, work, force, mass, etc., which physicists have with so much agony achieved, carries us back to a pre-Galilean age.

# **REPLICATING SONORITIES: TOWARDS A MEMETICS OF MUSIC**

**By**

**Steven Jan**

Department of Music

University of Huddersfield

Queensgate, Huddersfield HD1 3DH, UK.

[s.b.jan@hud.ac.uk](mailto:s.b.jan@hud.ac.uk)

## **Abstract**

The memetic paradigm is herein applied to music. While memetics has been used to elucidate a wide variety of cultural phenomena, its concerns to date have largely been with memes in the realm of verbally-expressible concepts. In view of this, this paper represents an attempt to integrate the central concerns of analytical musicology with a neo-Darwinian meme-selectionist perspective. Such a viewpoint may be used, it is argued, to unify, under a systematic new paradigm, understanding of both local issues of musical structure and organization, and global issues of musical style configuration and its diachronic change. Against the grain of several suggestions in the memetics literature, a minimalist view of the musical meme is taken, seeing it as consisting, at the lower extreme, of configurations of as few as three or four notes. The hierarchic location of musical memes is a central concern here, both in cultural hierarchies - i.e., the replication of patterning at different strata within a culture - and in structural hierarchies - i.e., the replication of patterning at different strata within a work, including the level of the global structural archetype. Leonard Meyer's perspective on culture is employed to frame consideration of the first phenomenon, whilst the analytical method of Heinrich Schenker is employed to comprehend the second. In order to understand how musical memes partake of meaning in association with verbally-mediated concept memes, the semiology of Ferdinand de Saussure and Jean-Jacques Nattiez is employed. The article concludes with observations on the transmission and mutation of musical memes, and an account of how this process engenders the evolution of musical styles.

**Keywords:** Meme, memetics, memplex, music, musicology, style, hierarchy, Meyer, Schenker, semiology.

## **1. Introduction: Music and the Memetic Paradigm**

In the relatively short time since its inception, contributors to the development of the memetic paradigm, including those writing in this journal, have made valuable progress in the development of memetics as, in Brodie's words, a "...long-awaited scientific theory unifying biology, psychology, and cognitive science..." (1996: 13). Understandably, the debate has largely concentrated on memes in the verbal-conceptual realm, exploring such questions as the growth and spread of religion (Dawkins 1993), the development of sexual taboos (Lynch 1996), and the evolution of information on the Internet (Pocklington and Best 1997).

Nevertheless, Dawkins' now celebrated original definition of the meme was catholic, and indeed began with music, maintaining that "[e]xamples of memes are tunes, ideas, catch-phrases, clothes fashions, ways of making pots or of building arches" (1989: 192). One must therefore remember that our modern view of evolution is, in Dennett's words, *substrate neutral*, meaning that, in the algorithm of evolution by natural selection, "[t]he power of the procedure is due to its *logical* structure, not the causal powers of the materials used in the instantiation, just so long as those causal powers permit the prescribed steps to be followed exactly" (1995: 50-51; his emphasis). [Note 1] This principle is the intellectual foundation of the 'universal Darwinism' Dawkins outlined (1983b), which Plotkin has recently expanded upon (1995), and of which memetics is an important subset.

Given its broad applicability, and observing a niche, it is to music that I aim to apply the memetic paradigm here, attempting to understand the central concerns of musicology [note 2] from a neo-Darwinian meme-selectionist perspective in order to show how such a standpoint may be used to unify understanding of both local issues of musical structure and organization, and global issues of musical style configuration and its diachronic change. [Note 3] Whilst I will largely use as my examples music written in the late-eighteenth and early-nineteenth centuries - a period of such stylistic clarity and coherence the epithet *classical* is deservedly applied to it by musicologists - many of the principles I adduce here are, I believe, applicable to all musics. [Note 4] Moreover, I hope that some observations, given their necessarily contextualizing purpose, make a more general contribution to the debate on memetics, particularly some of the content of Sections 3, 4, 7, and 8. What follows is, however, only a brief outline of what I believe to be the central issues, and I hope to develop the points made here on other occasions.

I begin by examining briefly the relationship between memes in music and those in the verbal-conceptual realm in terms of the linguistic concepts of phonetics, syntax, and semantics. I then consider how memetics offers a means of understanding the characteristics of style in music and, indeed, in culture generally. After exploring the application to musical memes of the genotype-phenotype distinction from biology, I move on to the central parts of the article, which define what I understand as the memetic elements in music, including the nature and hierarchic location of memplexes in music. The final sections of the article deal with the transmission and mutation of musical memes and the way in which this process engenders the evolution of musical style.

During the course of this article it will be necessary to use concepts and terms from music theory and analysis which may not be familiar to the reader. When these are used, notes provide what I hope are clear and non-patronizing explanations. For readers wishing to explore these concepts more fully, Bent and Drabkin 1987, Cook 1994, and Dunsby and Whittall 1988 are excellent sources. For more general information, Sadie 1980 is the most comprehensive reference source in English.

## 2. Musical Memetics and the Linguistic Analogy

One of the central issues encountered in developing a memetic theory of cultural evolution in analogy with the genetic theory of biological evolution is the fact that  
...what is preserved and transmitted in cultural evolution is *information* - in a media-neutral, language-neutral sense. Thus the meme is primarily a *semantic* classification, not a *syntactic*

classification that might be directly observable in "brain language" or natural language. In the case of genes, we are blessed by a gratifyingly strong alignment of semantic and syntactic identity: there is a single genetic language, in which meaning [i.e., what phenotypic feature(s) the gene codes "for"] is (roughly) preserved across all species.

(Dennett 1995: 353-354; his emphases)

Nevertheless, even though there is a long critical tradition [note 5] of conceiving music as in some sense akin to verbal language - with, accordingly, a distinct phonetic, syntactic, *and* semantic content - in reality music, unlike primarily referential symbolic systems such as language and painting, has the property of having no fixed semantic structure. To use the language of semiology, referential symbolic systems unite a signifier - a word or image - and a signified - an idea or concept. As Nattiez observes,

By making the sign a union of the signifier and the signified, Saussure conceived of the relationship between the two "faces" of the sign as stable and bi-univocal. Beyond this, the relationship is *arbitrary* [.]

(1990: 4; his emphasis)

In music, by contrast, a given configuration - a collection of pitches and rhythms, for instance - does not act as the stable *signifier* for a particular verbally-expressible *signified*, despite the propensity of music to exhibit to a degree of memetic coadaptation which permits specific musical patterns to become associated, in varying degrees of stability, with verbally expressible concepts. This question of music and the signifier-signified dualism - music as signified and music as signifier - is considered in more detail in Sections 6.3 and 6.4 respectively.

In this sense, music's semantic content is low, and its real substance could be said to inhere in its phonetic and syntactic structures. The phonetic presence of music is, as will be discussed in Section 4, a purely phenotypic property. Its syntax, however, whilst reflected phenotypically in, and only readily apparent from, the configuration of musico-graphical patterning in scores, and patterns of sound waves in performance, must presumably inhere at the level of the memotype, i.e., in the form of ordered, grammatical relationships between memes at the level of the neuron. Nevertheless, as Dennett implies, it seems the case that there is no "gratifyingly strong alignment" between neuronal configurations and the musico-syntactic structures they code for.

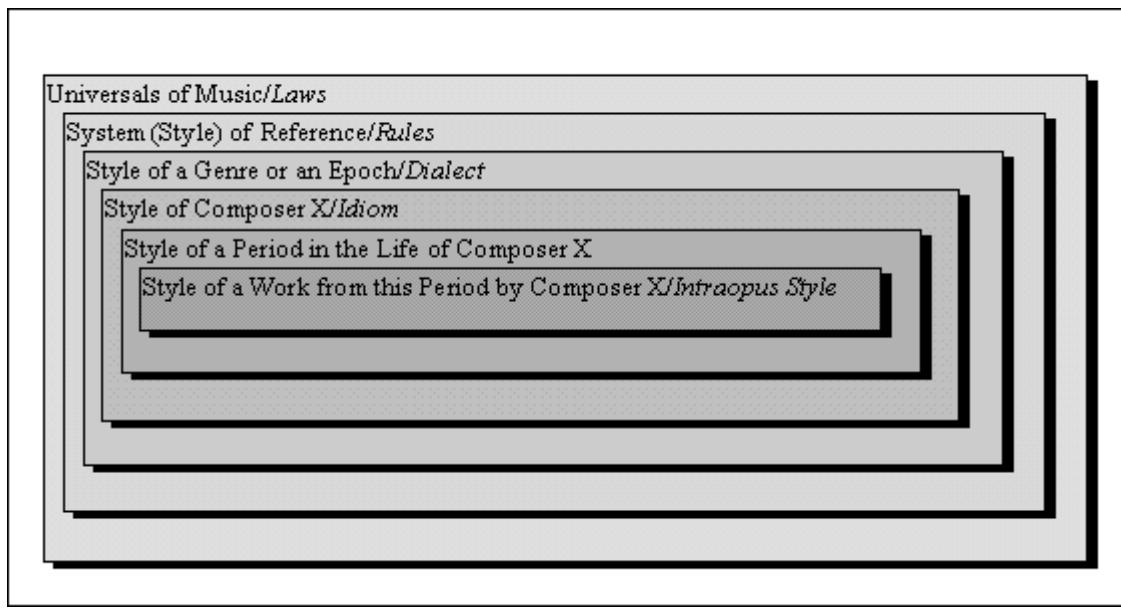
In contrast to the semantic aspect, the phonetic and syntactic realms are highly developed in the musics of all cultures, the sonorous (phonetic) field consisting of particles which have a combinatorial (syntactic) structure determined not by the constraints of verbal-conceptual logic but by the force of convention. In memetic terms, conventions arise from the differential replication of patterns. The most successful of these become to some degree *evolutionarily stable*, a concept examined in more detail in Section 8. Such patterns - Meyer terms them *strategies* played out within systems of *rules* (1989: 17, 20) - tend to oust rival configurations, which theorists, certainly in the western tradition, then belatedly condemn by deeming them aberrant.

As I hope will become evident from the following discussion, the rich phonetic and syntactic structure of music, together with its relative independence from the complexities engendered by

the rich semantic content of verbal-conceptual memes, makes it an especially lucid medium for memetic analysis. Although the discussion of the memotype-phemotype dualism in Section 4 will indicate this is to be an oversimplification, I shall, in conducting this analysis, take the phonetic and, perhaps more importantly, the syntactic aspects of music to be adequately represented by the musico-graphical content of the printed score.

### 3. Memetics and Musical Style

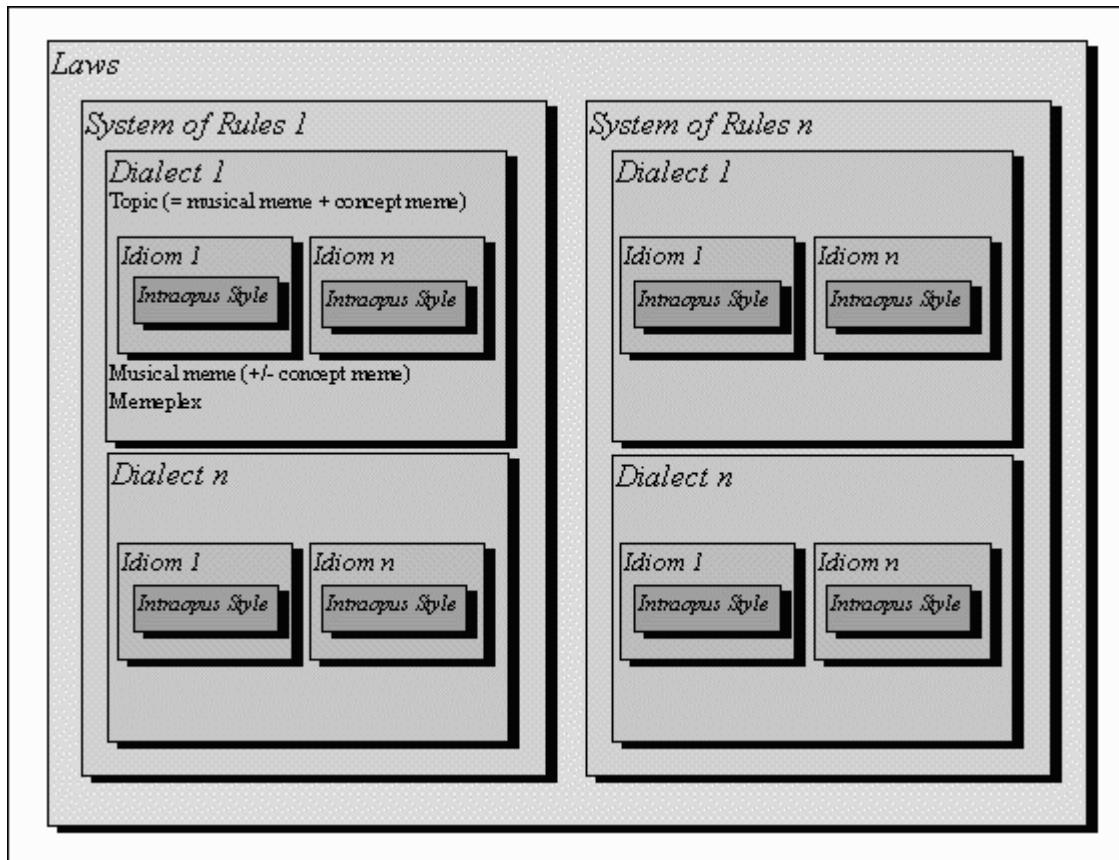
Before examining the nature of memes in music, a brief explanation is needed of the relationship between memetics and style analysis, which is deep-seated if one accepts Meyer's definition of style as "...*a replication of patterning*, whether in human behaviour or in the artifacts produced by human behaviour, that results from a series of choices made within some set of constraints" (1989: 3; my emphasis). Importantly, Meyer and Nattiez note that the pattern replication which defines style occurs at a number of discrete levels in a system of hierachic inclusion. The following diagram illustrates this concept, reworking the pyramidal representation of the notion of inclusive "levels of stylistic relevance" from Nattiez 1990, and adding corresponding terms from Meyer 1989 in italics. Whilst developed from sources discussing musical style, this perspective is applicable to the replication of memes in all cultural realms:



**Figure 1:** The Hierarchy of Style (after Nattiez 1990: 136, Fig. 6.2 and Meyer 1989: 13-24; see also Jan 1995: 6-7)

Here, *intraopus style* represents the style of - i.e., the totality of memes replicated within - a single work. *Idiom* represents the style of an individual composer within a particular cultural community, and is defined by the sum of the memes replicated within the totality of his or her intraopus contexts. A composer's idiom may, however, be subdivided into a number of chronologically-defined phases or style periods. [note 6] *Dialect* represents the style of a cultural community, defined in geographical (e.g., French, Viennese, Ghanaian) and/or chronological (e.g., Renaissance, Baroque, Classical) terms. *Rules*, the largest memetic aspect of the hierarchy,

are the large-scale systems of musical organization, such as the modal system [note 7] or the major-minor tonal system, [note 8] the generative elements of which are replicated in a number of separate dialects. Finally, *laws* represent invariable, biologically-defined attributes of human perception and cognition which ultimately constrain our responses to music and therefore set limits upon the nature of the memes we assimilate and the profile of their interaction. The *meme pool* might be taken to consist of the totality of memes at the level of dialect or rules:



**Figure 2:** Stylistic Hierarchies as Meme Pools [note 9]

#### 4. Musical Memes and the Genotype-Phenotype Distinction

Beginning with Dawkins (1983a: 109), many commentators have asserted that memes exist fundamentally as living biological structures, inhering in the patterning of neuronal interconnections. As Delius notes,

Any cultural trait [meme] that is taken over by a given individual from another individual must accordingly be thought of as the transfer of a particular pattern of synaptic hotspots within the associative networks of one brain to the associative networks of another brain. Different traits must be thought of as being coded by topologically different hotspot patterns. That is, a given cultural trait borne by an individual is encoded informationally as a particular pattern of modified synapses in his brain. Naturally the hotspot pattern that a trait has in one brain will not be geometrically arranged in exactly the same way as the pattern that the same trait has in another

brain...But functionally the two hotspot patterns would still be equivalent, at least to the extent that the effective memory contents representing the trait were identical. (1989: 44-45)

Acceptance of this point naturally suggests the application of the *genotype-phenotype* distinction - a differentiation between something hidden and its often far-reaching consequences - to the meme. Whilst care must be taken not to force our thinking on memetics into the straitjacket of earthly DNA, the assimilation of the genotype-phenotype principle is one area where analogizing seems warranted. [Note 10]

#### **4.1. The Memotype and (Extended) Phemotype of Musical Memes**

Of the often contradictory formulations of this analogy - essentially part of a wider debate on the ontology of the meme - surveyed by Blackmore (1999: 63-66), that articulated in Ball 1984 and Delius 1989 seems most compatible with the memetics of music I offer here. Accordingly, I take the *memotype* (Grant, in Blackmore 1999: 64) to be the totality of memes - including those coding for musical functions - resident in an individual's brain, existing physically as patterns of neuronal interconnection. The phenotypic - or, by way of distinction, *phemotypic* (Blackmore 1999: 63) - products of memes may be reduced to the two spheres of somatic and extrasomatic effects. The first of these is negligible and includes culturally-determined alterations to the form of a body holding the meme, such as tattooing and piercing. Much more important are the extrasomatic effects, which include patterns of behaviour and the resultant artifacts they motivate and control. These, in turn, facilitate the propagation of memes by engendering their reconstitution, as functionally - but not necessarily structurally - analogous neuronal configurations in the brains of other humans. A musical meme, for instance, may give rise to the behaviours of composition, improvisation, and performance; in terms of artifacts, the meme engenders representations of various kinds on paper, magnetic and compact discs, and, ultimately, the configurations of sound waves for which these media code.

This distinction is essentially that made several years before *The Selfish Gene* by Cloak (1973), who separated *i-culture*, the somatic cultural instructions, and *m-culture*, their resultant behaviour and artifacts. Whatever terms one uses, this perspective has the advantage of mirroring its biological analogue closely and being simple and clear-cut. Moreover, it loosely preserves the distinction Dawkins draws between *replicators* and *vehicles*, whereby the microscopic, invisible, units of replication are clearly distinguished from the large, visible, structures engendered by the replicants to facilitate replication (1983a: Chapter 6). It should be noted, however, that whereas in genetics, replicators are copied ("copy-the-instructions" - Blackmore 1999: 61-62), in memetics it is the vehicles which are copied ("copy-the-product").

Dawkins' concept of the *extended phenotype* enlarges the definition of the phenotype to incorporate

All effects of a gene upon the world. As always, "effect" of a gene is understood as meaning in comparison with its alleles. The conventional phenotype is the special case in which the effects are regarded as being confined to the individual body in which the gene sits. In practice it is convenient to limit "extended phenotype" to cases where the effects influence the survival chances of the gene, positively or negatively. (1983a: 286)

While the phenotypic effects of the gene are both somatic and extrasomatic, the phenotypic effects of the meme are, as noted, almost entirely extrasomatic; and it is in the nature of the meme that one cannot usefully distinguish between its phenotype and its extended phenotype. To paraphrase Dawkins, one may indeed speak of "the long reach of the meme" (after the subtitle to 1983a), meaning that a meme, being at "...the centre of a web of radiating power" (1983a: viii), may both influence the behaviour of its host and also exert its effects at considerable geographical and temporal remove.

As an example of this wide compass, it is clear that memes which appeared initially in symphonies composed by Haydn in the isolation of the Palace of Eszterháza in Hungary in the 1770s continue to copy themselves into the brains of concert goers on the West coast of America in the 1990s; paper-based phenotypic forms of these memes continue to roll from the printing presses of the developed world; and tons of plastic are poured into moulds to make compact disks sustaining other phenotypic incarnations of them.

To draw these points together, Table 1 compares genes and memes in terms of the two pheno/phenotypic spheres, somatic and extrasomatic, offering examples of each category:

		GENE/PHENOTYPE; MEME/PHENOTYPE		
REPLICATORS	SOMATIC	EXTRASOMATIC		ARTEFACTS
		BEHAVIOUR		
GENE/GENOME	feathers	nest building		nests
MEME/MEMOME	culturally-determined modifications to the body	singing, playing, conducting, improvising, composing		configurations of sound waves, written and printed scores, recordings

**Table 1:** Genetic and Memetic Pheno/Phenotypes (after Ball 1984: 156, Figure 2)

For a bird, for instance, some of its genes give rise to such phenotypic-somatic features as feathers, and to such phenotypic-extrasomatic behaviours as nest building and the resultant artifacts. For a human, some of its musical memes may give rise to culturally determined modifications to the body, such as the irreparable damage Schumann inflicted upon his hand as a result of using a crude device designed to strengthen the fingers (Daverio 1997: 77-79), and to such phenotypic-extrasomatic behaviours as composition, the artifacts of which are written and printed scores.

#### 4.2. The Ontology of the Musical Work

As noted in Section 4.1, the perspective afforded by the memotype-phenotype dualism raises the question of the ontological status of the meme and, by extension, the status of the larger musical work of which it is a component. In Nattiez's view, "...music's irreducible dimension is *sound*. The musical work *manifests itself*, in its *material* reality, in the form of sound waves" (1990: 69;

his emphases). Nevertheless, he acknowledges the philosophical complications which ensue when the work's "mode of *being*" is considered, noting Ingarden's assertion that ...the work is a purely intentional object, immutable and permanent, whose heteronomous *existence* is no more than a reflection of its *being*: the existence of the work finds its source in the "creative act" of the performer, and its foundation in the score....The score constitutes the work's "schema," which guarantees its identity over the course of history, even though numerous elements not fixed by the score "play an essential role for the aesthetic Gestalt of the work"..., and even though the "schema" allows an enormous number of possibilities for the work's realization. (1990: 69-70)

According to Siohan, however, the notation "...is a graphic element, is neither music, nor its reflection, but a solely mnemonic device. There is no music except in the state of sonorous manifestation" (in Nattiez 1990: 71).

In contrast to the score *versus* sound dualism underpinning these statements, the memotype-phemotype perspective indicates that the musical meme's principal state of being is the somatic neuronal configurations which define its information content. Despite being "...the thing that renders the work performable and recognizable as an entity, and enables the work to pass through the centuries" (Nattiez 1990: 71), the score (and its resultant sound image) is, from the meme's perspective, merely a phemotypic "vessel" to convey it on certain stages of the "long odyssey" from brain to brain (after Dawkins 1989: 33). Seen in this way, one arrives not at a sound *versus* score dualism but rather one of neuronal matrix *versus* sound.

Thus, contrary to Nattiez's assertions, one might, in the light of the memotype-phemotype distinction, maintain music's "irreducible dimension" to be somatic memes, and sound waves to be just one of their extended phemotypic artifacts, albeit overwhelmingly the one of greatest importance to most cultures. On this reasoning, the large-scale memeplex that is the musical work exists fundamentally as a temporary alliance of memes in the composer's brain which is subsequently reconstituted in the minds of listeners exposed to its phemotypic products. The freezing of the work by its extended phemotype should not blind us to the fact that the memes that constitute it merely used it as a means of furthering their replication.

Despite this, work-memeplexes have attained a high degree of ontological singularity for us in the West, owing to a crucial memetic event which occurred some two hundred years ago. Goehr, without using the memetic paradigm in her argument, asserts that the closing years of the eighteenth century - not, controversially, the early years of the sixteenth - saw the evolution of a memeplex she terms the *work-concept*. This is an array of memes which cause us ...to see works as objectified expressions of composers that prior to compositional activity did not exist. We do not treat works as objects just made or put together, like tables and chairs, but as original, unique products of a special, creative activity. We assume, further, that the tonal, rhythmic, and instrumental properties of works are constitutive of structurally integrated wholes that are symbolically represented by composers in scores. Once created, we treat works as existing after their creators have died, and whether or not they are performed or listened to at any given time. We treat them as artifacts existing in the public realm, accessible in principle to anyone who cares to listen to them. And when called, finally, to give examples of works, we

usually look to the tradition of western, European, classical, "opus" music, to works, in other words, of a "purely instrumental" or "absolute" sort. (1992: 2)

The propagation of this memeplex in the nineteenth century and afterwards gave work-memeplexes a far greater cultural salience than they had hitherto enjoyed, according them the status of exhibits in the "imaginary museum of musical works" (Goehr 1992: 8). Before the time of Beethoven, pieces of music were generally seen as transient and ephemeral, the utilitarian products of craftsmanship and labor. After Beethoven - and because of the work-concept memeplex - many were seen in terms of the new category of artwork; they came to acquire a degree of permanence and stability, and were supposed the transcendental products of genius and inspiration.

## 5. Defining the Musical Meme

In defining the meme, in any symbolic form, there are two fundamental and interdependent parameters which one has to consider: *particulateness*, the meme's status as a small, discrete, and self-contained entity, and *coequality*, its status as a unit of imitation, the *sine qua non* of memetic identity. The application of these principles to music is considered below.

### 5.1. Particulateness

Copying-fidelity upon replication is one of the defining characteristics of the gene - the others are longevity and fecundity (Dawkins 1989: 18, 194) - which Dawkins considers to be "...any portion of chromosomal material that potentially lasts for enough generations to serve as a unit of natural selection....a genetic unit that is small enough to last for a large number of generations and to be distributed around in the form of many copies....a unit which, to a high degree, approaches the ideal of indivisible particulateness" (1989: 28, 32, 33). As an undifferentiated segment of a chromosome, the longer a gene, the greater the likelihood that it will be divided - its information content destroyed - during the random cuttings and splicings of the chromosome ('crossing over') which occur during reproduction. Dawkins' definition is founded on his observation that the smaller the gene the greater the statistical probability that it will survive the process of crossing over intact. He does not, however, apply this same logic to the meme when discussing problematic aspects of its definition; again, he uses musical memes as his examples:

So far I have talked of memes as though it was obvious what a single unit-meme consisted of. But of course it is far from obvious. I have said a tune is one meme, but what about a symphony: how many memes is that? Is each movement one meme, each recognizable phrase of melody, each bar, each chord, or what? (1989: 195)

The biological analogy, indeed the replicator concept, would seem to suggest that memes tend toward the latter end of the continuum Dawkins outlines - toward existence as simple rather than complex units. Thus, the more concise a musical meme the greater the likelihood that it will survive the travails of replication by imitation intact. Ultimately, one might assert that the meme shares with the gene the fundamental attribute that it is "...a unit of convenience, a length of chromosome with just sufficient copying-fidelity to serve as a viable unit of natural selection" (1989: 195).

Dennett, however, is reluctant to admit very small particles as memetic. His definition of the meme asserts that:

These new replicators are, roughly, ideas....the sort of complex ideas that form themselves into *distinct memorable units*...Intuitively, we see these as more or less identifiable cultural units, but we can say something more precise about how we draw the boundaries - about why D-F#-A isn't a unit and the theme from the slow movement of Beethoven's Seventh Symphony is: the units are the smallest elements that replicate themselves with reliability and fecundity. We can compare them, in this regard, to genes and their components: C-G-A, a single codon of DNA, is "too small" to be a gene. It is one of the codes for the amino acid arginine, and it copies itself prodigiously wherever it appears in genomes, but its effects are not "individual" enough to count as a gene. A three-nucleotide phrase does not count as a gene for the same reason that you can't copyright a three-note musical phrase: it is not enough to make a melody. (Dennett 1995: 344; his emphasis)

Dennett seems to suggest that whilst three notes can't be memetic, four can, for he maintains that The first four notes of Beethoven's Fifth Symphony are clearly a meme, replicating all by themselves, detached from the rest of the symphony, but keeping intact a certain identity of effect (a phenotypic effect), and hence thriving in contexts in which Beethoven and his works are unknown. (Dennett 1995: 344)

Nevertheless, many listeners would recognize the following song, Johnny Burke's and Erroll Garner's 'Misty' of 1955, from its distinctive opening vocal gesture ("Look at me..."), a collection of just three notes, b<sup>b1</sup>-g<sup>1</sup>-d<sup>1</sup>. If the memory of a melody can be accessed by hearing its opening notes, this implies that these notes themselves have independent memetic status: [note 11]

**Example 1:** 'Misty,' bb. 4-7. [note 12]

The issue is, however, not quite as simple as it is presented, for Dennett is blurring the musical parameters of pitch and rhythm. Even the shortest and most self-contained musical meme is seldom an indivisible particle, by which I mean divisible not horizontally (by length), or vertically (by pitch strata), but parametrically. Almost invariably, it is itself a complex of memes in different musical parameters, principally of pitch and rhythmic configurations, but also structures in the realms of dynamics (patterns of loud and soft) and texture (patterns of thick and thin sounds). Example 2 i below, for instance, consists of configurations in pitch - bracketed and labeled meme x - and rhythm - meme y - which, as shown in ii and iii, have independent existence owing to their replication in other contexts:

The image contains three musical examples, labeled i), ii), and iii), each consisting of two staves: treble and bass. 
 Example i) shows two measures of music in common time (indicated by 'C'). Measure 1 starts with a quarter note followed by an eighth-note pair (x). Measure 2 starts with a quarter note followed by a sixteenth-note pair (y). 
 Example ii) shows two measures of music in common time (indicated by 'C'). Measure 1 starts with a dotted half note followed by an eighth-note pair (x). Measure 2 starts with a dotted half note followed by a sixteenth-note pair (y). 
 Example iii) shows four measures of music in common time (indicated by 'C'). Measure 1 starts with a quarter note followed by an eighth-note pair (x). Measure 2 starts with a quarter note followed by an eighth-note pair (y). Measure 3 starts with a quarter note followed by a sixteenth-note pair (x). Measure 4 starts with a quarter note followed by an eighth-note pair (y).

### Example 2: The Meme as A Parametric Complex

- i) Memes  $x$  (pitch) and  $y$  (rhythm): Mozart, Piano Concerto in  $B^\flat$  major K. 456 (1784) I, bb. 1-4.
- ii) Meme  $x$ : Pitch Analogue  $\uparrow 2 \downarrow 3 \uparrow 1$ ; [note 13] Mozart, *Die Zauberflöte* K. 620 (1791) no. 13, "Alles fühlt der Liebe Freuden," bb. 1-3, transposed, for ease of comparison, from C major to  $B^\flat$  major. [note 14]
- iii) Meme  $y$ : Rhythmic Analogue  $\downarrow \overline{J} \overline{J} \downarrow \downarrow$ ; Mozart, Piano Concerto in F major K. 459 (1784) I, bb. 1-4.

I will concentrate in this article exclusively on memes in the parameter of pitch, however, because that which Meyer terms the attribute of *syntacticism* is most clearly defined in the realm of pitch, rhythm possessing this property to a lesser extent. He notes that In order for syntax to exist..., successive stimuli must be related to one another in such a way that specific criteria for mobility and closure are established. Such criteria can be established only if the elements of the parameter can be segmented into discrete, nonuniform relationships so that

the similarities and differences between them are definable, constant, and proportional. (1989: 14)

In the western European tradition, at least, memetic structures have evolved most easily in the pitch and, less so, the rhythmic realms, the parameters of dynamics and texture lending themselves rather less readily to "...segment[ation] into discrete, nonuniform relationships....". Having established the parametric interaction of musical memes, Dennett's example, the theme from the opening of Beethoven's Fifth Symphony, could be understood as not just four "notes" but, fundamentally, two pitches, g<sup>1</sup> and e<sup>#1</sup>, articulated by four rhythmic attacks,  $\gamma \text{ } \text{ } \text{ } \text{ } \text{ }$ . One could even say that the most distinctive aspect of the theme is not the pitch content but this abstract rhythmic profile - bracketed and labeled as meme x - which, later in the first movement, becomes reduced to a *three* attack rhythm -  $\text{ } \text{ } \text{ } \text{ } \text{ }$ , meme y - a memetic unit of wide currency at this time:

The image shows three staves of musical notation from Beethoven's Fifth Symphony, labeled i), ii), and iii).

- i)** Staff 1 (Treble clef) and Staff 2 (Bass clef) in 2/4 time, key signature of C minor (one flat). Measure 1 starts with a single eighth note (pitch g<sup>1</sup>). Measures 2-4 show a rhythmic pattern of two eighth notes followed by a quarter note. Measure 5 shows a single eighth note (pitch e<sup>#1</sup>). Brackets above the staff are labeled 'x'.
- ii)** Staff 1 (Treble clef) and Staff 2 (Bass clef) in 2/4 time, key signature of C minor (one flat). Measure 179 starts with a sixteenth-note burst (pitch g<sup>1</sup>). Measures 180-181 show a rhythmic pattern of two sixteenth-note groups followed by a quarter note. Measures 182-183 show a sixteenth-note burst (pitch e<sup>#1</sup>). Brackets above the staff are labeled 'x' and 'y' respectively.
- iii)** Staff 1 (Treble clef) and Staff 2 (Bass clef) in common time, key signature of C major (no sharps or flats). Measure 147 starts with a sixteenth-note burst (pitch e<sup>#1</sup>). Measures 148-149 show a rhythmic pattern of two sixteenth-note groups followed by a quarter note. Measures 150-151 show a sixteenth-note burst (pitch g<sup>1</sup>). Brackets above the staff are labeled 'y' and 'y' respectively.

**Example 3:** Rhythmic Memes in Beethoven and Mozart

- i) Meme x: Beethoven, Symphony no. 5 in C minor op. 67 (1808) I, bb. 1-5.
- ii) Memes x and y: Beethoven, Symphony no. 5 I, bb. 179-190.

iii) Meme y: Mozart, Symphony no. 40 in G minor K. 550 (1788) I, bb. 146-150. Moreover, Dennett's assertion of the memetic status of the theme from the slow movement of Beethoven's Seventh Symphony - all forty-five notes of it (bb. 27-42) - is also problematic, in that he doesn't distinguish between *passive* listeners and *active* composers here. Although, clearly, "...there is a limit to the complexity of musical phrases that can be memorized" (Gatherer 1997: 81), this particular theme is probably not too complex to be held intact in the memory of many musically literate listeners. If so retained, it would qualify in its entirety as a meme. Of course, it also survives, rather more durably, in the form of countless phenotypic representations.

In the work of the composers who were Beethoven's contemporaries and successors, however, this theme, although also copied in their brains as a single-unit meme, was not imitated in their compositions *en bloc*, for what might be called the *anti-plagiarism* concept meme (prolific in western European culture certainly from the beginning of the nineteenth century) would have inhibited this, making it, in Dawkins' words, far too long to be a "...viable unit of natural selection" (1989: 195). Memorizing and then replicating whole chunks of the music of their predecessors and contemporaries is not what composers do: they tend, as I have shown, to replicate the small component units of the musical continuum - either remembered from musico-aural experience or viewed as musico-graphical symbols on paper - navigating the straits between the siren voices of memorability and the rocks of plagiarism.

On this reasoning, one would expect relatively short segments of the theme from the slow movement of Beethoven's Seventh Symphony to have memetic status. Indeed, as Example 4 illustrates, bb. 27-42 might be regarded as a memeplex, a collection of memes which assemble here into a larger pattern (this notion is taken up in more detail in Section 6). The example shows copies of just two of the memes - bracketed and labeled x and y in the example - which make up this melody: [note 15]

i)

413

**Example 4:** Memes in Beethoven's Seventh Symphony

- i) Meme *x*: Mozart, Piano Concerto in C major K. 467 (1785) I, bb. 413-417.
- ii) Memes *x* and *y*: Beethoven, Symphony no. 7 in A major op. 92 (1812) II, bb. 27-42.
- iii) Meme *y*: Schubert, Piano Quintet in A major D. 667 ('Trout') (1819) I, bb. 1-6.

A further complicating factor in the definition of musical memes is the fact that, discounting all parameters except pitch, one seldom deals with one *stratum* of pitch; rather, in most music, several strata function together, giving rise to harmony, the vertical dimension, and counterpoint, the horizontal dimension. A fragment of music will, typically, consist of a melody line, a bass line, and one or more inner parts, each of which may have its own linear-memetic identity, but which may also form a component of a higher-order meme, a phenomenon discussed in Section 6. Here, for instance, is a three-component meme, a version of the so-called *perfect cadence* (i.e., the harmonic progression V-I, in this case preceded by an altered form of chord II), made up, contrary to Dennett's assertion, of three distinct vertical units (chords) and four replicated horizontal units (melodic lines):

The musical score consists of two parts, i) and ii), each with two staves: soprano (G clef) and basso continuo (F clef). Part i) starts at measure 3, indicated by a box containing the number 3. It features a soprano line with eighth-note chords and a basso continuo line with eighth-note chords. A large bracket groups measures 3 through 5. Part ii) starts at measure 9, indicated by a box containing the number 9. It features a soprano line with eighth-note chords and a basso continuo line with eighth-note chords. A large bracket groups measures 9 through 11.

### Example 5: A Three-Component Cadential Meme

- i) Bach: *Matthäus-Passion* BWV 244 (c. 1729-c. 1744) no. 25, Chorale "Was mein Gott will, das g'scheh' allzeit," bb. 3-5.
- ii) Bach: *Matthäus-Passion* no. 40, Chorale "Bin ich gleich von dir gewichen," bb. 9-11.  
One could even assert the memetic status of the following two-component perfect cadential patterns, for the end-segments, the harmonic progression V-I, are identical, yet are preceded by different penultimate configurations:

The image contains two musical examples, labeled i) and ii), each consisting of two staves of music.

**i)** The top staff is in treble clef and common time (C). It starts with a eighth note followed by a sixteenth note, then a eighth note, a sixteenth note, a eighth note, a sixteenth note, a eighth note, a sixteenth note, and a eighth note. The bottom staff is in bass clef and common time (C). It starts with a eighth note, followed by a sixteenth note, a eighth note, a sixteenth note, a eighth note, a sixteenth note, a eighth note, a sixteenth note, and a eighth note. Both staves end with a vertical bar line and a repeat sign, followed by a colon and a vertical bar line.

**ii)** The top staff is in treble clef and common time (C). It starts with a eighth note, followed by a sixteenth note, a eighth note, a sixteenth note, a eighth note, a sixteenth note, a eighth note, a sixteenth note, and a eighth note. The bottom staff is in bass clef and common time (C). It starts with a eighth note, followed by a sixteenth note, a eighth note, a sixteenth note, a eighth note, a sixteenth note, a eighth note, a sixteenth note, and a eighth note. Both staves end with a vertical bar line and a repeat sign, followed by a colon and a vertical bar line.

### Example 6: A Two-Component Cadential Meme

i) Mozart: Piano Sonata in D major K. 284 (205b) (1775) I, bb. 126-127.

ii) Beethoven: Piano Trio in D major op. 70 no. 1 ('Ghost') (1809) I, bb. 268-270.

Ultimately, however, Dennett concedes that "...there is no 'principled' lower limit on the length of a sequence that might come to be considered a gene or a meme" (1995: 344), despite previously asserting four as the nominal value in discussion of Beethoven's Fifth Symphony. Observing that "...Beethoven is the favorite example for illustrating this problem," Blackmore considers this question of meme size, essentially taking Dennett's "...smallest elements that replicate themselves with reliability and fecundity" as the most secure criterion (1999: 53-56). The concept of *coequality* is offered in Section 5.2 as a means of resolving difficulties of this kind.

Given these various qualifications, in defining the musical meme, one may therefore assert that any discrete musical segment which a composer assimilates from his or her environment may be held to be memetic, and that such a segment will tend to be the union of a *short* melodic fragment, a *simple* chord progression, or a *concise* rhythmic pattern. Thus, to demarcate musical memes one needs, in the words of Bent and Drabkin, "...to break the stream of music into component units (or 'unities' - i.e. units that either cannot be further subdivided or do not need to be because their sub-units never occur independently)" (1987: 96). Whilst offered as a summary

of the working methods of *distributional analysis*, [note 16] this principle both holds for the definition of the meme, and neatly parallels that Dawkins gives for the gene.

## 5.2. Coequality

From a memetic standpoint, however, the subdivision into "unities" takes place by reference to other copies of that unit, identifying that portion of the pitch and/or rhythmic continuum which is replicated. "[A]n item of brain-stored memory" has been termed a *mnemon* by Lynch (1998). For such a particle to be regarded as a meme, a unit of imitation, one must, by definition, isolate the copy/copies - what one might term the coequal/s - of the particle, from which the particle is imitated, or which is/are derived by imitation of the particle, for "[w]hen copied from one brain to another, [a mnemon] becomes a meme" (1998). It will be understood that particularity, the segmentation of the symbolic stream, can only be defined by reference to coequality, the presence of analogous segments of the same or another symbolic stream, and not by means of potentially deceptive surface articulations in the symbolic stream.

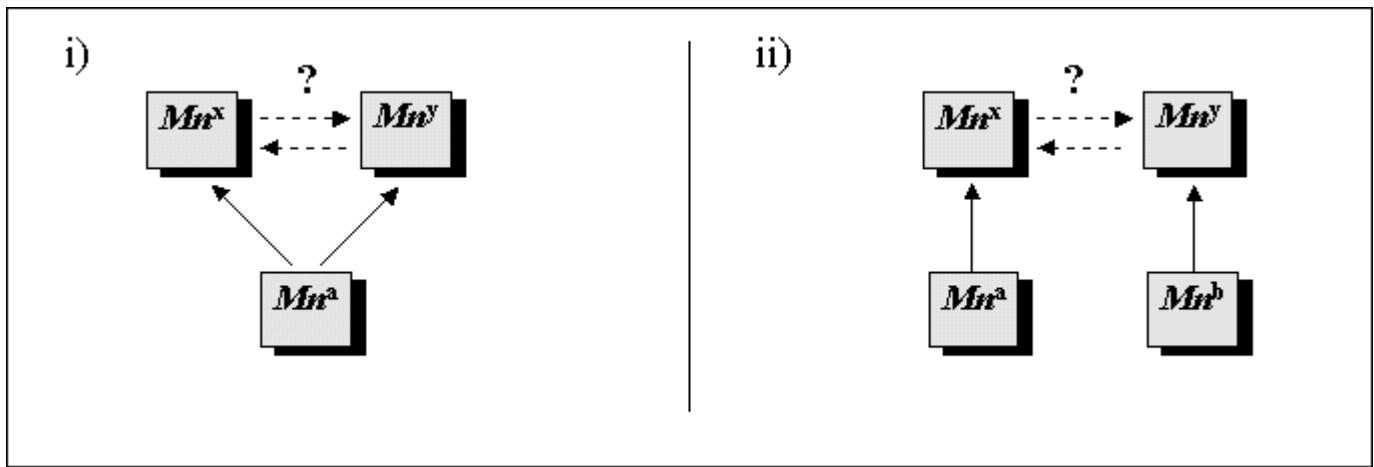
A coequal can exist at any level of the stylistic hierarchy discussed in Section 3, although in reality the coequal can exist in concrete form only at the level of intraopus style. Therefore a (musical) meme may exist abstractly at the level of dialect or idiom, but can only exist in reality at the level of the individual work. Here, as explained in Section 4, it manifests itself extrasomatically/phemotypically, but primarily derives from a somatic form in the composer's brain. In the absence of a proven coequal (or because the sheer magnitude of that which one has to investigate militates against identification of coequality, a particular problem in the memetic analysis of music), one might perhaps instead need to invoke a probabilistic perspective, positing the statistical probability of the existence of a coequal of given meme - without which, by definition, that mnemon would not attain memetic status - somewhere in the dialect at a particular point in time.

If the coequal of a mnemon  $Mn$  is identified one has to determine whether it is possible to verify if the coequal is the direct - i.e., causal - antecedent or consequent of  $Mn$  or whether the resemblance is adventitious. Dennett notes that "[t]his is the same epistemological problem, in the science of culture, that taxonomists confront when they try to sort out homology [resemblance due to direct descent] from analogy [resemblance due to fortuitous *convergent evolution*], ancestral from derived characters, in cladistic [species' family tree] analysis" (1995: 357). Ultimately, it may be impossible to resolve such questions, for the complexity of human communities and the convoluted modes of interaction between their members - particularly the admixture of horizontal and vertical modes of transmission - make verification of causality highly problematic. A workable definition of the meme probably ought not, in consequence, to contain a stipulation that a causal connection be verified between two equivalent particles, merely that a given particle has a coequal.

It is clear that a definition of coequality is more difficult in the case of two mnemons,  $Mn^x$  and  $Mn^y$ , which are similar in configuration but not identical to each other, for in such situations one must determine, using substrate-appropriate criteria, whether they merit the status of coequals, and hence memes. In such cases,  $Mn^y$  might be judged to be sufficiently different from  $Mn^x$  to be regarded as a mutant form of it, although it is not always apparent, even using chronological

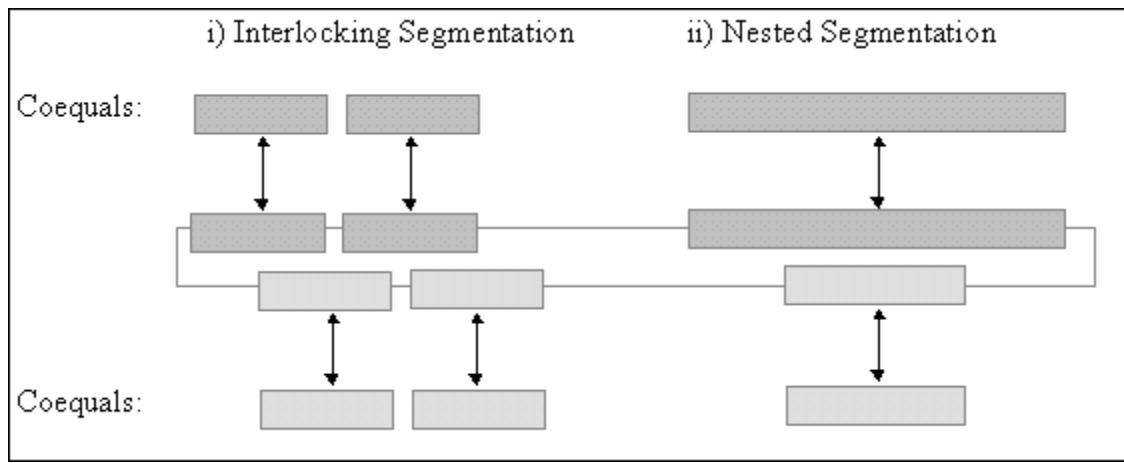
evidence, which is the antecedent and which is the consequent in a pair of mnemons in an assumed evolutionary relationship. If  $Mn^x$  and  $Mn^y$  are considered sufficiently different to preclude identity as coequals, their individual memetic status depends, of course, upon the existence of a coequal for each.

As Figure 3 illustrates, to regard  $Mn^y$  as a mutant form of  $Mn^x$  (or *vice versa*) implies direct evolution, by definition a causal process, whereas in reality the relationship might be one between independently generated variants of a single antecedent,  $Mn^a$  (Figure 3 i); or between two independently generated mnemons, each deriving from its own antecedent,  $Mn^a$  and  $Mn^b$  (Figure 3 ii):



**Figure 3:** Potential Ambiguities in Relationships Between Mnemons

It is often the case that a given span of music may contain two or more *interlocking* and/or *nested* memes, each segmentation of the pitch or rhythm matrix being defined and verified by reference to a coequal in another context - either the same work or a different one. Figure 4 represents this diagrammatically, with a musical illustration of nesting given in Example 7 i-v. Here a phrase by Mozart is shown to contain several nested memetic melodic segments, bracketed and labeled with the letters *a-d* (Example 7 ii is transposed to the key of C, the key of Example 7 i, iii, iv, and v). Example 7 vi (transposed to C) and vii, whilst not demonstrating interlocking or nesting, are added in order to show the coequals - bracketed and marked *e* and *f* - of other memes in i:



**Figure 4:** Interlocking and Nested Memetic Segmentations

i)

1 b  
a  
d  
c  
e  
f

ii)

15 a  
b  
c  
d

1  
 b

iii)

7  
 c

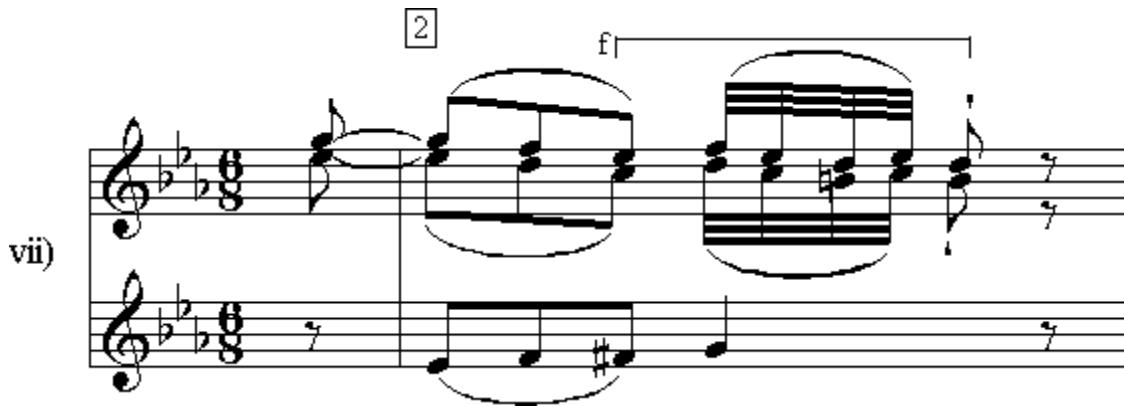
iv)

13  
 d

v)

66  
 e

vi)



**Example 7:** Nested Memetic Segmentations

- i) Mozart: Five Divertimenti for Three Bassett Horns K. 439b (1783-1788) no. 1, III, bb. 1-9.
- ii) Mozart: Serenade in B<sup>b</sup> major K. 361 (370a) (c. 1781) I, bb. 15-18.
- iii) Mozart: Missa Brevis in C major K. 220 (196b) (*Spatzen-Messe*) (1775) Agnus Dei, bb. 1-2.
- iv) Mozart: Missa Brevis in C major K. 220 (196b) Gloria, bb. 7-9.
- v) Mozart: *Die Entführung aus dem Serail* K. 384 (1782) no. 11, "Martern aller Arten," bb. 13-15.
- vi) Mozart: Serenade in C minor K. 388 (384a) (1782) I, bb. 66-68.
- vii) Mozart: Adagio and Rondo in C minor/major K. 617 (1791), bb. 1-2.

Note that the principle of coequality imposes articulation and segmentation on the two or more contexts in which replication occurs. A particulate, quasi-digital order - inhering ultimately at the level of the psychological representation - is engendered within what would otherwise be undifferentiated, quasi-analogue streams of musical data.

## 6. Structural Hierarchies and the Musical Memeplex

### 6.1. Pitch-Based Structural Hierarchies

In Section 3 I discussed the question of the hierarchy of musical style, defining style in terms of the cultural level at which memes were replicated, from the individual works of a composer to dialects spanning great swathes of time and space. There is, however, another and more fundamental sense in which musical memes exist hierarchically, as the following passages illustrate:

i)

[1]

3 1

1 2 4 5

1 2 5

**≡**

ii)

3 N 3 2 ||

3 2 1

I V I V I

**≡**

iii)

[1] 1

3 1

2 5 4

1 2 5

**≡**

iv)

$\begin{smallmatrix} \hat{3} \\ N \end{smallmatrix}$   $\begin{smallmatrix} \hat{3} \\ 2 \end{smallmatrix} \parallel$

$\begin{smallmatrix} \hat{3} \\ N \end{smallmatrix}$   $\begin{smallmatrix} \hat{3} \\ 2 \end{smallmatrix} \hat{1}$

I      V      I      V      I

v)

[1]

$\begin{smallmatrix} \# \\ \# \\ \# \end{smallmatrix} \frac{3}{4}$

$\begin{smallmatrix} \# \\ \# \\ \# \end{smallmatrix} \frac{3}{4}$

vi)

$\begin{smallmatrix} \hat{3} \\ N \end{smallmatrix}$   $\begin{smallmatrix} \hat{3} \\ 2 \end{smallmatrix} \parallel$

$\begin{smallmatrix} \hat{3} \\ N \end{smallmatrix}$   $\begin{smallmatrix} \hat{3} \\ 2 \end{smallmatrix} \hat{1}$

I      V      V      I

vii)

a) [91] 1

$\begin{smallmatrix} \# \\ \# \end{smallmatrix} C$

$\begin{smallmatrix} \# \\ \# \end{smallmatrix} C$

### **Example 8:** Memes at Various Hierarchic Levels

- i) Mozart: Flute Quartet in A major K. 298 (?1786) II, Minuet, bb. 1-8. [note 17]

- ii) Voice-Leading Reduction of i.
- iii) Mozart: Adagio in C major for Glass Harmonica K. 356 (617a) (1791), bb. 1-8, transposed to D major.
- iv) Voice-Leading Reduction of iii.
- v) Mozart: Flute Quartet in A major K. 298 II, Trio, bb. 1-8.
- vi) Voice-Leading Reduction of v.
- vii) Coequals of Memes 1-5, transposed to D major:
  - a) Meme 1: Beethoven, Symphony no. 9 in D minor op. 125 (1824) III, bb. 91-93.
  - b) Meme 2: Mozart, Symphony in C major K. 338 (1780) I, bb. 1-4.
  - c) Meme 3: Haydn, *Die Schöpfung* Hob. XXI: 2 (1798) no. 32, "Hölle Gattin!", bb. 14-16.
  - d) Meme 4: Beethoven, Symphony no. 8 in F major op. 93 (1812) I, bb. 5-8.
  - e) Meme (1+) 5: Mozart, Piano Quartet in G minor K. 478 (1785) II, bb. 1-4.

Examining the extracts at i and iii it will be evident that several segments from the flute quartet passage are copied almost exactly in the Adagio; these correlations are bracketed in the example and numbered 1-5. It is possible automatically to define these segments as memetic, for those in the first passage are copied in the second; indeed, the segmentation of the first passage is carried out, as discussed in Section 5.2, on the basis of that which is copied in the second. Nevertheless, an alternative reading of the passage would regard it as possibly controlled not by several discrete memes, but rather by a *single unit meme*, even though the length of the passage - and the presence of connecting material not replicated in both examples - somewhat militates against this interpretation.

To discount the single unit meme reading properly, it is necessary to ascertain the memetic status of the individual bracketed particles. Although one might only imply - on probabilistic grounds - the existence of other copies of these memes elsewhere in the dialect, it is prudent to attempt definitively to locate copies. Part vii a-e of Example 8 shows such copies, giving a coequal of each of memes 1-5 from other works from the dialect of Viennese Classicism.

Perhaps the best way to regard the passages in i and iii is to regard them as two copies of a *coadapted meme complex* or, to use Cees-Speel's term, a *memeplex* (in Blackmore 1999: 19). [note 18] As Dawkins notes,

Memes, like genes, are selected against the background of other memes in the meme pool. The result is that gangs of mutually compatible memes - coadapted meme complexes or memeplexes - are found cohabiting in individual brains. This is not because selection has chosen them as a

group, but because each separate member of the group tends to be favored when its environment happens to be dominated by the others. (in Blackmore 1999: xiv)

Having ascertained the memetic status of the individual segments in the passages, the replication, in order, of the memes of extract i in the passage at iii allows us to speak with some confidence of the replication of a memeplex in these phrases. As a general principle, then, a memeplex may only consist of units which are also replicated independently of their preferred partners and which in association form a class of at least two. If a replicated complex of discrete units exists only in this form, one must regard this larger, encompassing, unit as the meme, in which case no memeplex exists. Conversely, what might initially appear to be two large single unit memes - as in Example 8i and iii - might be better regarded as memeplexes. Such a determination, however, hinges upon the hierarchic location of the replicated unit(s). A middle ground structure - as in Example 8 ii and iv - might be regarded as a single unit meme, with copies existing at various hierarchic levels in other contexts, whereas the configurations generating this pattern may also demonstrate memetic existence independently of the memeplex into which they ostensibly assemble.

The 'voice-leading reductions' at Example 8ii and iv strip i and iii down to their fundamental framework using a method of analysis developed by the influential Austrian music theorist Heinrich Schenker (1868-1935) in which notes deemed, on various criteria, to be of lesser importance are gradually removed. [note 19] It is evident that the *foreground* events of the passages in i and iii generate a *middle ground* structure - the *interrupted linear progression* 3/I-4-3-2/V|| 3/I-4-3-2/V-I/I [note 20] - which is essentially common to both passages, aside from a few minor details of voice alignment and disposition of scale degrees. Because this structure is replicated - it occurs more than once - it is, by definition, of memetic status. It is a meme at a deeper level of structure, but is not necessarily *confined* to that level; one might easily find examples of replicated patterns based upon this scale degree succession situated closer to or at the foreground level.

Part v of Example 8 shows the theme of the first half of the Trio of K. 298 III. Despite the fact that its foreground harmonic and melodic organization is quite different from that of the first half of the Minuet, i, and of the opening of the Adagio, iii, its middle ground organization is essentially that of the passages in which the memeplex is replicated, ii and iv. This illustrates the point that a meme at a level of structure below the foreground may be expressed by a variety of foreground configurations. In the present case, the middle ground meme in ii and iv is expressed both by a memeplex located at the foreground, i and iii, and by a sequence of foreground events, v, which - to my knowledge - is not a memeplex.

On the evidence of the above examples, it is logical to suggest that, in addition to foreground-located memeplexes, such complexes may also be situated at the middle ground level. In such a situation, discrete middle ground events in one piece would be replicated in sequence in another, both complexes generating a common *background* pattern. This is, of course, analogous to the situation presented in Example 8 i-iv, but located at one level of structure deeper (assuming, for simplicity's sake, three principal structural levels; in some analyses, Schenker often identifies more than one middle ground stratum). Figure 5 represents these relationships diagrammatically: [note 21]

i)	<i>Urlinie</i> $\hat{5}-\hat{1}$ .				
	$\hat{5}$	$\hat{4}$	$\hat{3}$	$\hat{2}$	$\hat{1}$
ii)	Middleground Structures: Constituent Scale Degrees.				
a)		b)	c)	d)	e)
	$\underline{\hat{5}-\hat{4}-\hat{3}-\hat{2}-\hat{1}}$	$\underline{\hat{4}-\hat{5}-\hat{4}}$	$\underline{\hat{3}-\hat{4}-\hat{3}}$	$\underline{\hat{5}-\hat{4}-\hat{3}-\hat{2}}$	$\underline{\hat{5}-\hat{4}-\hat{3}-\hat{2}-\hat{1}}$
iii)	Foreground Memes $M^1-M^{20}$ .				
	$M^1-M^5$	$M^6-M^8$	$M^9-M^{11}$	$M^{12}-M^{15}$	$M^{16}-M^{20}$

**Figure 5:** Memeplexes at Various Structural Levels

At level iii of this hypothetical example, twenty foreground memes,  $M^1-M^{20}$ , each express (i.e., are principally based around) a single scale degree. These memes are shown as associating to form five groups. The first group of memes,  $M^1, M^2, M^3, M^4$ , and  $M^5$ , for instance, might express the scale degrees  $\hat{5}, \hat{4}, \hat{3}, \hat{2}$ , and  $\hat{1}$  respectively, giving rise to a middle ground linear progression  $\hat{5}-\hat{1}$ , shown in Figure 5 ii a. Each of these middle ground structures might, in turn, express a component (underlined in Figure 5 ii) of the  $\hat{5}-\hat{4}-\hat{3}-\hat{2}-\hat{1}$  *Urlinie* shown in Figure 5 i. The fifth middle ground structure, ii e, for instance, another  $\hat{5}-\hat{1}$  linear progression and itself generated by the group of memes  $M^{16}-M^{20}$ , might express the final pitch, the  $\hat{1}$ , of the *Urlinie*.

It will be evident that each middle ground structure in Figure 5 ii is a meme; examples may be found throughout the dialect of  $\hat{5}-\hat{4}-\hat{3}-\hat{2}-\hat{1}$  middle ground linear progressions, such as that expressed in bb. 18-22 of the first movement of Mozart's Piano Sonata in C major K. 545 (1788), illustrated in Schenker 1979 as Fig. 88 4 c. Moreover, if, by analogy with the foreground memes in the passage in Example 8 i, these five middle ground structures a-e were replicated *in sequence* in another work at the same level, as is analogically the case in Example 8 iii, they would clearly constitute a middle ground memeplex. As with the memeplex in Example 8 i and iii, such a memeplex will generate a meme at a still deeper structural level: in the same way that the foreground memeplex of Example 8 i and iii gives rise to the middle ground meme of Example 8 ii and iv; correspondingly, the middle ground memeplex of Figure 5 ii and its coequal in another work will give rise to the background meme of Figure 5 i and its coequal in that other work.

Furthermore, just as the middle ground structure shown in Example 8 ii, iv, and vi was not dependent upon the *specific* set of foreground memes present for its expression - this being evident from the divergence at the foreground level between the passages in Example 8 i and iii on the one hand and Example 8 v on the other - so the background structure in Figure 5 i is not contingent upon a *specific* set of middle ground memes for its expression. Indeed, one of the basic tenets of Schenkerian theory proposes a dualism between the limited number of background forms and the virtually unlimited number of middle ground and foreground patterns. Nevertheless, it is logical to infer, as does the previous paragraph, that the replication of a middle

ground memeplex - whether or not it is *itself* engendered by a foreground memeplex - will give rise to identical background memes, just as the replication of a foreground memeplex will generate identical middle ground memes. [Note 22]

This relationship between proximate events and more remote phenomena is essentially that conceptualized by Gjerdingen as the *schema-features* dichotomy. Both Gjerdingen (1988) and Meyer (1989) have written at length on musical patterns they term variously *archetypes* and *schemata*. [note 23] These are perceptual/conceptual frameworks being determined by, and determining, the nature and disposition of a number of lower-level elements, termed *features* by Gjerdingen. Thus, "[s]chemata can be defined as meaningful sets of features, and features can be defined as meaningful elements in sets of schemata" (1988: 6). From the point of view of the composer, schemata are primarily synthetic, generative tools, guiding the creative process; from the point of view of the listener they are analytical aids, directing the processes of perception and cognition.

Implicit in the Meyer-Gjerdingen notion of a schema is the principle of recurrence: schemata are replicated patterns of relationships and are, by virtue of this replication, necessarily memetic. If one regards the schema as a middle ground event and its constituent features as foreground phenomena, it will be evident, in the light of the discussion of Example 8 and Figure 5 that a diversity of foreground features may assemble to generate a specific middle ground schema.

## 6.2. Structural Memes in Music

It is logical to extend the hierachic perspective adopted in discussion of the relatively circumscribed examples of Section 6.1 to the complex, large-scale patterns which composers repeatedly employ as the overarching structural framework of their compositions. Such archetypal configurations exist because they are repeatedly reinstated - replicated - by memes at more immediate levels of musical organization.

Cook argues, on the basis of several empirical studies, that these structures seem to have little tangible reality for listeners, suggesting that

...few people actually experience musical compositions as such, in the sense of constituting them as fully co-ordinated, objective structures. Unless they have both the training and the inclination to track the form of a piece of music in theoretical terms as they listen, people experience recurrence without actually observing what it is that recurs; they experience coherence but not the unitary organization in terms of which a theorist or analyst would explain that coherence. People enjoy musical compositions, in other words, without really perceiving them at all; rather than listening *to* them...they "just listen." (1990: 68)

Nevertheless, they apparently had clear existence for composers, for as Drabkin notes, apropos of Beethoven:

...we know that he studied the music of his predecessors Haydn and Mozart and modeled a number of sonata-type works of the 1790s and early 1800s on specific works by Mozart [Beethoven's String Quartet in A major op. 18 no. 5 (c. 1799) is, for instance, closely modeled on

Mozart's String Quartet in A major K. 464 (1785)]. If a piece by Mozart could suggest to Beethoven a set of procedures to be followed - with modifications - to produce another piece, then he probably did not regard the growth processes in a particular piece as the exclusive property of that piece. (In Kinderman 1991: 15)

This is equivalent to saying that Beethoven - indeed probably most composers working in traditions with strongly developed notions of form and genre - perceived a fairly clear separation between a work's global organization and its detailed content. He realized that a given structural configuration might receive many instantiations each of which was unique in its local details but all of which shared broadly analogous "growth processes."

Given this process of replication by repeated instantiation, it is a small step to the view that the fundamental archetypes of music are memes - *structural memes* - at the largest dimensions and highest hierarchical levels of musical organization; they are, to use Gjerdingen's terminology, superordinate schemata which may be instantiated by a diversity of subordinate features. Given this, one can conceive of interplay, perhaps indeed tension, between the top-down processes of the structural memes (which require and perhaps select memes at lower hierarchic levels for their generation and replication), and the bottom-up processes of the memes at lower hierarchic levels (which selfishly pursue their own replication without regard for the structural memes which they generate and replicate). Moreover, one might even envision a process of allelic rivalry between lower-level memes in the competition to occupy the *loci* in a structural archetype.

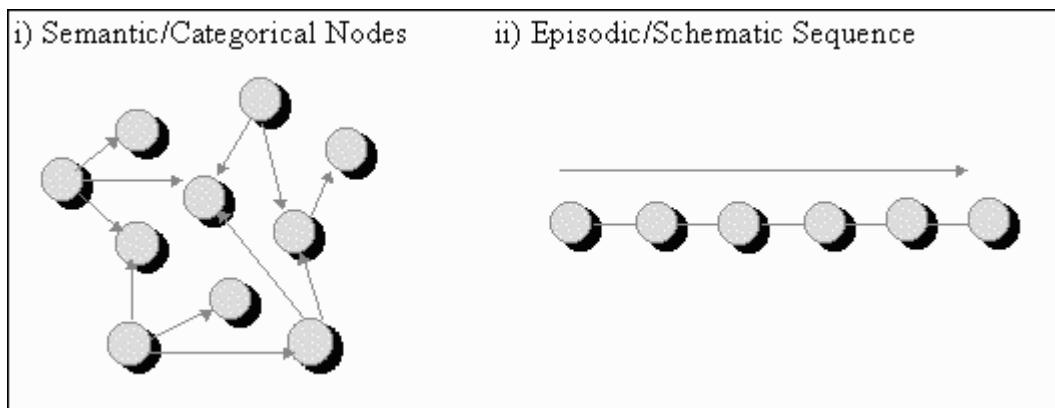
How structural memes should be conceptualized and described has been one of the great debates of music theory and, latterly, music analysis. These disciplines have, moreover, had to keep pace with the evolution of structural memes over the last four hundred years. Conceptual memplexes assembled to account for such configurations include, in roughly chronological order, Quintillian rhetoric (Burmeister, c. 1600), phrase assemblage (Koch, c. 1780), formal models (Marx, c. 1840), voice-leading strata (Schenker, c. 1930), motivic elaboration (Réti, c. 1950), and Chomskyan generative grammar (Lerdahl and Jackendoff, c. 1980). [note 24] It will be realized that these conceptual memplexes are themselves subject to variation and selection (including by the criterion of fit with that which they aim to elucidate) over time and, as a result, undergo evolution in association with the structures they endeavor to describe.

I hope to return to this question of deep structural archetypes in music and their evolution - the musical analogue to Dawkins' 'blind watchmaker' principle (1991) - on another occasion. [note 25] Before leaving the issue, however, it should be noted that structural memes clearly retain the fundamental characteristics of memes at lower hierarchic levels, in that they are composed of a relatively small number of discrete components in a clear and distinct relationship with each other. The observations made in Section 5.1 on the implications of the memetic attributes of copying fidelity and fecundity would therefore appear to apply to memes irrespective of the hierarchic level at which they are propagated. Whether one chooses to speak in terms of traditional formal models, such as binary, ternary, or sonata form; [note 26] the Schenkerian *Ursätzen* 3-1, 5-1, or 8-1; or the *normative prolongational structure* proposed by Lerdahl and Jackendoff, [note 27] one is essentially dealing with small memes - particulate entities consisting of relatively few elements - despite their generally expansive temporal arch.

### 6.3. The Psychological Organization of Musical Memes

As is discussed in Section 4, the patterns analyzed in Example 8, as with all memes, exist fundamentally as neuronal configurations within the brain. Whilst a detailed treatment of the neurobiology of memory is beyond the scope of this article, [note 28] the matters covered in Section 6.1 warrant a return to this question from the perspective of psychology, examining the issues of hierarchical location in musical memetics in terms of the psychology of memory organization. In discussing the schema-features dualism, Gjerdingen notes that we demonstrate two types of memory structure, termed *semantic* and *episodic* by Tulving; these correspond to the classifications *categorical* and *schematic*, respectively, proposed by Mandler (Gjerdingen 1988: 59). It should be noted, however, that these are psychological, not neural models, nor are they as separate as the following necessarily simplified exposition implies.

Semantic/categorical memory takes the form of an *associative network* of information-encoding memory nodes (memes) and their interconnections. Delius notes that "[s]uch plastic synapses have to be thought of as the critical components of neural networks that function as associative arrays" (1989: 44; see also his Figure 2; and Gjerdingen 1988: Figures 4-1, 4-2). The memory patterns for foreground features such as memes 1-5 in Example 8 i and iii would seem to be subject to this type of organization. In contrast, episodic/schematic memory structures the recall of discrete events anchored in linear time. The memory patterns for a middle ground schema such as that shown in Example 8 ii, iv, and vi, with its clear sequence of component events, would seem to be subject to this type of organization. These two models of memory organization may be represented diagrammatically as follows:



**Figure 6:** Semantic/Categorical and Episodic/Schematic Memory (after Gjerdingen 1988: Figures 4-1, 4-2, 4-3)

Having outlined their differences, Gjerdingen asserts that

These two types of memory organization commingle if we allow the term *feature* for event schema to coincide with the term *concept* for semantic[/categorical] networks....[I]n music, *dominant seventh chord*, for example, can be both a concept linked to the memory nodes *instability*, *dissonance*, and so on, and a feature in the [episodic/]schemata[tic] *perfect authentic cadence*.... (1988: 61; his emphases)

In the light of the discussion of Example 8, one might suggest that foreground-orientated memes (features) are coded by the individual neural nodes in a semantic/categorical associative network; and that middle ground-orientated memes (schemata) are the consequence of the organization of these nodes into sequential configurations fixed in and mediated by episodic/schematic memory. Moreover, these middle ground memes then act as features coded by the nodes of a higher-order semantic/categorical associative network; these are themselves fixed sequentially into a superordinate episodic/schematic configuration (the highest level schema), of which the *Ursatz* is one representation. In this sense, the hierarchic, embedded, nature of memory organization gives rise to the hierarchic, embedded, nature of musical organization.

Despite its congruence with observed musical phenomena, Gjerdingen's synthesis of the two forms of memory organization in the extract cited above tends to blur the distinction between signifiers and signifieds. One method of clarification would be to regard the spoken (verbal-phonetic) and written (verbal-graphical) patterning 'dominant seventh chord' as the phemotypic products of somatic memes which act as signifiers. [note 29] To this category one should add the memes antecedent to such phemotypic products as the written (musico-graphical) patterning of the dominant seventh chord:



**Figure 7:** The Musico-Graphical Symbology of the Dominant Seventh Chord

Together, and in coadaptation with an array of other memes, these form the matrix of concept memes - the network of interlinked notions Peirce termed *interpretants* (in Nattiez 1990: 6) - which structure our verbal-conceptual cognition of this chord.

Whilst to some extent independent, these three signifier memes also exist in coadaptation with the somatic *sonority meme*, our mental image of the sound - i.e., the sound internalized ('heard') in the brain - which operates as their signified and whose phemotypic product, via the intercession of musical instruments, is the physical sonority through which the dominant seventh chord impinges upon us most directly. This complex relationship is illustrated below: [note 30]

	MEMOTYPE	PHEMOTYPE
SIGNIFIERS	Verbal-Phonetic Meme Dominant Seventh	Verbal-Phonetic Pattern "Dominant Seventh"
SIGNIFIED	Sonority Meme 	Physical Sonority 

**Figure 8:** The Memetic Constellation of The Dominant Seventh Chord

It will be understood that in verbal signification systems, the somatic signifier - the verbal-phonetic and/or verbal-graphical memes antecedent to the language-specific spoken and/or written word - and the somatic signified - the idea or concept with which they are stably united - give rise to a unitary phemotypic manifestation, a spoken and/or written verbal configuration. In music, by contrast, a separation is maintained, in that the somatic signifier - the memes antecedent to the verbal-phonetic and/or verbal-graphical phrase 'dominant seventh chord' and/or the musico-graphical image shown in Figure 7 - and the somatic signified - the memes antecedent to the physical sonority - give rise to separate manifestations which preserve the signifier-signified dualism at the level of the phemotype.

#### 6.4. Topics

In the previous section, musical memes were understood to act as the signified element of semiological memplexes in which verbal-conceptual memes acted as the signifier. Another sense in which music forms semiological memplexes is in contexts where a musical meme acts, conversely, as a signifier, becoming associated with verbal-conceptual memes, which are the signified. Musicologists such as Ratner (1980: 9 ff.) and Allanbrook (1983: 1 ff.) tend to refer to such memplexes as *topics*. Another figure in this tradition, Agawu, observes that

A topic, T, may be defined as a musical sign. Each T embodies the union of a signifier and a verbally mediated signified. Signifiers are purely musical dimensions such as texture, timbre, rhythm, melody, and harmony *in a particular disposition*....The signified or concept is represented by an arbitrary label drawn from [the] U[niverse of] T[opic and] normally point[s] to dimensions of historical or sociocultural specificity, including the elements of a contemporary compositional code. (1991: 128-129)

An example of a topic is the chromatic descending tetrachord pattern, which consists of the motion from the tonic of the key to its dominant (the fifth note) downwards by semitones. [note 31] From Monteverdi to Beethoven, this meme was regularly replicated in coadaptation with texts expressing death, grief, and the funereal. In the following passage from Purcell, the tetrachord is bracketed:

The musical score consists of two staves of three-measure sections. Measure 6 starts with a half note on G in the treble clef staff. Measure 7 begins with a half note on F# in the bass clef staff. Measure 8 starts with a half note on G in the treble clef staff. Measure 9 starts with a half note on A in the bass clef staff. Measure 10 starts with a half note on B in the bass clef staff.

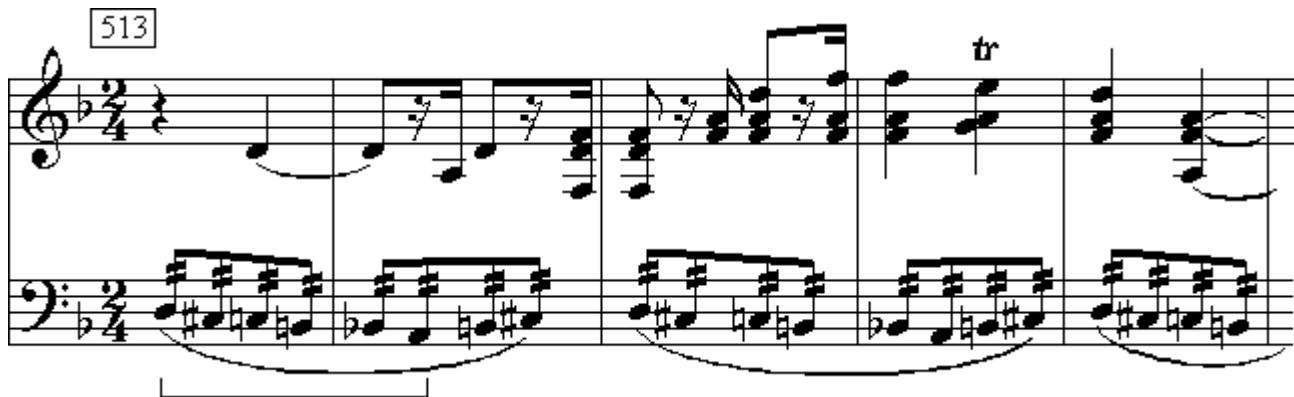
**Example 9:** A Descending Chromatic Tetrachord in Purcell

i) Musical Signifier: Purcell, Dido's Lament from *Dido and Aeneas* (1689), bb. 6-12.  
Text: When I am laid in earth, may my wrongs create no trouble in thy breast.

ii) Verbal-Conceptual Signified: Coadapted Concept Memes

Death    Grief    Lamentation    Sorrow

If theorists of this period are to be believed, the strength of the association was such that even without the verbal-conceptual signified, the musical signifier alone called forth the coadapted concept memes in the minds of contemporary listeners. Indeed, these associations endure today. Hopkins, for example, describes the following passage, from Beethoven's Ninth Symphony, as "...a funeral march. Its foundation is a chromatic wailing in the bass [the tetrachord, bracketed] above which brass and wind proclaim funereal fanfares" (1981: 259):



**Example 10:** A Descending Chromatic Tetrachord in Beethoven: Symphony no. 9 in D minor op. 125 (1824) I, bb. 513-517 (Hopkins 1981: 259, Ex. 255)

As Figure 2 indicates, the topic exists at the level of dialect - the association of musical meme with concept meme is copied in all the brains of the members of a given cultural community - but also shares the meme pool with analogous memplexes consisting of personal associations of musical memes and concept memes at the level of idiom. A composer may, in other words, replicate a private association between musical patterning and verbal concepts in his or her works which remains inscrutable to the wider musical community.

## 7. The Transmission and Mutation of Musical Memes

The center of our theoretical models is the concept of transmission: for biologists, and in particular geneticists like ourselves, the centrality of the concept of transmission to evolutionary thinking is patently obvious. In biological material, however complex, there exists an almost perfect mechanism of genetic transmission. Evolution is the result of mutational exceptions to perfect transmission and the differential success of the products of this imperfect transmission. (Cavalli-Sforza and Feldman 1981: 341)

In cultural material, the "mechanism of transmission" is less perfect than is the case with biological material and, as a result, the "mutational exceptions to perfect transmission" are more numerous. These deficits in copying-fidelity are thus intrinsic to all replicants, but are a particular propensity of the meme; the parlor game of 'Chinese Whispers' represents this process in isolated and compressed form. The tendency towards miscopying makes available memes which are variants of existent forms - i.e., mutations - and which are subject, as are all memes, to the pressure of 'cultural selection', the analogue, proposed by Cavalli-Sforza and Feldman, to the natural selection which regulates the replication of genes within living systems.

They define this as

...the rate or probability that a given innovation, skill, type, trait, or specific cultural activity or object - all of which we shall call, for brevity, *traits* - will be accepted in a given time unit by an individual representative of the population. (1981: 15)

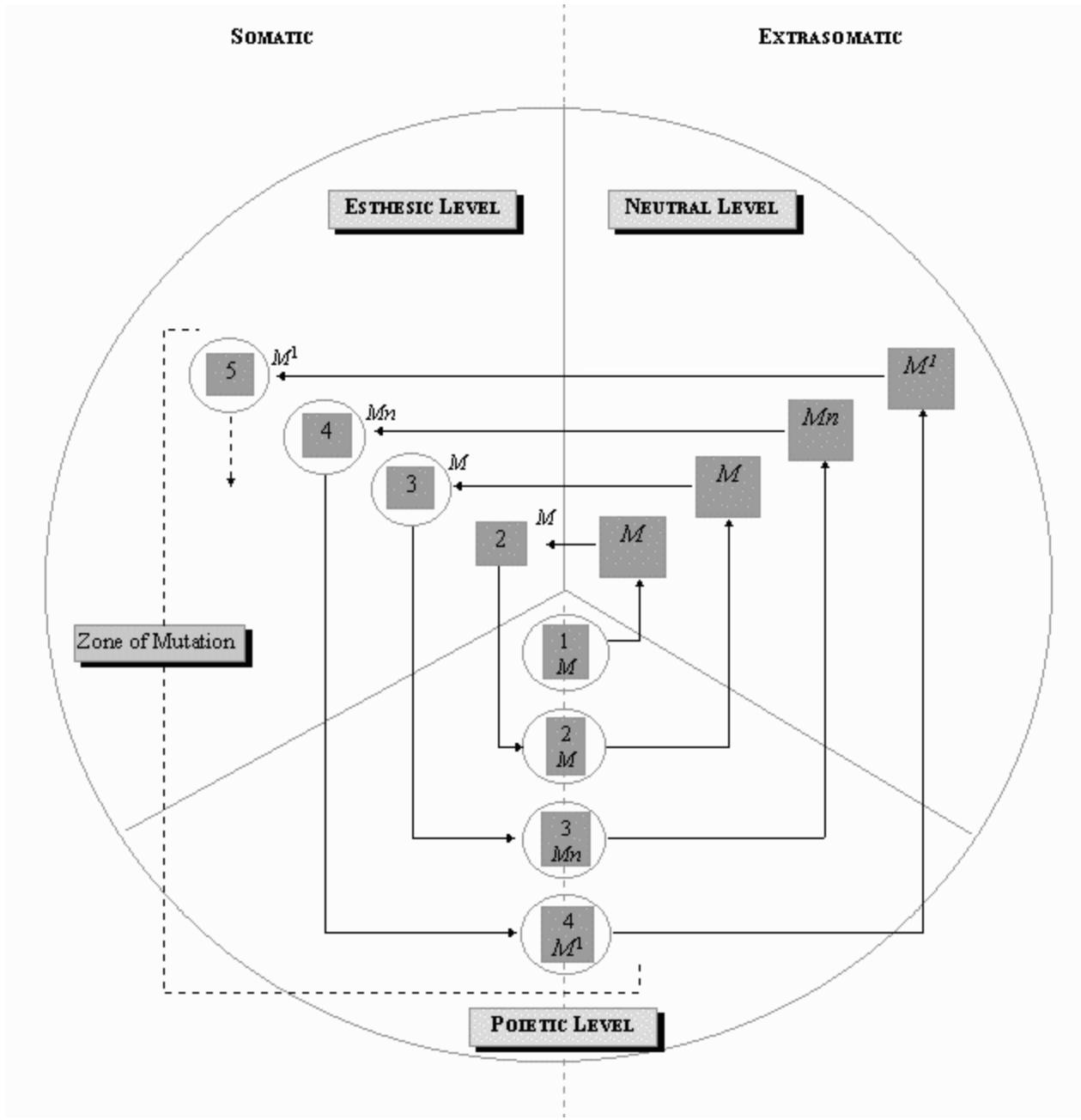
Like natural selection, cultural selection works upon the totality of memes (traits, in the terminology of Cavalli-Sforza and Feldman) within a cultural community and leads to their

differential replication. That is, if a meme arises which is a clear variant of an existing meme it may persist and increase its representation or it may become extinct, according to whether it is more or less susceptible to imitation as a result of the changes it has undergone. The net result of this process of cultural selection is change in the relative frequency of memes in the meme pool over time. The concept of susceptibility to imitation is directly analogous to the biological notion of reproductive fitness; it is an index of the replicant's capacity for replication which Cavalli-Sforza and Feldman term 'cultural fitness' (1981: 17). In biological replication, a gene's fitness is contingent upon the phenotypes it plays a role in engendering. In cultural replication, an analogous situation obtains, for a meme is not selected directly; rather, it is selected by virtue of its phenotypic consequences. [Note 32]

Figure 9 represents the processes of memetic transmission and mutation schematically. Because cultural products are transmitted from a producer to a receiver via an intermediate (phenotypic) form, one means of representing the process is to adapt the Molino/Nattiez notion of the *semiological tripartition* (Nattiez 1990: 10 ff.). The figure traces the movement of memes as they pass between the *poietic*, *neutral*, and *esthetic* levels and across the somatic-extrasomatic boundary which bisects the diagram. [Note 33] Five individuals, or hosts, are shown; these individuals, 1-5, need not be in immediate geographical or chronological relationship with each other, although they often are:

Here meme *M*, shown located in the brain of the first host, 1, is situated at the poietic level and, initially, exists in somatic form. The process of conversion into the extrasomatic form, a musico-graphical or musico-aural version, first moves the meme across the somatic-extrasomatic boundary (the vertical line bisecting the diagram) within the poietic level - at this stage the meme may exist in the context of compositional sketches and drafts - and then across to the neutral level, where the meme exists as part of a completed, notated (or even improvised) composition. The second host, 2, initially encounters the meme from the standpoint of the esthetic level, either by hearing the meme in its musico-aural form or by seeing the meme in its musico-graphical notation. The movement from the esthetic back to the poietic level within the second host corresponds broadly to the two principal stages of memetic transmission identified by Cavalli-Sforza and Feldman: *awareness* followed by *adoption* (1981: 62-65).

Once the adoption stage is accomplished, meme *M* now resides in the brains of both the first and second hosts and may then be transmitted to other hosts. Of course, a meme is never imitated with complete accuracy over any length of time; some changes, albeit very slight, are almost inevitable during this process. As the diagram shows, a third host, 3, might encounter *M* and subject it to alteration at some stage in the course of its progress through the sectors esthetic-somatic (awareness), poietic-somatic (adoption), poietic-extrasomatic (compositional sketches), and finally neutral-extrasomatic (the musico-graphical and/or musico-aural phenotypic form). The sectors of the diagram within which mutation can occur are indicated by dotted brackets. The changes imposed on *M* might be *intentional*, *accidental*, or indeed *obligatory*. In the latter case, this is determined by the new context into which the meme must be inserted, for memes, as has been seen in Section 6.1, are rarely transmitted singly but as components of larger memplexes.



**Figure 9:** Memetic Transmission and Mutation and the Three Levels of the Semiological Tripartition

At its inception, the mutant is a mnemon,  $Mn$ , rather than a meme, for if a particle defined as or assumed to be memetic gives rise to a mutant form, the mutant, at its moment of genesis, is not memetic; at this moment it is theoretically unique, unimitated. When the mnemon is copied into the brain of a subsequent host - 4 in Figure 9 - the particle attains memetic status, as meme  $M^1$ . The mutant form may *overwrite* the original meme in the brain of the third host or coexist with it. It may even be (re-)encountered by the first and second hosts and copied into their brains,

possibly overwriting the original copies of  $M$  or coexisting with them. As Figure 9 shows, meme  $M^1$  is then subject to further imitation, and possible mutation, by a fifth host, 5, and subsequent hosts *ad infinitum*.

This account hopefully makes clear that, whereas the gene is propagated non-phenotypically (ultimately genes, not bodies and other gene products, reproduce themselves), the meme, of necessity, must be propagated phenotypically. To recall Blackmore's terminology cited in Section 4.1, memetic propagation is an example of "copy-the-product," as opposed to the "copy-the-instructions" of genetic transmission (1999: 61).

## 8. The Evolution of Musical Style

It is perhaps true to say that the central achievement of *The Selfish Gene* was to recast an old argument in evolutionary biology in the most radical and striking terms, and in so doing breathe new life into the New Synthesis. Biologists have long debated the question of the level at which natural selection operates - the 'units of selection' issue. Does it operate at the level of the species (i.e., are species selected for survival or extinction *en bloc*); at the level of the group within a species (i.e., is the group favored at the expense of the larger collective); or at the level of the individual organism (i.e., is the species simply a useful term for discussing a collection of genetically and morphologically similar but functionally independent agents)? Dawkins dismissed all of this, insisting that ... the best way to look at evolution is in terms of selection occurring at the lowest level of all....the fundamental unit of selection, *and therefore of self-interest*, is not the species, nor the group, nor even, strictly, the individual. It is the gene, the unit of heredity. (1989a: 11; my emphasis)

Given that the principle of Universal Darwinism holds that these remarks must be taken to apply analogically to all replicators, the same considerations must be applied to memetics. Thus, the single unit meme is the fundamental unit of selection, and not any larger configuration. Without wishing to labor analogies between genetics and memetics, it is nevertheless interesting to note that Meyer's hierarchical view of cultural structure outlined in Section 3 maps fairly neatly onto the biological hierarchy Dawkins describes, helping us to comprehend more clearly the implication of the unit of selection principle to musical memetics:

NATURE	CULTURE
Species	Dialect
Group of Organisms	Idiom/Genre/Formal-Structural Type
Individual Organism	Work
Gene	Meme

**Table 2:** Correspondences between Natural and Cultural Hierarchies

Here, the cultural analogue for the species may be taken to be the dialect; a group of organisms within the species can be compared with either the composer's personal style, with a musical genre (such as the symphony, the string quartet, or the opera), or with a formal-structural type (such as those considered in Section 6.2); and the individual organism maps onto the self-contained musical work, the individual representative of a genre. Lastly, the memes whose

phenotypic products constitute a musical work are analogous to the genes whose phenotypic products constitute the individual organism.

It is, of course, intuitively clear that the meme is the unit of selection in the evolution of musical style: whole dialects, such as the Austro-German Classical style, are not selected for or against; neither are the styles of composers, musical genres, or formal-structural types; and neither are whole works. What are selected are the memes which constitute those works. In this view, cultures evolve from the bottom up, not the top down, the nature of the system being a function of the properties of the units at the lowest hierachic level.

In the evolution of musical style, a mutant meme appears in the work of a composer, the level of intraopus style, by means of the processes discussed in Section 7. Providing it possesses high longevity, fecundity, or copying-fidelity, it may then be copied in other works of that composer, thereby moving to the level of idiom. Further imitation might disseminate the meme among the composers in a community, taking it to the level of the dialect, the configuration of the which is slowly and subtly changed.

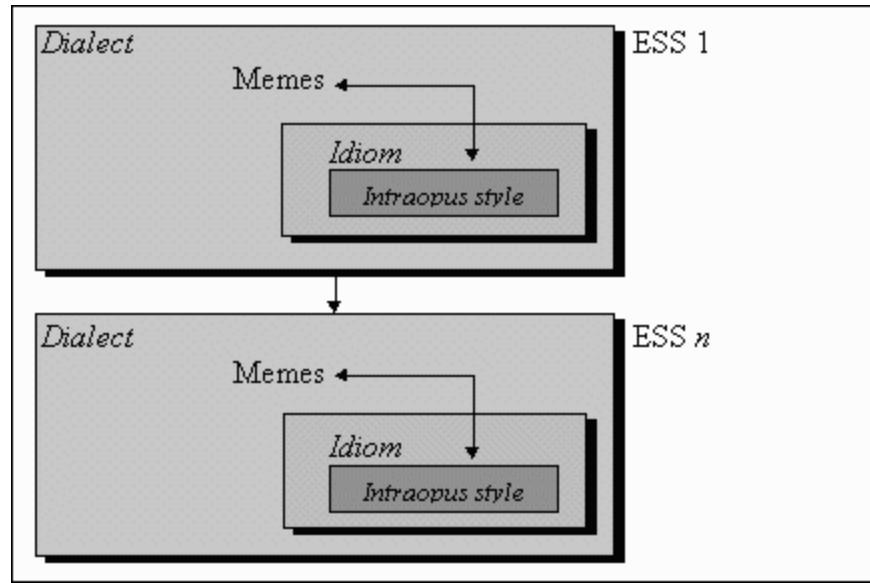
However, it is the case, as Gatherer observes apropos the evolution of jazz, that New memes, such as novel chord patterns, rhythmic changes, or alterations in instrumentation are subject to scrutiny as to how they cohere with the pre-existing whole. Sudden innovation is not permitted but small innovations may cumulatively have large results. (1997: 80)

This point is developed by Dawkins in a passage in which the word 'meme' may be substituted for 'gene' throughout. He asserts that

The gene pool will become an *evolutionarily stable set* of genes, defined as a gene pool that cannot be invaded by any new gene. Most new genes that arise, either by mutation or reassortment or immigration, are quickly penalized by natural selection: the evolutionarily stable set is restored. Occasionally a new gene does succeed in invading the set: it succeeds in spreading through the gene pool. There is a transitional period of instability, terminating in a new evolutionarily stable set - a little bit of evolution has occurred....Progressive evolution may be not so much a steady upward climb as a series of discrete steps from stable plateau to stable plateau. It may look as though the population as a whole is behaving like a single self-regulating unit. But this illusion is produced by selection going on at the level of the single gene. (1989: 86)

The notion of the *evolutionarily stable strategy* (ESS), formulated by J. Maynard Smith, is a concept fundamental to understanding cultural evolution. In biological evolution, an ESS is essentially a behavioral strategy which, "...if most members of a population adopt it, cannot be bettered by an alternative strategy," "[a] strategy that does well in a population dominated by the same strategy" (Dawkins 1989: 69; 1983a: 286). Many patterns of animal behaviour, including aggression and territoriality, in which the conflicting self-interests of the members of a group reach an equilibrium, are understandable in terms of this notion. Individuals deploying deviant strategies against the majority strategy generally cannot better it, and they and their strategies are consequently penalized by natural selection.

In cultural evolution, the ESS may be understood as an index of the resistance of a meme pool to incursions by mutant memes. However, because the meme "...is already achieving evolutionary change at a rate that leaves the old gene panting far behind" (Dawkins 1989: 192), even a year of memetic evolution is comparable to many millions of years of genetic evolution. Consequently, the generation and propagation of mutant memes, and the resultant motion from "stable plateau to stable plateau" as one ESS is supplanted by another, occurs at a rapid, and probably indistinguishable rate. This process is represented, in simplified form, below:



**Figure 10:** Successive Evolutionarily Stable Strategies

Acceptance of this perspective forces one to question Meyer's notion of dialect, which he defines as "...substyles that are differentiated because a number of composers...employ (choose) the same or similar rules and strategies" (1989: 23). This differentiation may be on the basis of geography, nationality, cultural movements, social class, or function within the culture; "...but most often dialects are distinguished historically" (1989: 23). From this one might interpret the period *c.* 1700-*c.* 1900 as comprising three historical dialects, the Baroque (the period of Bach and Handel), Classicism (the period of Haydn, Mozart, and Beethoven), and Romanticism (the period of Schumann, Wagner and Brahms), operating within the constraints of a single set of rules, that of tonality. Yet whilst no musicologist would claim a strict articulation between these periods, if one regards the evolution of musical style as driven by a fluid meme pool configuring itself into a succession of transient ESSs, then the notion of discrete historically distinguished dialects cannot logically remain tenable. The meme pool of Bach's time evolved through a near infinity of ESSs until it became transmuted into that of Brahms'.

Moreover, skepticism is needed concerning the rigidity of the distinction Meyer draws between diachronic changes from one dialect to another, which he terms *trended* changes (for example, between Classicism and Romanticism, which occurred *c.* 1800-*c.* 1820); and changes from one system of rules to another, which he terms *mutational* changes [note 34] (for example, between tonality and serialism, [note 35] which occurred *c.* 1900-*c.* 1920). He conceives this distinction as one between differences of degree and kind respectively, asserting that the two modes of style

change are not necessarily related (1989: 102), and suggesting that "[i]t does not follow...that the trend *caused* the mutational change, only that the stylistic/cultural situation created by the trend made it possible for the mutation to endure - be replicated....Nor is it supposed that if a stylistic consensus takes place, the new rules (mutations) will be a result of prior trends or strategies" (1989: 103, notes 79, 80).

Whilst there is no teleological inevitability of a given rule change's arising from antecedent changes in the dialect(s) it encompasses, it must be the case - accepting the bottom-up perspective of memetic evolution - that the large-scale reconfiguration represented by a change in rules can only be defined in terms of the changes occurring upon lower hierarchic levels. A neo-Darwinian, meme-selectionist perspective would regard mutational changes as the inexorable outcome of cumulative trended changes, the latter acting not only as a necessary condition for the former but also as determinants of the configuration of the new higher-order system. In this respect, cultural evolution again parallels biological evolution, for "...all sane Darwinians are gradualists in the extreme sense that they do not believe in the *de novo* creation of very complex and therefore statistically improbable new adaptations like eyes. This is surely what Darwin understood by the aphorism 'Nature does not make leaps'" (Dawkins 1983a: 287-288). [Note 36]

Lastly, just as the notion of distinct, historically sequential dialects is undermined by a view of musical style as a succession of transient ESSs, so the idea of distinct, historically sequential systems of rules cannot logically be sustained. Whilst the division of style along the vertical (synchronic) axis of rules, dialects, and idioms is a useful intellectual tool for structuring a complex field, retaining its clarity at any given point in musical time, when considered in association with the horizontal (diachronic) axis representing temporal progression its categories of rules and dialects are ultimately blurred by the inexorable, random gradualism of memetic mutation.

## **9. Conclusion: Towards a Memetics of Music**

Given the great complexity of the subject it addresses, this article has only scratched the surface of what might constitute a memetics of music. I hope, nevertheless, to have outlined a number of what I see as fundamental points which, in conclusion, are now briefly reviewed.

Whilst commentators such as Dawkins and Dennett have seen musical memes as, to re-employ the linguistic analogy, comparable with sentences, I have attempted to show that they are more akin to the individual word - that music is made up of small units of pitch and rhythm which, by virtue of their replication, should be regarded as memetic.

The examples have illustrated that music is a rich field of interacting parameters, some of which are capable of sustaining memetic replication. Whilst I have largely confined myself to the parameter of pitch, in common with the practice of much analytical musicology, a comprehensive memetics of music must account for these individual parameters and, more importantly, their interaction in the form of parametric memplexes.

In the realm of pitch, music, perhaps more than any other symbolic system, sustains the propagation of memetic structures at a variety of hierarchic levels, from its immediate surface to

more remote levels. At the deepest levels of structure are propagated structural memes, the superordinate configurations repeatedly reinstated by composers as the guiding architectural frameworks of their compositions. The analytical method of Schenker is an effective, although by no means unproblematic, method of representing these structures. Whilst not discussed here, it seems the case that rhythmic memes are similarly propagated at several hierarchic levels - a subject to which Yeston 1976 is addressed, although not in terms of the memetic paradigm. I hope to have shown that, as with all cultural evolution, the evolution of music occurs because of the differential selection and replication of mutant memes within idioms and dialects. Slowly and incrementally, these mutations alter the memetic configuration of the dialect they constitute. Whilst gradualistic, this process eventually leads to fundamental changes in the profile of the dialect and, ultimately, to seismic shifts in the overarching principles of musical organization, the rules, propagated within several dialects.

Finally, I hope to have indicated that the application of musicological insights to memetics and the memetic paradigm to musicology offers considerable benefits to both areas of inquiry. To memetics, music offers a field of study in which the units of selection are, arguably, clearer and more amenable to analysis than verbal text- and concept-based memetics. To musicology, memetics offers a new viewpoint on its concerns, one which has the power to unify a variety of disparate musical subdisciplines - composition, performance, improvisation, listening, notation, theory and analysis, history and criticism, musical style and style-history, music-psychology, aesthetics and philosophy, and influence [note 37] - under the cool logic of a neo-Darwinian, meme-selectionist perspective.

## Notes

1. One might qualify Dennett's assertion by noting that for any substrate there is a limited set of algorithms it can instantiate. Moreover, substrates vary in their ability to carry out an algorithm; there is a continuum of efficiency between a set of substrates that are able to carry out the algorithm and a set that are not. A nonarbitrary relationship therefore exists between the logical structure of an algorithm and the materials that instantiate it. It is, of course, my proposition that the algorithm of evolution by natural selection *is* one music instantiates well, a claim for which this paper is hopefully seen as offering evidence.
2. The various strands constituting the discipline of musicology are surveyed in Kerman 1985.
3. Gatherer's (1997) contribution to the development of a memetics of music is acknowledged. A longer treatment of the subject of musical memetics than the present article may be found in my forthcoming article 'The Selfish Meme: Particularity, Replication, and Evolution in Musical Style' (2000). A few paragraphs in the present article, particularly in Section 8, are adapted from 'The Selfish Meme.'
4. Constraints of space prevent the discussion of music outside the western European art canon. Traditional and popular musics are, I believe, readily amenable to examination in memetic terms, although arguments presented here involving notation are not always pertinent to non-western musics, in which graphical fixation often plays little or no role. This is partly a consequence of

the fact that the western trichotomy between composer, performer, and listener is rarely so sharply drawn in non-western musics.

5. See Scruton 1997: Chapter 7 for a detailed examination of the relationship between music and language.

6. The works of Beethoven (1770-1827), for instance, are generally divided into three style-periods, albeit with very fuzzy boundaries at c. 1803 and c. 1815.

7. Modality is the organization of the seven pitches C, D, E, F, G, A, and B into a number of different scale patterns, each with its own distinctive internal intervallic (i.e., pitch spacing) pattern. It was the system of pitch organization employed during the Medieval (c. 500-c. 1450) and Renaissance (c. 1450-c. 1600) periods, but has its origins in the music of ancient Greece.

8. Tonality is the organization of the twelve pitches of the chromatic scale into a hierarchy centered around one note, the *tonic* or *key-note*. It was the system of pitch organization employed during the Baroque (c. 1600-c. 1750), Classical (c. 1750-c. 1820), and Romantic (c. 1820-c. 1900) periods. Boundaries between these periods, and between modality and tonality and between tonality and serialism (see note 35), are, for reasons discussed in Section 8, highly fluid.

9. The concept of *topic* referred to in this Figure is examined in Section 6.4.

10. Sereno 1991 offers a stimulating critique of interdisciplinary analogy.

11. A now-defunct game show featured a competition ('Name that Tune') between two finalists to name a well-known melody, the contestants gambling on being able to name the melody on hearing only a few of its notes, within a range of one to seven. Some, I believe, were able to name melodies having heard as few as the opening two notes. Accepting the element of fortune here, and the fact that the themes were played fully harmonized, some melodies clearly begin with distinctive three-note memes.

12. Abbreviations used in musical examples in this article are as follows: b. N = bar N; bb. N-N = bars N-N inclusive; I, II, etc. = movement number of a multi-movement work.

13. The up and down arrows and associated integers indicate the direction of melodic movement and its size in semitones (the octave, in western music, is divided into 12 semitones). It seems logical to suggest that intervallic structure, in addition to absolute pitch location, should be a criterion in the establishment of memetic equivalence classes. On this principle, the opening gesture of 'Misty' could be represented, independently of key, as ↓3↓5.

14. In many of the examples here, particularly Example 8, I have transposed pitch memes from passages in different keys into one key in order to facilitate comparison, but it is worth noting that a given meme in two different keys is essentially the same meme, in that, by definition, it retains its internal intervallic structure; what differs is the absolute pitch location.

15. In the passage from the 'Trout' Quintet, meme *y* appears in the key of A major, whereas in the Beethoven symphony it appears in C major, at the point where the melody modulates from the tonic (home) key of A minor to C, its *relative major* (i.e., the major key with the same *key signature*).
16. Distributional analysis is a method of analyzing music developed by musical semiologists such as Jean-Jacques Nattiez and Nicolas Ruwet in which, as with the analysis of language, music is segmented into its component phonetic units (*paradigmatic analysis*) and then the syntactical (grammatical) distribution of those units is determined (*syntagmatic analysis*).
17. The minuet movement of a Classical multi-movement work (such as a symphony or a string quartet) is tripartite, consisting of the minuet proper, a central *trio* section (i.e., a contrasting minuet), and a closing repeat of the minuet.
18. The 'parametric complex' discussed in Section 5.1 is, of course, a memplex composed of elements in different parameters; that considered here exists in only one parameter, that of pitch.
19. In the 1920s and 1930s Schenker developed a theory which conceived music as consisting of three interrelated levels (*Schichten*): the *foreground* (*Vordergrund*), the immediate musical surface; the *middle ground* (*Mittelgrund*), a deeper level revealed by the removal of superficial, decorative notes; and the *background* (*Hintergrund*), the most remote and fundamental level. Forte and Gilbert 1982 and Neumeyer and Tepping 1992 are good introductions to his thinking, presenting the essence of Schenker's seminal work, *Der freie Satz* of 1935 (1979), in accessible form. The following Schenkerian conventions are adopted here: the degrees of the scale are indicated using Arabic numerals topped by the caret symbol ( $\hat{1}$ ,  $\hat{3}$ ,  $\hat{4}$ , etc.); and important harmonic steps (*Stufen*) are indicated by Roman numerals (I, IV, V, etc.).
20. *A linear progression* (*Zug*) is a stepwise melodic motion generally spanning the intervals (gaps) of a third, fourth, fifth, sixth, or octave. The forward slash, in Schenkerian analysis, indicates both vertical alignment and structural synchrony. Thus,  $\hat{5}/\text{I}-\hat{4}\hat{3}$  means that whilst the tonic (I) chord is operative, the scale degree  $\hat{5}$  is sounded, followed by the descending motion  $\hat{5}\hat{4}-\hat{3}$ . In an *interruption* (*Unterbrechung*) - symbolized by "||" - the upper line descends to  $\hat{2}/\text{V}$  and then returns to the pitch from which it began its descent. This is then followed by a second descent which reaches the  $\hat{1}/\text{I}$  averted by the interruption.
21. Schenker proposed that all tonal pieces could ultimately be reduced to one of three *Ursätzen*, or *fundamental structures*. The *Ursatz* is a two-part structure, consisting of a lower voice or *Bassbrechung* (*bass arpeggiation*) outlining the harmonic progression I-V-I, and an upper voice or *Urlinie* (*fundamental line*), expressing the linear progression  $\hat{8}\hat{7}\hat{6}\hat{5}\hat{4}\hat{3}\hat{2}\hat{1}$ , or  $\hat{5}\hat{4}\hat{3}\hat{2}\hat{1}$ , or  $\hat{3}\hat{2}\hat{1}$ . It will be realized that the three *Ursätzen*, repeatedly reinstated by a variety of foreground and middle ground memes, are memes at the deepest levels of musical structure (see Section 6.2 for further treatment of this concept). A memetic view of musical structure tends, however, to overturn Schenker's view that "...all the foreground diminutions, including the apparent 'keys' arising out of the voice-leading transformations [of the middle ground], ultimately emanate from the diatony [i.e., the *Ursatz*] in the background" (1979: 11). Instead of this *top-down* perspective, memetics suggests the *bottom-up* generation of middle ground

structures by patterning at the foreground and, in turn, the generation of background structures by middle ground events.

22. From this discussion, it will be understood that 'Schenkerism' is a rich memplex of musical and extra-musical memes fronted by a set of graphical memes. As such, one must be aware that it filters, and inevitably colors, the memes that make up the music which is its subject and with which it interacts. In other words, the view of musical structure one formulates via Schenkerian analysis is, possibly, as much an artifact as those distortions which sometimes result from the treatment of samples in preparation for electron microscopy. Throughout this article, my reliance on Schenkerian memes is a testament to their virulent infective powers; no other comparable music-analytical memplex has such high fecundity at present. Nevertheless, I am fortunately also a host to other memes which impart to me some degree of critical detachment, and I hope the reader is likewise infected. For vigorous arguments *contra*-Schenker, see Narmour 1977.

23. See Meyer 1989: 50-54 for a consideration of such schemata, which "...are patterns that, because they are congruent both with human perceptual/cognitive capacities and with prevalent stylistic (musical and extra-musical) constraints, are memorable, tend to remain stable over time, and are therefore replicated with particular frequency" (1989: 51). Meyer also uses the term *archetype* to refer to structural classes which are the result of innate rather than learned responses; in practice the distinction is often difficult to draw (Gjerdingen 1988: 46 f.).

24. For a comprehensive survey, see Bent and Drabkin 1987, Chapters 2 and 3; see also note 22.

25. It is worth observing that in distinguishing between the process of memetic conglomeration which gives rise to the individual work of art and the long process by which the structural archetype instantiated by that work evolves, we mirror the clear biological parallel between *ontogeny* - the (embryonic) development of the individual organism - and *phylogeny* - the development of the species of which the organism is, in a sense, an instantiation.

26. Binary form consists of two repeated sections, the first of which (A) moves from the tonic key to a contrasting key, the second (B) from this contrasting key back to the tonic. Ternary form consists of three sections, the first (A) being a closed unit, the second (B) offering a strong contrast, and the third being either identical to A or a modified version ( $A^1$ ) of A. Sonata form may be understood as an evolutionary descendant of binary form in which the B section of the binary form becomes divided into two distinct phrases, the second of which *recapitulates* the A part of the form ( $A | B A^1$ ). This summary is, however, highly oversimplified: see Rosen 1988 for a full discussion of the relationships between eighteenth-century binary forms and sonata forms.

27. A normative prolongational structure is "...specified in terms of relationships of prolongational connection among structurally significant events, not in terms of any specified sequence of events" (Lerdahl and Jackendoff 1983: 248).

28. Constraints of space prevent my addressing the burgeoning neuro-scientific literature on music, central as this will be to further research in musical memetics. See, especially, Bharucha 1999, in which a neural net model of musical organization is examined. Such models "...have a

number of properties that recommend them as models of music cognition....they can account for how we learn musical patterns....their assumptions are either known or plausible principles of neuroscience....they are capable of recognizing varying shades of similarity and are therefore well suited to modeling similarity-based accounts...of tonality or modality....[and] they can discover regularities in musical styles that may elude formal music-theoretic analysis..." (1999: 413). Moreover, "...a neural net can learn temporal composite patterns so that they function as schemas and as sequential memories" (1999: 424).

29. The particular signifier 'dominant seventh chord' would, of course, be propagated principally in the brains of Anglophones. Owing to language differences, the same signified musical meme would have been associated with different signifying concept memes in the brains of, for instance, Mozart, Rameau, and Vivaldi.

30. This account clearly oversimplifies matters considerably by ignoring the motor-control memes which govern the muscular actions engendering writing, speaking, and the production of musical sounds.

31. Another example is the coadaptation between the rhythmic meme  (seen in Example 2 i and iii, where it is labeled meme y) and verbal-conceptual memes articulating notions of pompous militarism.

32. To speak of a single gene's or meme's fitness is somewhat artificial: these replicators are selected and propagated as complexes and exert their pheno/phemotypic effects in concert. A given pheno/phemotypic effect may be engendered by several genes or memes; conversely, several effects may be brought about by a single gene or meme. To use terms from population biology, musical memes would appear to be subject to *r*-selection, which is "[s]election for the qualities needed to succeed in unstable, unpredictable environments..." (Dawkins 1983a: 293). *K*-selection, by contrast, is "[s]election for the qualities needed to succeed in stable, predictable environments..." (1983a: 288).

33. Nattiez defines these three levels as follows: "(a) *The poetic dimension*: even when it is empty of all *intended* meaning...the symbolic form results from a *process of creation* that may be described or reconstituted. (b) *The esthetic dimension*: 'receivers,' when confronted by a symbolic form, assign one or many meanings to the form.... (c) *The trace*: the symbolic form is embodied physically and materially in the form of a *trace* accessible to the five senses. We employ the word *trace* because the poetic process cannot immediately be read within its lineaments, since the esthetic process...is heavily dependent upon the lived experience of the 'receiver.' Molino proposed the name *niveau neutre* [neutral level] or *niveau matériel* [material level] for this trace" (1990: 11-12; his emphases).

34. Meyer is using the term *mutational* in a different sense to that used in this article: a memetic standpoint regards *trended* changes as just as much the result of mutations as *mutational* changes.

35. Serialism is the organization of the twelve pitches of the chromatic scale into a relationship of equal status. It is a system of organization developed by Arnold Schoenberg (1874-1951) by the early 1920s and employed in the works of many later twentieth-century composers.

36. Thus, whilst many commentators see the appearance of the serial system as a revolutionary change, in the gradualist interpretation advocated here it is an evolutionary transformation grounded in antecedent developments, particularly the steady attenuation of tonality in the decades after Wagner's *Tristan und Isolde* (1865), and its eventual disintegration into "free atonality" in the works of progressive composers c. 1910.

37. Studies of influence - now often subsumed under Kristeva's term *intertextuality* (1969) - in music, such as Korsyn 1991 and those sources listed in his note 5, are increasingly common in the musicological literature. To my knowledge, however, none invokes the memetic paradigm.

## Acknowledgements

My thanks are extended to Dr. Yo Tomita of Queen's University Belfast for the use of his *Bach* musicological font. The *Bach* website is at <http://www.music.qub.ac.uk/~tomita/bach-mf.html>. The music examples were set using *Sibelius* by Ian Phillips-Kerr at *MusicWorks* (ipk@musicworks2000.freeserve.co.uk).

## References

- Agawu, V.K. (1991) *Playing With Signs: A Semiotic Interpretation of Classic Music* (Princeton: Princeton University Press).
- Allanbrook, W.J. (1983) *Rhythmic Gesture in Mozart: "Le nozze di Figaro" and "Don Giovanni"* (Chicago: University of Chicago Press).
- Ball, J.A. (1984) Memes as Replicators, *Ethology and Sociobiology* 5:145-161.
- Bent, I.D. and Drabkin, W. (1987) *Analysis* (The New Grove Handbooks in Music) (London: Macmillan).
- Bharucha, J.J. (1999) Neural Nets, Temporal Composites, and Tonality, in *The Psychology of Music* (Academic Press Series in Cognition and Perception), 2nd edn., ed. Deutsch, D. (San Diego: Academic Press): 413-440.
- Blackmore, S.J. (1999) *The Meme Machine* (Oxford: Oxford University Press).
- Brodie, R. (1996) *Virus of the Mind: The New Science of the Meme* (Seattle: Integral Press).
- Cavalli-Sforza, L.L. and Feldman, M.W. (1981) *Cultural Transmission and Evolution: A Quantitative Approach* (Monographs in Population Biology, no. 16) (Princeton: Princeton University Press).

- Cloak, F. (1973) Is a Cultural Ethology Possible?, *Human Ecology* 3: 161-182.
- Cook, N. (1990) *Music, Imagination, and Culture* (Oxford: Clarendon Press).
- Cook, N. (1994) *A Guide to Musical Analysis* (Oxford: Oxford University Press) (repr. from London: Dent, 1987).
- Daverio, J. (1997) *Robert Schumann: Herald of a "New Poetic Age"* (New York and Oxford: Oxford University Press).
- Dawkins, R. (1983a) *The Extended Phenotype: The Long Reach of the Gene* (Oxford: Oxford University Press) (repr. from London: W.H. Freeman and Company Ltd., 1982).
- Dawkins, R. (1983b) Universal Darwinism, In *Evolution from Molecules to Men*, ed. Bendall, D.S. (Cambridge: Cambridge University Press): 403-425.
- Dawkins, R. (1989) *The Selfish Gene*, 2nd. edn. (Oxford: Oxford University Press) (from Oxford: Oxford University Press, 1976).
- Dawkins, R. (1991) *The Blind Watchmaker* (London: Penguin Books) (repr., with an Appendix, from London: Longman, 1986).
- Dawkins, R. (1993) Viruses of the Mind, in *Dennett and His Critics: Demystifying Mind*, ed. Dahlbom, B. (Oxford: Blackwell): 13-27.
- Delius, J.D. (1989) Of Mind Memes and Brain Bugs: A Natural History of Culture, in *The Nature of Culture: Proceedings of the International and Interdisciplinary Symposium, October 7-11, 1986 in Bochum*, ed. Koch, W.A. (Bochum: Brockmeyer): 26-79.
- Dennett, D.C. (1995) *Darwin's Dangerous Idea: Evolution and the Meanings of Life* (London: Penguin Books).
- Dunsby, J. and Whittall, A. (1988) *Music Analysis in Theory and Practice* (London: Faber Music).
- Forte, A. and Gilbert, S.E. (1982) *Introduction to Schenkerian Analysis* (New York: Norton).
- Gatherer, D. (1997) The Evolution of Music: A Comparison of Darwinian and Dialectical Methods, *Journal of Social and Evolutionary Systems* 20/1: 75-92.
- Gjerdingen, R.O. (1988) *A Classic Turn of Phrase: Music and the Psychology of Convention* (Philadelphia: University of Pennsylvania Press).
- Goehr, L. (1992) *The Imaginary Museum of Musical Works: An Essay in the Philosophy of Music* (Oxford: Clarendon Press).

- Hopkins, A. (1981) *The Nine Symphonies of Beethoven* (London: Heinemann).
- Jan, S.B. (1995) *Aspects of Mozart's Music in G Minor: Toward the Identification of Common Structural and Compositional Characteristics* (London and New York: Garland).
- Jan, S.B. (2000) The Selfish Meme: Particularity, Replication, and Evolution in Musical Style, *International Journal of Musicology* 8 (in press).
- Kerman, J. (1985) *Musicology* (London: Fontana).
- Kinderman, W. (ed.) (1991) *Beethoven's Compositional Process* (Lincoln, Nebraska: University of Nebraska Press).
- Korsyn, K.E. (1991) Towards a New Poetics of Musical Influence, *Music Analysis* 10/1, 2:3-72.
- Kristeva, J. (1969) *Semiotiké* (Paris: Seuil).
- Lerdahl, F. and Jackendoff, R. (1983) *A Generative Theory of Tonal Music* (Cambridge, Massachusetts: The Massachusetts Institute of Technology Press).
- Lynch, A. (1996) *Thought Contagion: How Belief Spreads Through Society - The New Science of Memes* (New York: Basic Books).
- Lynch, A. (1998) Mnemon 1998a: Y2K Memes (Issue 1).  
<<http://www.mcs.net/~aaron/Mnemon1998a.html>>
- Meyer, L.B. (1989) *Style and Music: Theory, History, and Ideology* (Philadelphia: University of Pennsylvania Press).
- Narmour, E. (1977) *Beyond Schenkerism: The Need for Alternatives in Music Analysis* (Chicago: University of Chicago Press).
- Nattiez, J.-J. (1990) *Music and Discourse: Toward a Semiology of Music*, tr. Abbate, C. (Princeton: Princeton University Press).
- Neumeyer, D. and Tepping, S. (1992) *A Guide to Schenkerian Analysis* (Englewood Cliffs: Prentice-Hall).
- Plotkin, H.C. (1995) *Darwin Machines and the Nature of Knowledge: Concerning Adaptations, Instinct and the Evolution of Intelligence* (London: Penguin Books).
- Pocklington, R. and Best, M.L. (1997) Cultural Evolution and Units of Selection in Replicating Text, *Journal of Theoretical Biology* 188/1:79-87.
- Ratner, L.G. (1980) *Classic Music: Expression, Form, and Style* (New York: Schirmer).

- Rosen, C. (1988) *Sonata Forms*, 2nd, revd., edn. (New York: Norton).
- Sadie, S. (ed.) (1980) *The New Grove Dictionary of Music and Musicians*, 6th edn. (London: Macmillan).
- Schenker, H. (1979) *Free Composition*, tr., ed. Oster, E. (New York: Longman).
- Scruton, R. (1997) *The Aesthetics of Music* (Oxford: Clarendon Press).
- Sereno, M.I. (1991) Four Analogies Between Biological and Cultural/Linguistic Evolution, *Journal of Theoretical Biology* 151: 467-507.
- Yeston, M. (1976) *The Stratification of Musical Rhythm* (New Haven: Yale University Press).

# MATHEMATICAL MODELS FOR MEMETICS

By

**Jeremy R. Kendal and Kevin N. Laland**

Sub-Department of Animal Behaviour, Department of Zoology,

University of Cambridge, Madingley, Cambridge CB3 8AA

Tel.: 01954 210301 Fax: 01954 210247

[jrk31@cam.ac.uk](mailto:jrk31@cam.ac.uk)

## Abstract

The science of memetics aims to understand the evolution of socially transmitted cultural traits. Recently attention has focused on the interaction between memetic and genetic evolution, a phenomenon described as meme-gene co-evolution. Whether cultural evolution occurs purely at the level of the meme, or through meme-gene interaction, a body of formal theoretical work already exists that can be readily employed to model empirical data and test theoretical hypotheses. This is cultural evolution and gene-culture co-evolutionary theory, a branch of theoretical population genetics (Cavalli-Sforza & Feldman [6]; Boyd & Richerson [3]; Feldman & Laland [12]). We reject the argument that meaningful differences exist between memetics and these population genetics methods. The goal of this article is to point out the similarities between memetics and cultural evolution and gene-culture co-evolutionary theory, and to illustrate the potential utility of the models to memetics. We illustrate how the theory can be applied by developing a simple illustrative model to test a hypothesis from the memetics literature.

**Keywords:** brain size, cultural evolution, gene-culture co-evolution, meme, memetics

## 1 Introduction

Memetics endeavours to understand culture in terms of the evolution of socially transmitted cultural traits. Recent instantiations have acknowledged that memetic and genetic evolution can interact in rich and complex ways, a phenomenon described by Blackmore [2] as "meme-gene co-evolution". Memetics has developed in virtually complete isolation from a related discipline with very similar concerns, namely, cultural evolution and gene-culture co-evolution. Cultural evolution is a branch of theoretical population genetics established in the early part of the 1970s, and dedicated to using population genetics models to investigate the evolution and dynamics of cultural traits equivalent to memes. Gene-culture co-evolution employs the same methods to explore the co-evolution of genes and cultural traits. Whether meme evolution occurs exclusively at the cultural level, or through meme-gene interaction, a body of formal theoretical work already exists that can be readily employed to model empirical data and test theoretical hypotheses (Cavalli-Sforza & Feldman [6]; Boyd & Richerson [3]; Feldman & Laland [12]). The goal of this article is to point out the parallels between the logic and assumptions of memetics and cultural evolution/gene-culture co-evolutionary theory, and to illustrate the potential utility of the models to memetics.

We begin by pointing to the similarities between memetics and cultural evolution theory. Both

disciplines explicitly assume that at least some aspects of human culture can be broken down into discrete traits, or component elements, henceforth memes (although both approaches have also explored the evolution of continuous traits). Both disciplines believe that memes can increase or decrease in frequency through the processes of Darwinian evolution (i.e., the differential survival and reproduction of individuals), as well as through similar or analogous processes affecting cultural selection (i.e., the differential survival and reproduction of memes). Both disciplines stress that memes should be treated as replicators in their own right, and most proponents of each would probably accept that the adoption of at least some memes is not independent of the individual's genotype. We suspect that most advocates of the population genetics approach would be sympathetic to the argument that many of the dynamics of contemporary human culture are best understood at the level of the meme. The allocation of parameters to describe rates of cultural transmission in cultural evolution and gene-culture co-evolution models is formally equivalent to treating cultural transmission based on cultural replicators that are distinct from genes (Feldman & Cavalli-Sforza [10]; Cavalli-Sforza & Feldman [6]). We also suspect that many memeticists would be sympathetic to Feldman & Laland's [12] claim that much human evolution over the last 2 million years can be best understood as gene-culture (or meme-gene) co-evolution. Indeed the basic philosophy and parallels between the two approaches are so similar, it is striking that they have not converged.

We suggest that there are two main reasons why memetics has not yet embraced cultural evolution, neither of which is of sufficient validity to warrant their continued divergence. The first reason is that cultural evolution and gene-culture co-evolutionary theory are relatively small and specialized branches of population genetics, with a lot of the research published in highly-rated yet somewhat esoteric journals, and containing material of an extremely technical and mathematical nature, but which rarely adopts the term meme, or refers to the memetics literature. As a result of this, many meme enthusiasts are simply unaware of this population genetics literature, and among those that are, many find it difficult to comprehend. This article is written with the belief that these technicalities and linguistic differences should not be allowed to drive a wedge between two disciplines that otherwise have such similar goals, and for which the opportunities for cross-fertilisation of ideas and methods are considerable. It is quite clear that memetics as a discipline is not at all hostile to theory, and that many meme enthusiasts are sufficiently mathematically sophisticated to understand the theoretical papers. One of the goals of this paper is to introduce cultural evolution methods to a memetics audience.

A second reason why memetics has distanced itself from cultural evolution theory is the belief among meme enthusiasts that advocates of cultural evolutionary theory assume genes or genotypes to control the adoption and transmission of memes. To use Lumsden and Wilson's [17] metaphor, memeticists think that cultural evolutionists believe memes are on a genetic leash. This is not the case. Although Lumsden and Wilson used the terms cultural evolution and gene-culture co-evolution, their work had more in common with the earlier sociobiologists than with modern cultural evolution and gene-culture co-evolutionary theory. The population genetics approach to culture grew out of a rejection of sociobiology. This rejection stems, in part, from sociobiology's failure to take cultural transmission seriously, or to recognise that not all cultural traits are to the ultimate benefit of the genes. Cultural evolution and gene-culture co-evolutionary models have developed a system of formalism that allows for leashes of variable length, ranging from extremely tight genetic control, to no meaningful genetic influence at all (see Feldman &

Laland [9], for a discussion of this issue). The models do this by allocating parameters that determine the probability that an individual will acquire a particular meme that may, or may not, differ according to the individual's genotype. Moreover, gene-culture co-evolutionary theory has long explicitly asserted that the leash tugs both ways, and that under some circumstances genes can be under cultural control (Feldman & Cavalli-Sforza [10]; Feldman & Laland [9]). If contemporary practitioners of gene-culture co-evolutionary theory no longer ask Dennett's [7] *Cui bono?* question, it is because two decades earlier they developed their own method of answering this question, namely, the *phenogenotype*. We describe how phenogenotypes can be used to explore meme-gene co-evolution below. For the moment it is sufficient to note that there is nothing in cultural evolution theory that precludes cultural adaptations "for the good of the meme", or a "meme's-eye-view" approach.

We reject the argument that meaningful differences exist between memetics and the population genetics methods. We also believe that cultural evolution and gene-culture co-evolutionary theory will be much enriched by embracing memetics. *Prima facie* complex social phenomena, for example, institutions or belief systems, may be treated as memplexes made up of a plethora of attributes, many of which are memes in their own right, each with varying distribution and dynamics (Plotkin [18]; Blackmore [2]). Social anthropologists limit their ability to carry out predictive and comparative analyses, by insisting on an all-encompassing, holistic approach, rather than breaking culture down into manageable units, and performing rigorous quantitative analyses. We do not refute the idea that the distributions of cultural traits may be complex, but suggest that real progress can be made by addressing those cultural traits that have manageable, easily identified and quantifiable distributions. In the case of our worked example, imitate the imitators, we are interested in the presence or absence of a particular cultural trait, therefore it is appropriate to consider discrete memes. Part of the excitement generated by the memetics movement surely reflects the fact that the meme concept by its very existence operationalises culture, breaking it down into the kind of discrete packages that can be subject to scientific enquiry. We believe that mathematical analyses, using cultural evolution and gene-culture co-evolutionary models, can be profitably employed in restricted domains, where cultural traits or memes are easily defined and quantitative studies are possible. We suggest that the techniques of cultural evolution and gene-culture co-evolution can be employed to describe and predict how cultures change, explain current trends in cultural attributes, and test hypotheses concerning the evolution of human behaviour and culture.

## 2 Introducing an Established Modelling Paradigm

Let us return to Dennett's *Cui bono?* question. Gene-culture co-evolutionary theory has devised its own simple and pragmatic solution to the question of whether genes or memes control cultural (and genetic) change.

Population geneticists recognise that a salient model should be developed at the level of the smallest non-divisible unit, as this is the natural currency by which changes in replicator frequency will occur. For illustration, imagine the co-evolution of two genetic loci subject to natural selection in a population of animals. We will call the loci **A** (with alleles *A* and *a*) and **B** (with alleles *B* and *b*). A two-locus population genetics model would have to track the frequencies of each of the four types of chromosomal combination (*AB*, *Ab*, *aB*, and *ab*), usually

describing them as gametes. The two loci cannot be assumed to be independent because there may be complex interactions or non-random associations between loci, depending on the nature of the selective regime. Subsequent analysis of the system will tell us whether allele *A* is favoured over allele *a*, and describe the evolution at the **B** locus. The same applies to gene-culture co-evolution, except here one of the loci (say **B**) contains variants of meme (meme variants *B* and *b*) rather than alleles of a gene. The (rather cumbersome) name given to particular combinations of genotype and meme is the phenogenotype. In population genetics models, the genotypes are allocated fitnesses and the rate of transmission is tracked using recursions which describe the frequency of each genotype in terms of the frequencies of genotypes in the previous generation. Likewise, for the gene-culture co-evolutionary model, each phenogenotype is allocated a "fitness", representing its rate of reproduction. However, while genes are transmitted to each individual once a generation, meme transmission is more complex, and may occur on a within-generational timeframe, depending on who acquires the meme from whom. Nonetheless, it is frequently mathematically convenient to track rates of change of phenogenotypes using a generational timeframe. Of course, individuals change their memes frequently throughout their lifetimes, and this will be reflected in changing frequencies of phenogenotypes. In tracing the fortunes of the meme, the model only makes assumptions as to the path of transmission and not the processes that underlie replication. As long as the meme is observable and measurable, analysis of the dynamics of the system can be used in conjunction with empirical data to question the nature of these assumptions, explore the replication process, and test predictions. While there may be nothing directly equivalent to a 'locus' or 'allele' for memes, we believe that sufficient memes compete directly with each other for this to be a useful abstraction.

The gene-culture co-evolutionary framework embraces the full range of selective influences upon the evolution of memes: memetic fitness may be quite independent of genetic influence (there is no genetic leash); memes may influence the fitness of genes (genes are held on a memetic leash); or genetic fitness can determine the success of the meme (memes are held on a genetic leash). For purposes of exposition, we introduce the use of the gene-culture co-evolutionary methods under each of these circumstances.

If cultural inheritance is completely independent of genetic inheritance, that is, if memes are acquired independently of the learner's genotype, then the unit of cultural selection collapses to that of the meme. Under such circumstances, tracking gene frequencies is of no value, gene-culture co-evolutionary methods are inappropriate, and all evolution occurs at the level of the meme. It is here that cultural evolution methods are useful. Cultural evolutionary theory is a modelling paradigm developed by Cavalli-Sforza & Feldman and others to analyse cultural change (Feldman & Cavalli-Sforza [10]; Cavalli-Sforza & Feldman [6]; Boyd & Richerson [3]). This framework has been used to explore phenomena as diverse as linguistics, epidemics, social values and customs, and the diffusion of innovations (i.e. the same subject matter as memetics). Cultural evolutionary models do not use the term meme because they were first developed prior to the publication of *The Selfish Gene*. Instead they use the phrase "cultural trait" or "cultural phenotype". Nonetheless, Cavalli-Sforza and Feldman's methods effectively track changes in meme frequency in a population through cycles of mating, differential social transmission of the meme (cultural selection) and natural selection. The modes of social transmission describe the routes by which memes spread (Cavalli-Sforza & Feldman [6]). Social transmission can occur

vertically (from parents to offspring), obliquely (from parental to offspring generation) or horizontally (within generation transmission), or some combination of these.

Cavalli-Sforza & Feldman define cultural selection as a Darwinian process by which particular memes increase or decrease in frequency due to their differential probability of being adopted by other individuals. In contrast, for these authors natural selection refers to the differential survival of individuals expressing different types of memes. For instance, the spread of contraceptive use through cultural selection processes could alter natural selection pressures induced by sexually transmitted diseases. Therefore the analysis is based on the explicit assumption that cultural traits evolve by Darwinian selection processes, whereby individuals can be selected purely on the basis of their memes. This allows interesting questions to be asked such as, 'Under what conditions might non-adaptive cultural traits evolve?', (Cavalli-Sforza & Feldman [6]). One of the earliest findings to emerge from this analysis is that there are countless circumstances when memes with sufficiently high cultural fitness can increase in frequency despite being maladaptive from a genetic perspective, and decreasing genetic fitness (Feldman & Cavalli-Sforza [10]). Cavalli-Sforza and Feldman's framework also considers cases in which cultural selection operates without affecting Darwinian fitness (e.g. a preference for a particular soft drink).

The models' strengths are in their simple formalization, the ability to predict the frequencies of cultural traits, and the direct applicability to empirically derived transmission coefficients.

Cavalli-Sforza and Feldman [6] demonstrated that models can be applied to empirical data, estimating the transmission coefficients for five vertically transmitted traits that would be of interest to meme enthusiasts: salt usage, frequency of praying to God, frequency of swimming, belief in ability versus luck, and political interest. The models can be used to describe patterns of change, make sense of meme variation, extrapolate back into the past to predict past cultural usage, and extrapolate into the future to predict future trends. We have only begun to describe the breadth of cultural evolution models, and countless elaborations and refinements of the methods have emerged, for example, investigating frequency-dependent transmission, the evolution of continuous (quantitative) traits, and the effects of migration, population structure, innovation, and other factors (Cavalli-Sforza & Feldman[6]; Boyd & Richerson[3]; see Laland[15] for a simple introduction).

There are also situations where there is an *interaction* between memes and genes, such that an individual's propensity to adopt a particular meme depends on his or her genotype. This is likely to be more common in investigations of human evolution than in explorations of contemporary human culture. Under such circumstances where there may be meme-gene co-evolution, gene-culture co-evolutionary methods are appropriate. For example, Feldman and Cavalli-Sforza [11] used gene-culture co-evolutionary theory to investigate the evolution of lactose absorption. By defining genotypes that differ in terms of their ability to process lactose, and by describing individuals as either having a meme for milk consumption or not, Feldman and Cavalli-Sforza were able to develop a population genetics model to explore how dairy farming and milk use might co-evolve with genes for lactose absorption. In this case, the cultural trait is treated as a discrete meme (milk user or nonuser), rather than as a continuous distribution, which simplifies the model without compromising the qualitative results. The analysis suggested that whether or not the allele for absorption achieves a high frequency depends critically on the probability that the children of milk users themselves adopt the meme. The analysis is able to account for both the spread of lactose absorption, and the culturally related variability in its incidence. Moreover,

Feldman & Cavalli-Sforza found a broad range of conditions under which the absorption allele does not spread despite a significant fitness advantage. Meme transmission complicates the selection process to the extent that the outcome may differ from that expected under purely genetic transmission.

In the lactose absorption example, the leash tugs both ways: whether or not an individual drinks milk depends on his or her genotype, but also whether or not the genes for lactose absorption are present in a population depends on the history of dairy farming in that culture. Such two-way influences are likely to have been common throughout recent human evolution, and may have been particularly important in the evolution of the human brain. A number of theories have related the change in brain size and associated anatomical effects with cultural change (Humphrey [14]; Holloway [13]; Wilson [20]; Byrne & Whiten [5]; Dunbar [8]; Aiello & Wheeler [1]). For example, the evolution of the brain with a costly, high mass specific, metabolic rate may have been paid for by a reduction in the size of the gastro-intestinal tract (Aiello & Wheeler [1]). Aiello and Wheeler [1] suggested that a reduction in gut size was only possible because the larger brains were capable of culturally mediated improvement in diet, in proportion to the loss of gut. Humphrey [15] emphasised the effect of social behaviour and hierarchy upon the evolution of brain size, resulting in the acceleration of cultural evolution. Wilson et al. [20] suggested the behavioural drive hypothesis, predicting that the dual capacity for behavioural innovation and social propagation of new habits accelerated anatomical evolution. Gene-culture co-evolutionary theory is suited to investigate the plausibility of such hypotheses.

One recent interesting addition to hypotheses concerned with brain evolution has been made by Blackmore [2]. Blackmore suggests that social learning and memetic drive are the cause behind the evolution of increased brain size, proposing a three stage process. Firstly, individuals with a predisposition for imitating (we will call them 'generalist' imitators, as they are not biased to learn from any particular individual) would succeed over those that only learn directly from the environment. Secondly, in a population of imitators, those with a predisposition to imitate from the best imitators would be selected (we will call them 'selective' imitators). Thirdly, preferential mating between enhanced imitators would produce the most successful offspring. The selective imitators are assumed to require a greater than average cognitive capacity, therefore there is selection for an increase in the average brain size. As this hypothesis comes directly from the memetics literature, in order to illustrate the gene-culture co-evolutionary method we show how one aspect of Blackmore's hypothesis can be tested using a phenogenotype model.

### 3 Imitate the Imitators

The first stage of Blackmore's model can already be regarded as to some extent supported by several gene-culture co-evolutionary models exploring the circumstances under which natural selection should favour reliance on social learning (Boyd & Richerson [3, 4]; Laland, Richerson & Boyd [16]; Feldman, Aoki & Kumm [9]). Here we concentrate on exploring the plausibility of stage two. We start with a population of 'generalist imitators', who show no bias to learn from any particular individual. We ask if a small number of individuals with a more selective capacity for imitation, in that they concentrate on imitating only other individuals with this selective tendency, would be at a selective advantage. If so, genes for imitating the imitators selectively would be expected to spread through the population. Blackmore assumes that the more selective

imitators will by definition be the best imitators, and will be more likely to adopt advantageous memes than generalist imitators. Here we show that this assumption is not necessary for selective imitation to spread, although it could easily be incorporated into the model and would almost certainly increase the likelihood that selective imitation would be favoured. Although we take issue with Blackmore's claim that imitation is the only process that can support meme transmission (Reader & Laland [19]), to accurately reflect Blackmore's position, here we restrict our use of the term imitation to the direct copying of a motor pattern. As the model considers the *consequences* of social learning, rather than the underlying psychological processes, it applies equally to any social learning processes, to the extent that they support meme transmission.

We assume that individuals adopt one of two memes ( $B_1$  or  $B_0$ ) with the assumption that meme  $B_1$  is manifest as a novel behaviour that enhances the individual's fitness by  $\gamma_2$  [note 1], relative to  $B_0$ .  $B_1$  may, for example, represent the ability to exploit a new, highly nutritious food source. We define as generalist imitators individuals of genotype  $ii$ , while selective imitators are of genotype  $Ii$ . The heterozygote,  $Ii$ , has intermediate imitative tendencies, depending on the level of dominance of  $I$  over  $i$ , given by the coefficient,  $b$ . As Blackmore suggests that individuals who can imitate the best imitators require a greater brain size, we attribute a fitness cost,  $\gamma_1$  to the  $I$  allele, reflecting the extra neural investment, but nonetheless leave open the possibility that  $\gamma_1$  may be set to zero. With three possible genotypes and two memes, there are six possible associations or phenogenotypes. Each phenogenotype has an associated fitness (see Table 1).

Genotype Meme	$II$	$Ii$	$ii$
$B_1$	$1 - \gamma_1 + \gamma_2$	$1 - \beta\gamma_1 + \gamma_2$	$1 + \gamma_2$
$B_0$	$1 - \gamma_1$	$1 - \beta\gamma_1$	$1$

**Table 1:** The fitnesses of the six phenogenotypes. Each fitness, reflecting the relative change in that phenogenotype frequency per generation with respect to selection pressures, is in the cell corresponding to the combination of the particular genotype (defined by column) and meme (defined by row).

Genotype	$II$	$Ii$	$ii$
Probability of $B_1$	$p1 = \frac{F_{1I}}{(F_{1I} + F_{0I})}$	$p2 = c \left[ \frac{\beta F_{1I}}{(F_{1I} + F_{0I})} + (1 - \beta)(F_{1I} + F_{0I}) \right]$	$p3 = c(F_{1I} + F_{0I})$

**Table 2:** The probability of acquiring the  $B_1$  meme given genotype of the individual, where  $F_{pg}$  is the frequency of the phenogenotype in the parental generation. Note that the probability of acquiring the  $B_0$  meme for an individual of a given genotype is just  $1 - Pr(\text{acquiring the } B_1 \text{ meme})$ .

After random mating with respect to the two memes and two genotypes, the  $B_1$  meme is passed down from the parental generation to the offspring generation in a manner dependent on its genotype (see Table 2).

The probability that an individual acquires the  $B_1$  phenotype depends upon its genotype, on the probability of social learning,  $c$  and the frequencies of  $B_1$  individuals of the parental generation. Whereas generalist imitators can potentially learn from all  $B_1$  individuals in the parental generation, selective imitators constrain themselves to copying the sub-population of selective imitators in the parental generation. In either case, the probability of adopting the  $B_1$  meme is directly proportional to its frequency in the relevant population.

Using the assumptions in tables 1 and 2, we can construct recursions defining the relationship between the frequencies of each phenogenotype in the parental and offspring generations. In each offspring generation, phenogenotype frequencies depend upon the probability of acquiring the genotype following random mating in the parental generation, the probability of acquiring the particular phenotype given the genotype (Table 2), and the fitness associated with the phenogenotype in question (Table 1). In this case, the system can be reduced from six phenogenotype recursions to a simpler system, with the following set of four recursions (1a-d). These recursions give the frequency of the  $I$  and  $i$  alleles in individuals with  $B_1$  and  $B_0$  memes, respectively.

(1a)

$$WF'_{1I} = (F_{1I} + F_{0I}) [p_1(F_{1I} + F_{0I})(1 - \gamma_1 + \gamma_2) + p_2(F_{1i} + F_{0i})(1 - \beta\gamma_1 + \gamma_2)]$$

(1b)

$$WF'_{0I} = (F_{1i} + F_{0i}) [p_3(F_{1i} + F_{0i})(1 + \gamma_2) + p_2(F_{1I} + F_{0I})(1 - \beta\gamma_1 + \gamma_2)]$$

$$(1c) WF'_{1i} = (F_{1I} + F_{0I}) [(1 - p_1)(F_{1I} + F_{0I})(1 - \gamma_1) + (1 - p_2)(F_{1i} + F_{0i})(1 - \beta\gamma_1)]$$

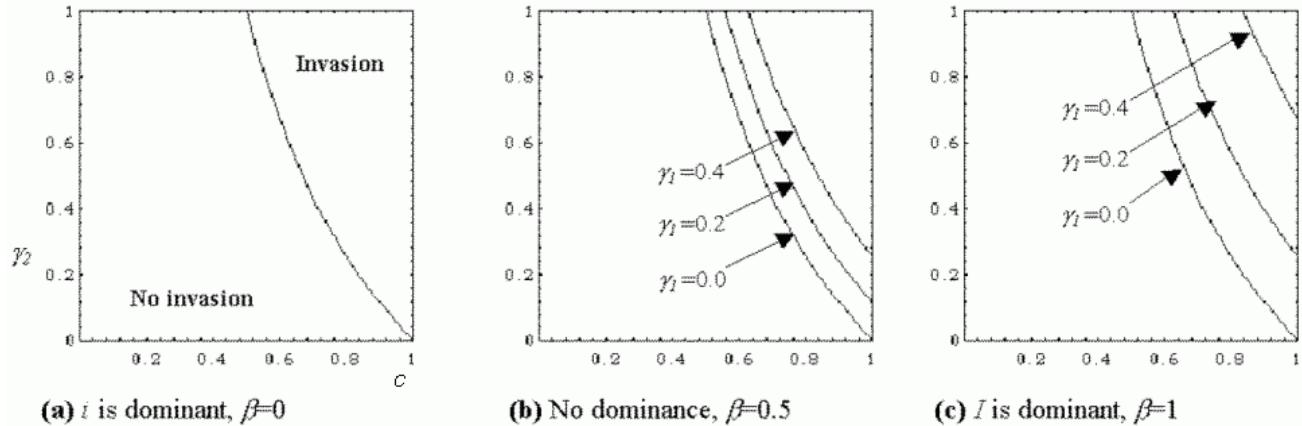
(1d)

$$WF'_{0i} = (F_{1i} + F_{0i}) [(1 - p_3)(F_{1i} + F_{0i}) + (1 - p_2)(F_{1I} + F_{0I})(1 - \beta\gamma_1)]$$

where  $F'_{pg}$  is the frequency in the offspring generation, and  $W$  is the sum of all the right hand sides of equations 1a-d.

Analysing this mathematical system [note 2] reveals that whether selective imitators evolve in a population of generalist imitators depends on whether the following inequality is satisfied, namely

$$(2) \left| c^{\frac{1}{2}} [1 - \beta\gamma_1(1 - \gamma_2) + \gamma_2]^{\frac{1}{2}} \right| > 1$$



**Figure 1:** The conditions under which a population of  $B_0i$  individuals is susceptible to invasion from the  $I$  allele under different magnitudes of fitness cost,  $\square_I$  and levels of dominance (a-c), where  $\square_2$  is the fitness advantage of meme  $B_I$ , and  $c$  is the probability of social learning.

Invasion occurs in populations with parameters to the right of the curved line.

The conditions under which a population of  $B_0i$  individuals is susceptible to invasion from the  $I$  allele under different magnitudes of fitness cost,  $\square_I$  and levels of dominance (a-c), where  $\square_2$  is the fitness advantage of meme  $B_I$ , and  $c$  is the probability of social learning. Invasion occurs in populations with parameters to the right of the curved line.

This result confirms Blackmore's intuition that there are conditions under which the selective imitation  $I$  allele can invade the population. These conditions are illustrated in Figure 1. It turns out that invasion is most likely when the fidelity of social learning,  $c$  and the fitness advantage of the  $B_I$  meme,  $\square_2$  are high, whilst the dominance of  $I$  over  $i$ ,  $\square$  and the cost of selective imitation,  $\square_I$  are low. The conditions for invasion rely most heavily upon the reliability of social learning,  $c$ , that is, the fidelity of meme transmission, with no chance of invasion if  $c < 0.5$ . However, it is possible that a high fidelity for meme transmission would already exist if we assume part 1 of Blackmore's hypothesis that effective imitation has already evolved in the population. The result shows clearly that if  $\square \square_I$  is high, there is less chance of invasion, as the inequality is less likely to be satisfied. If  $\square \square_I=1$  the left hand side of inequality 2 reduces to 0, making it impossible for allele  $I$  to invade. Interestingly,  $I$  is less likely to invade if it is dominant over  $i$ .

It is possible for selective imitators to evolve if either  $\square$  or  $\square_I$  is small, even if the other is large. That is, if there were little cost to selective imitation, it might evolve even if allele  $I$  is strongly dominant. Likewise, if allele  $i$  is strongly dominant, selective imitators could evolve even with a high cost. The effect of these scenarios is shown clearly in inequality 3. If we assume high transmission fidelity of the memes ( $c \approx 1$ ) and  $\beta \gamma_1 \approx 0$ , then the condition for the evolution of selective imitators merely requires a fitness advantage to  $B_I$  over  $B_0$  ( $\square_2 > I$ ),

$$(3) \left| \left(1 + \gamma_2\right)^{\frac{1}{2}} \right| > 1$$

It is also worth stressing that if inequality 3 is not satisfied then selective imitators will not be favoured in a population of generalist imitators, and that there is an equivalent criterion by which generalists could invade a population of selective imitators. Were we to make the assumption that selective imitators are more likely to acquire  $B_1$  than generalists (e.g. if  $c$  was not constant but had a higher value for selective over generalist imitators) then the probability of alleles for selective imitators invading and their stability to invasion would be enhanced.

We have developed this example to illustrate how a hypothesis can be neatly formalized and tested under the phenogenotype paradigm. In our judgement the analysis suggests that there are realistic circumstances under which Blackmore's hypothesis may be operational. However, given that advantages of imitating appear to be dependent on the pattern of spatial and temporal variation in the environment, a more sophisticated analysis incorporating a changing environment might be an appropriate next step. Having illustrated a method for addressing the general feasibility of Blackmore's hypothesis, it is now open for further analysis.

## 4 Conclusions

The cultural evolution and gene-culture co-evolutionary modelling paradigms can be effectively employed to enhance the quantitative study of memetics. Simple and complex cultural phenomena such as behaviour patterns, belief systems and institutions can be analysed by characteristics of associations between easily definable and quantifiable memes. The quantitative approach can be used to describe meme diffusion dynamics, and make sense of patterns of variation in memes. The methods can also ask why and how human attributes evolved in conjunction with memes, how they continue to evolve, and what is the basis of any stability or maintenance of the trait.

## Notes

1. If your browser has written 'g' instead of the Greek letter gamma, then you should note that  $\gamma_2 = \gamma_2$ ;  $\gamma_1 = \gamma_1$ ; and  $b = \beta$ .
2. We consider the evolution of selective imitators, by assessing the local stability of the initial population consisting solely of  $B_{0i}$  individuals. Inequality (2) shows the conditions under which the  $B_{0i}$  population would be unstable, and susceptible to invasion from the  $I$  allele, and is

calculated by finding the Eigenvalues of the Jacobian matrix at  $\hat{F}_{0i} = 1$ . Invasion is restricted to parameter values that give the leading Eigenvalue a magnitude greater than unity, indicating that  $B_{0i}$  is not asymptotically stable with a small perturbation from equilibrium.

## References

- [1] Aiello, L. & P. Wheeler (1995). The Expensive-Tissue Hypothesis. *Current Anthropology*. 36(2):199-221.
- [2] Blackmore, S. (1999). *The Meme Machine*, Oxford University Press.

- [3] Boyd, R. & P. Richerson (1985). *Culture and the evolutionary process*, University of Chicago, Chicago.
- [4] Boyd, R. & P. J. Richerson (1988). An evolutionary model of social learning: the effects of spatial and temporal variation. In *Social learning: psychological and biological perspectives*, ed. Zentall. T.R. & Galef Jr, G. B. pp. 29-48, Hillsdale NY: Erlbaum.
- [5] Byrne R.W. & Whitten, A. (1988). Tactical deception in primates. *Behavioural and brain sciences*. **11**(2):267-271.
- [6] Cavalli-Sforza, L. L. F., M.W (1981). *Cultural transmission and evolution: a quantitative approach.*, Princeton: Princeton University Press.
- [7] Dennett, D. C. (1997). The evolution of evaluators. International School of Economic Research, Siena.
- [8] Dunbar, R. I. M. (1993). Coevolution of neocortical size, group size and language in humans. *Behavioural and brain sciences*. **16**:681-735.
- [9] Feldman, M., K. Aoki & J. Kumm (1996). Individual versus social learning: Evolutionary analysis in a fluctuating environment. *Anthropological Science*. **104**(3):209-232.
- [10] Feldman, M. & L. Cavalli-Sforza (1976). Cultural and biological evolutionary processes, selection for a trait under complex transmission. *Theoretical Population Biology*. **9**:238-259.
- [11] Feldman, M. & L. Cavalli-Sforza (1979). Aspects of variance and covariance analysis with cultural inheritance. *Population Biology*. **15**:276-307.
- [12] Feldman, M. W. & K. N. Laland (1996). Gene-culture coevolutionary theory. *TREE*. **11**:453-457.
- [13] Holloway, R. L. (1981). Cultural systems and brain evolution. *Dialectical Anthropology*. **5**:287-303.
- [14] Humphrey, N. (1976). The social function of intellect. In *Growing points in ethology*, ed. Bateson. P.P.G., Hinde, R., pp. 303-317, Cambridge University Press.
- [15] Laland, K. (1993). The mathematical modelling of human culture and its implications for psychology and the human sciences. *British journal of psychology*. **84**:145-169.
- [16] Laland, K. N., P. J. Richerson & R. Boyd (1996). Developing a theory of animal social learning. In *Social learning in animals: the roots of culture*, ed. Heyes, C.M. & Galef Jr., B.G. pp. 129-154, Academic Press, London.
- [17] Lumsden C.J., Wilson., E.O. (1981). *Genes, mind and culture: the coevolutionary process*, Harvard University Press.

- [18] Plotkin, H. (1994). *Darwin Machines and the Nature of Knowledge*, Penguin, London.
- [19] Reader, S. M. & K.N. Laland (1999). Do animals have memes? *Journal of Memetics - Evolutionary Models of Information Transmission*, 3(2), [http://jom-emit.cfpm.org/1999/vol3/reader\\_sm&laland\\_kn.html](http://jom-emit.cfpm.org/1999/vol3/reader_sm&laland_kn.html).
- [20] Wilson, A. C. (1985). The molecular basis of evolution. *Scientific American*. 253(4):148-157.

# **DIFFERENT TYPES OF MEMES: RECIPEMES, SELECTEMES, AND EXPLANEMES**

**By**

**John Z. Langrish**

Manchester Metropolitan University

JLangrish@aol.com

## **Abstract**

This paper attempts to provide three directions for advancing Dawkins' original description of memes as 'units' of cultural transmission'. These three directions are

1. There could be different types of memes.
2. Memes are not 'units'.
3. The transmission of memes might not be restricted to something called 'culture'.

A biological perspective allows for different types of memes with different transmission mechanisms. The disease type mechanism (epidemiology) has been over emphasised and is only one such mechanism. Three types of memes are suggested and given the names recipemes, selectemes and explanemes. Their use is illustrated by the evolution of technology in which black box systems are used to do things requiring recipemes - ideas about how to do it - and selectemes - ideas about what is a desirable output. Explanemes are ideas about the inside of the 'box'.

**Key words:** memes, evolution, culture, technology, biology.

## **1 A Biological View**

The three lines of advance, mentioned above, stem from a 'biological' view of the world as opposed to a 'physics' view. This 'biological' view (abbreviated to **B**) is based on Darwinian evolution but has roots going back to Aristotle's tendency to develop typologies i.e. accounts of different types of things. The 'physics' view (abbreviated to **P**) is based on Newtonian mechanics but has roots going back to Plato's concentration on the essence of things rather than their differences. The big difference between the **B** and **P** views is that **B** welcomes diversity and **P** hopes for one theory of everything. ('Real' Physics and Biology are not quite the same as **P** and **B**. Fundamental particles begin to look like botany and reductionist biologists seem to think like physicists)

Scientific thinking is so imbued with the **P** view that it sometimes seems that the existence of an alternative scientific way of looking at things has been kept a carefully guarded secret. One of the few people to attempt an account of **B** is Ernst Mayr whose 'The Growth of Biological Thought' (Mayr 1982) is recommended to economists, sociologists and anyone interested in evolution.

Certain features of the **B** view need highlighting to provide a background for the **B** view of memes. These are:-

i) There is no such thing as a single cause for any event or process in biology. For example Steven Rose (1997 p 10) gives five different explanations for a frog jumping away from a snake. These range from its evolutionary history to the biochemical properties of its muscles. The processes by which ideas are transmitted are surely more complex than the reaction of a frog to a snake.

ii) **B** concepts are fuzzy patterns and advances are made through subdivision. (This is the opposite of **P** which has mathematically precise concepts and likes to reduce the number of concepts, not increase them).

As Mayr (1982) puts it-

*"Particularly important (in scientific progress) is the occasional recognition that a more or less technical term, previously believed to characterise or designate a certain concept, was in reality used for a mixture of two or more concepts, like 'isolation' for geographical and reproductive isolation or 'variety' for individuals and populations."* p43

It seems reasonable to suggest that advances in memetics could also be made by regarding concepts as mixtures or patterns rather than by attempting precise definitions.

iii) The need for descriptive studies. Biology used to be sneered at as one of the descriptive sciences. Its higher status today has partly come from adopting a **P** approach, sometimes referred to as reductionist molecular biology.

Yet Mayr (1982) claims,

*"In evolutionary biology almost all phenomena and processes are explained through inferences based on comparative studies. These in turn are made possible by very careful and detailed descriptive studies. It is sometimes overlooked how essential a component in the methodology of evolutionary biology the underlying descriptive work is."* p.70

What is needed to advance memetics is detailed studies of memes competing and surviving. The history of ideas is a start. This can become the evolution of memes but more is needed; the day to day natural history of memes needs to be added to existing notions of influential ideas, their propagation and recombination (e.g. Lovejoy 1936). Darwin's theories came after years of detailed study, not before.

iv) The nature of evolution. From the **P** view, evolution is 'unfolding' but biological evolution is NOT a gradual unfolding of a predetermined pattern. The growth of an embryo or the growth of an oak tree from an acorn are examples of 'unfolding' or development.

Confusing unfolding which is predictable with Darwinian change which is not, is a mistake made mainly by physicists who tend to believe that everything in principle is predictable. Thus a report

in the *Times Higher Educational Supplement* (12 April 1996 p3) describes the computer modelling of the growth of the Universe. The article refers to 'the latest step in unravelling the evolution of the universe'. It quotes Carlos Frenk, Physics Professor at Durham University, as saying, "*Imagine you could do a computer simulation that would take the fossil record and evolve it until we see the human growing*". The article also refers to the 100,000 years old universe as, "*the embryo universe*".

The development of an embryo is actually quite a different process from the changes that led from the life represented by the fossil record to our present existence. It is only with the advantage of hindsight that we might reconstruct the evolution of humanity. If some scientific being came from outer space and only had fossils up to the time of say the reptiles, there is no way that humanity could be predicted from these fossils. At any stage in biological evolution, the number of possible future states is enormous. Those varieties that do survive and lead to further varieties do so partly by luck and coincidence and because of environmental changes which are also not predictable.

This is why some physicists claim that the Darwinian theory of evolution is not a scientific theory at all (the only science is physics; the rest is stamp collecting or social work or cookery or qualitative description, depending on which quote you prefer). Nonetheless, the **B** view is a scientific view. It can make predictions that can be tested - predictions about the fossil record for example. Any good classification system (stamp collecting) provides predictions. If a new chemical is identified as being secondary, it is possible to predict some of its properties and ways of making it from other chemicals. Mendeleev's periodic table, a two dimensional classification system or pattern, predicted the existence of two unknown elements which were soon discovered. This prediction was possible without any knowledge of electrons, the nature of the chemical bond, wave mechanics etc. which are the product of **P** type approaches. What Darwinian evolution cannot do is predict the future results of change; it is not an unfolding development.

v) The importance of variety (of whatever kind). This can be summarised as: Without variety there is no competition. Without competition there is no evolution. Without evolution there is no biology. In contrast, the **P** view likes things to be the same. Where variety does exist, this is reduced to 'deviation' from the 'norm'. The emotive connotations of these two words demonstrate the preferences of **P** type thinkers. Mayr's version of what I label **P** and **B** is his essentialist vs. populationist divide. He claims (Mayr 1982),

*"The statistics of the essentialist are quite different from those of the populationist...Differences in height among a group of people are real and not the result of inaccuracies in measurement. Darwin could not have arrived at a theory of natural selection if he had not adopted populational thinking. The sweeping statements in the racist literature, on the other hand, are almost invariably based on essentialistic thinking."* p.47

In other words, variety is to be embraced and not regarded as something that gets in the way of prediction. Different kinds of memes having different methods of propagating and altering have to be the starting pattern for advances in memetics. (**B** advances from a pattern; only **P** has a

starting point). Armed with a **B** view of the world, we can take a fresh look at Dawkins original statement on memes.

## 2 Three Lines of Advance

As readers of the Journal of Memetics will know, Dawkins (1976 - Chap 11) described memes as units of cultural transmission. The word is a contraction of mimeme. Examples of memes given by Dawkins are "*tunes, ideas, catch phrases, clothes fashions, ways of making pots or of building arches*". Richard Dawkins is described by Niles Eldredge (1995) as the ultimate ultra-Darwinian and Eldredge claims, "*Ultra-Darwinians are guilty of a form of 'physics envy'*". The failure of Dawkins and others to develop the meme of memes may be due to **P** type thinking, especially concentrating on 'units'. In calling memes 'units of cultural transmission', Dawkins could have been mistaken about three things, representing three lines of advance; these are:-

### 2.1 Memes are not all the same

There are different types of memes. Memes are biological entities - not physics type particles. Progress can be made by identifying different kinds of memes with different ways of competing and different ways of being transmitted and changed.

### 2.2 Memes are not units; they are patterns - or Russian dolls

**P** type units are essentially the same as each other. An atom of sodium may be a different isotope or in a different energy state but an atom of sodium is the same now as it was a million years ago or as it would be a million miles away. (Once you have seen one you have seen them all). **B** type things are different - they vary and change over time and place. An elephant in India is different from an elephant in Africa and in fact every elephant is unique. The 'idea' of a meme is not a unit; it is a complex concept meaning different things to different people. Some see memes as being infectious parasites invading the mind i.e. different from 'ordinary' ideas. Others do not draw this distinction.

The 'idea' of a railway is not a 'unit'. When the Liverpool - Manchester railway was built, many people could 'catch' the idea of a railway from looking at it working. This did not mean that they could then go and build railways somewhere else. The amount of information needed to construct and run a railway system is enormous. The meme of a railway is a Russian Doll, which includes knowledge of suspension systems, traction systems, signalling, time tables, track maintenance, safety, finance, customer preferences etc. No single human has the complete 'idea' of a railway in a brain or a computer. From the point of view of an observer rather than an operator, the idea of a railway might be thought of as a pattern consisting of fragments of memory assembled around a label - railway. From the point of view of a user, a railway is an idea in competition with ideas about other forms of transport. Whatever is meant by the idea of a railway, it is different for different people and it is not a unit.

Some writers (e.g. Ball 1984 and Durham 1991) have seen memes as being more than unit bits of information or unit ideas. They can be complexes of things, behaviour and ideas but still transmitted as units and firmly placed in something called culture. Now it is possible to define

culture in such a way as to include almost everything 'out there' but such definitions usually include something about cultural transmission. This brings us to the next point.

### **2.3 Memes are not restricted to cultural transmission; other forms of transmission are possible**

Some writers with an interest in cultural evolution anticipated Dawkins 'memes'. There is a tradition which sees culture as something having an evolutionary life of its own. (Spencer's social evolution which preceded Darwin by three years, described culture as super organic, coming on top of the organic - living - and the inorganic - inanimate). Thinking along these lines produced antecedents to Dawkins' memes. If there is an independently evolving culture, it might have something analogous to genes or it might actually have its own form of genes. Suggested names for these genes of culture include mnemotype (Blum 1963), culturetype (Burhoe 1967), sociogene (Swanson 1973) and culturgen (Lumsden and Wilson 1981). These writers share with Dawkins a tendency to view the world from a **P** perspective which causes problems because evolution needs a **B** view.

A **P** view of things is pleased by identical fundamental particles - hence units. Such a view also likes fields in which to place its basic units. Culture to Dawkins and others seems to provide the field in which the meme particles move about. Culture, however is not **P**; it is **B**. It is not like a gravitational field in which projectiles have predictable trajectories (until humans intervene). It is a biological field in which things grow and compete. A problem with culture is knowing which way round to look at the relationship between culture and humanity. Is culture something created by humans or are humans at the mercy of whatever culture they happen to be born into? Is it an agricultural field, artificially selected or is it a wild field, naturally selected? (Or from a chemical perspective, is there a 'bond' between humans and culture? - Bonds go both ways).

From the many definitions of culture, the following might help to show why I think this term does not help memetics -

*"culture is viewed as man's way of maintaining life and perpetuating his species, a system of learned and socially transmitted ideas, sentiments, social arrangements and objects that depend for their formulation and continuation upon man's ability to create symbols." (Gamst and Norbeck 1976).*

Quite apart from the outdated emphasis on 'man' having all the fun, I am in disagreement with two aspects of this definition, the inclusion of objects - not ideas about objects but the actual objects - with everything else and the necessity of having symbols for transmission. Now certain kinds of ideas do require symbols. The ideas of science, for example, and explanations in general need words or maths for their transmission. (It might be that having explanations is the one real difference between the human and other animals). This kind of transmission is not always necessary for other types of ideas such as ideas about how to do things. Ideas about playing the piano or making even apparently simple things, like pots, are very difficult to put into words or other symbols. They involve 'finger tip' knowledge and ideas about things feeling right. An idea of how to do something can be transferred more easily by doing it than by describing it. Written music, of course, involves symbols but written notes do not tell you how to play the piano. Tunes

can be transmitted by written music but more often their transmission has no dependence on symbols. Cymbals can be better than symbols for the transmission of musical ideas.

Another description of culture is provided by Lumsden and Wilson (1983), who state,

*"Human cultures consist of artifacts, such as knives of a certain shape and function; behaviours, such as initiation rites of a particular form; and mental constructions having little or no direct correspondence to reality such as myths."*

To lump together knives, initiation rites and myths under the label 'culture' may be all right for those who want to see the world that way. I don't. I have a suspicion that the knives are more important and that transmission of ideas about knives is different from the transmission of mythical ideas. Though I am sure that myths need words or other symbols for their transmission, I am also sure that knives don't. A knife is its own picture- a representation not a symbol. Using a knife can be copied directly.

To make any progress, we have to unpack this 'culture includes everything' approach to look at different kinds of memes with the possibility of different means of transmission. To those who cannot imagine a meme without cultural transmission, I would say, what is the meme of a railway? Does it require a railway culture in order to move from one country to another and if not, why not?

Before attempting to describe different types of memes, it is necessary to discuss another aspect of transmission that has held up memetics, namely a fascination with memes as parasites or viruses.

### **3 Epidemiology - a side track**

Much of what has been written about memes can be summed up by the title of Aaron Lynch's book (1996), 'Thought Contagion'. This approach has been used to attack both Communism and religion, provoking a counter attack by John Bowker whose 'Is God a Virus?' (1995) attempts to bring free will into the discussion.

If some extra-galactic intelligence discovered our earth, they might be expected to be interested in carbon based life. However, if after twenty years they had only studied viruses and parasites we might think they were deficient in curiosity or just weird. Why no interest in ants, elephants, spiders, trees, eagles, coral, sharks or fungi? It is a bit like that with memes; why such little interest in the memes of technology, poetry, design, economics and all those interesting human activities involving ideas. After all, Dawkins' original description of a meme included 'ways of making pots and building arches'.

There are two further problems with the memes as viruses school of thought. One is that it ignores Dawkins original use for memes - as the basis for a new kind of evolution, acting on top of genetic evolution. Epidemiology is not in itself evolutionary unless it asks historical questions about the viruses. The second problem is that it has not found a use for memes as such. Ideas about the spread of 'foreign' ideas have been around a long time. Have they been improved by

the addition of memes? Older readers will remember Vance Packard's *Hidden Persuaders* (1957) and might have come across William Sargent's 'Battle for the Mind' (1957) which has the opening remark,

*"...this book is not concerned with the truth or falsity of any particular religious or political belief. Its purpose is to examine some of the mechanisms involved in the fixing or destroying of such beliefs in the human brain."* p.9

Memetics has added some new words to describe the transmission and effects of memes, e.g. Glenn Grant (1990) has bait, dormant, hook, membot, meme allergy, sociotype, vaccine and so on, which are fun but don't really add very much to a discussion of ideas outside religion and politics. This emphasis on religion and politics and the disease-like nature of the spread of 'foreign' ideas would not have mattered too much if this tendency had been confined to those Xenophobic Americans wanting to believe that Communism was a dangerous infection capable of sapping the minds of true Americans (e.g. Peter Vajk, 1989) claiming that perestroika was a deliberate strategic deception by the infectious meme of Communism). Unfortunately, the meme of the infectious meme seems itself to be rather infectious. The author of the immortal quote, "*Memes share the farmyard with us and it is up to us which of them we let push us around*" (Westoby 1994) was not a hawkish American - he was an English Lecturer in Education, with so called left wing views.

Dawkins himself seems to have added little to memes beyond his diatribe against religion (Dawkins 1993) apparently fuelled by his six year old daughter being given instruction by a nun without his knowledge. The big problem with memes as viruses, parasites or even symbiotes is how do you know which memes are 'you' and which are 'invaders'. Dawkins (1993) glimpses this problem with an afterthought about science. Couldn't the ideas of science be just the same as the ideas of religion; that is 'caught' from teachers, books etc.

'No', says Dawkins (1993):-

*"The rapid spread of a good idea through the scientific community may look like a description of a measles epidemic. But when you examine the underlying reasons you find that they are good ones, satisfying the demanding standards of scientific method. For scientific belief, epidemiology merely comes along afterwards and describes the history of its acceptance. For religious belief, epidemiology is the root cause."* p.23

To me this really does not help. How do you know that a reason is a good one? You have to have ideas about what is a good idea - selectemes. Such ideas can be 'caught' as easily as any other kind of idea. As Dennett puts it (1995):-

*"it cannot be 'memes versus us' because earlier infestations of memes have already played a major role in determining who or what we are. The independent mind struggling to protect itself from alien and dangerous memes is a myth."* p.365

Dennett removes the polarity between us and the invaders by seeing our notion of self as the result of 'infestation'. The idea that the mind begins its life as a clean slate or an empty vessel

waiting to be filled goes back a long time (e.g. John Locke (1690) wrote about the mind being at first a 'yet empty cabinet' which was furnished by experience.) This notion leads to thinking about the source of the ideas that are supposed to fill the empty mind and some people have seen a quasi independent culture as the filling agent. It can be seen that the two uses for memes - memes as nasty invaders of an independent self and memes as the units of culture without which there would not be a 'self' - are as different as sociology and psychology or as macro and micro economics. They ought to meet but they don't. My aim is to find something different from either of these limited perspectives.

I want to know the difference between a belief that is somehow based on evidence and one that is not. I want to know how ideas change with time - how they evolve. To answer these kinds of questions, I need a biological evolutionary approach using different kinds of memes involving different transmission methods.

## 4 Different kinds of memes

### 4.1 Memes for evolution

My own interest in memes lies in the possibility of using them in an evolutionary framework for discussing technological innovation. An evolutionary viewpoint requires:

- Variety
- Competition between the varieties
- Replication of the 'winners'
- A mechanism for the production of further variety (followed by further competition)

A mechanism for changing the 'rules' of the competition. (Without this requirement, evolution stops when diminishing returns from further variety lead to equilibrium)

From a modern viewpoint, the evolution of technology is the evolution of ideas about artifacts within environments which select certain ideas and reject others. Competing ideas are first subject to selection within a mind- they compete for attention. They then compete for approval by people who control resources. A very small proportion of ideas end up incorporated in something else, an actual artifact, a process, a system etc. These things then seem to compete with each other in an exosomatic (outside the body) evolutionary process.

Jean and Peter Medawar's book(1978) has a chapter on exosomatic evolution which starts,

*"Everybody has observed that the human artifacts which serve as tools are to some extent extensions of the body".*

They are probably rather optimistic about the observational power of 'everybody' but that is just a way of saying that they cannot remember how they got such an idea. Way back in 1872 Samuel Butler expressed this idea in 'Erewhon' (almost 'nowhere' spelt backwards but pronounced with three syllables) pointing to a spade as an external limb, a note book as an external memory, a magnifying glass as an external eye etc. The Medawars then add on to Butler's concept of

evolutionary external organs the point that it is not the tools (organs) that evolve but ideas about tools:

*"It is very clear that these exosomatic parts of ourselves undergo a slow systematic secular change of a kind which it is perfectly possible to describe as involution' - exosomatic evolution - provided of course one realises that it is the design of these instruments that undergoes the evolutionary change and not the instruments themselves, except in a quite unnecessarily figurative sense".*

Instead of drawing a parallel between genes and designs (both are 'instructions' that affect the production of something), the Medawars add on ideas from cultural evolution. They say:

*"Ordinary organic evolution is mediated through a genetic mechanism but exosomatic evolution is made possible by the transfer of information from one generation to the next through non-genetic channels. By far the most important of these non-genetic agencies is language. It is because of the primacy of language... that exosomatic evolution is often referred to as cultural or psychosocial evolution". chap.6*

So culture and language are still in the picture. Ideas about technology, however, can be transmitted without words and without invoking some cultural 'field'.

Competition between ideas leading to some ideas being replicated, modified or added to can be discussed without reference to 'culture' (which is not the same as saying ideas are culture free - just that some ideas, like some microbes can be transmitted between cultures even though both ideas and microbes can flourish within an appropriate 'culture')

Technology moves across cultures quite easily. If your tribe is attacked by a new tribe armed with bows and arrows which kill humans and animals then your tribe absorbs the idea of bows and arrows. You don't have to learn the language of the new tribe to do this nor do you need the special prayers and ceremonies which the new tribe associate with making the bows and arrows; you have to watch them being used and perhaps capture some. Bows and arrows can be imitated. This is not the same as saying all artifacts can be easily imitated; nor is it the same as saying artifacts are culture-free. It is just saying that the meme of using bows and arrows is better described as a technological pattern, a Russian doll idea of how to do something rather than a unit of cultural transmission. Since I think about technology as a way of doing something better than it could be done by an unaided human, a convenient class of memes is ideas about ways of doing something. I propose to call these recipemes.

## **4.2 Recipemes, selectemes and black boxes**

Recipemes are competing ideas of how to do things. Successful recipemes are replicated - sometimes with modification or addition. Recipemes have a different method of transmission from the other memes. Knowing how to do something often involves 'finger tip' knowledge which can only be obtained through doing. You cannot tell someone how to ride a bicycle, when the clay for a pot feels right, how to play the piano or how to knap a flint. You can write down the recipe for a cake but this will assume some shared practical knowledge.

Once we have different ways of doing different kinds of things, we have ideas of success and betterness. Some things and some ways are 'better' than alternatives. I propose to use the word, 'selecteme' to mean ideas that form the basis of selection. Selectemes are competing ideas of betterness. They provide the mental environment in which other memes compete for selection. (Selectemes, of course, compete with other selectemes!) The transmission of selectemes is closely bound to something that might be called a society - in an early version, I tried calling them societemes but decided that was too restrictive.

Many people, of course, don't accept all the selectemes that are offered them. Some have Friday night selectemes that are different from those that are present on Monday morning. So how do we select our selectemes? P-type thinkers will immediately see an infinite regression of minor selectemes being selected by higher order selectemes which are... . However, in a B world, things form non mathematical patterns. A selecteme is a Russian doll type pattern which forms a whole. When we feel that something is wrong, we do not think, "I will not do that because it would be stealing and stealing is against my religion which I have chosen to abide by even though I do not believe in God and in any case I might get found out and that would bring shame which I do not like..." No, we either just feel it would be wrong or we feel it is worth the risk. Either way that feeling can be described as a pattern of selection which the pattern of proposed action either fits, does not fit, or is repelled by. When a pattern of action fits a pattern of selection, we have a 'click' which Maria Abu-Risha (1999) calls *Purposive Pattern Recognition*. An important point is that patterns are not units nor are they always made of the same units. The pattern of the letter A can be constructed out of many things including a hole.

Recipemes and selectemes can be combined in black box systems. A black box can be a rubber tree, a chemical reaction, a loom, a bow and arrow or anything that has inputs and outputs under some degree of control. Recipemes are ideas about inputs into boxes (raw materials, energy and conditions) and about alternative boxes. Selectemes are ideas about outputs and their relative desirability. Since boxes can be connected - the output from one being the input to another - there can be long chains of recipemes and selectemes. The loom maker makes a loom using selected materials made elsewhere; the weaver selects a loom to make cloth from selected yarns made elsewhere; the tailor turns the cloth into clothes and the clothes are purchased and worn. At each stage there are recipemes - ideas about how to do things - and selectemes - ideas about what sort of loom, yarn, cloth and clothes might be desirable or undesirable.

In addition to forming long chains of inputs and outputs, recipemes and selectemes can form Russian doll nesting structures. The human body is a special black box. Part of its input is food. If someone decides to bake a cake, selectemes for healthy life style and sensuous enjoyment compete, leading perhaps to a creamy, fatty chocolate cake or a fat-free cake made from organic whole meal flour. Inside these selectemes and recipemes will be other more specific ones. Which recipe for a healthy cake or a delicious cake might be used? Someone might have a selecteme for Delia Smith; someone else for mother's trusted recipe. The person making the choice does not work through a decision tree. They do not think, "I am in favour of being healthy so I must have a healthy diet and that means using a recipe from 'Recipes for a Healthy Life' in which is a rather nice cake which I have made before and my friends enjoyed". No, they just think, "I'll make that cake again". The actual cake when made and eaten by other people might influence future

choices made by others. Recipes do get copied and ideas of what is a 'good' cake can be changed. Such changes usually involve real cakes rather than symbols of cakes.

So we now have a view of the world consisting of interacting black boxes. Humans attempt to alter this world through thinking and doing. Black boxes can be observed and their inputs can be altered. Different inputs give different outputs some of which are judged to be desirable or more desirable than existing outputs. Different black boxes can be tried out. Ideas about inputs and boxes are recipemes. Ideas about desirability of outputs are selectemes. These ideas compete firstly in one mind and then in the world and then in the minds of other people.

### 4.3 Explanemes and Institutions

A third type of meme is required to describe ideas about what is happening inside the black box. I propose to call such ideas explanemes. Explanemes are competing ideas that are used in answering questions about why things work or work better. Selectemes associated with both curiosity and communication enable the survival of ideas about the insides of the black boxes. These 'inside' ideas are explanemes.

Some explanemes lead to suggestions for new black boxes or improvements in existing boxes. Some are 'just - so' stories. Some provide words to enhance communication. Some are highly sophisticated and live in special institutions. Explanemes are competing ideas that are used in answering questions about why things work, work better or don't work. Recipemes just tell you how to make something. Explanemes provide a story about what is going on. The transmission of explanemes always requires a language - maybe maths - maybe symbols - but you can not tell a story about something without a story telling language. Stories do translate from one culture to another and so do special symbols such as algebra which moved from an Arabic origin to a European usage.

Explanemes may be expressed in a specialist vocabulary to assist communication about black boxes, their recipemes and appropriate selectemes and they may fit into an environment of curiosity. Competition between explanemes takes place within the mind of anyone who is wondering 'why?'. Competition also takes place within complex organisational structures of the kind that are usually called institutions. Scientific institutions provide an environment for the competition between scientific explanemes. This way of looking at science has been neatly described by John Ziman (1998) as,

*"a peculiar type of social institution, devoted to the production of public, communally acceptable knowledge about the natural and social worlds through a delicately balanced tension between originality and criticism".*

Ziman's description of science meets the requirements of an evolutionary system (though that is not his intention). Originality is a variety production system. Criticism is the arena of competition. The 'winners' are reproduced through becoming communally acceptable. His delicately balanced tension is also descriptive of biological evolution in that the rate of mutation of genes is not so high as would lead to the loss of successful characters nor so low as would lead to a lack of adaptability.

An explicitly evolutionary approach to science has been provided by David Hull (1988) who claims that the pressures that shape the rise and fall of species are similar to those acting on scientific ideas. He makes use of a selection system involving replicators and interactors. His conceptual replicators could be seen as memes but there is a potential clash here with those such as Susan Blackmore (1999) who see memes as units of imitation. Scientific ideas are passed on but not by imitation. The potential clash is avoided if we have different types of memes with different methods of transmission. Explanemes differ from other memes in that they are not transmitted by imitation. They are still memes, however and they form part of an evolutionary system which sometimes involves institutions.

It is not just the explanemes of science that compete within institutional frameworks. Other types of memes also have special institutions with 'a delicately balanced tension between originality and criticism'. Thus manufacturing industry provides an arena for competition between different technological recipemes. The institutions of regulation, law and government oversee competition between those selectemes that are written down and enforced.

## 5 Memes and Evolution

We now have to add the dimension of time. Given different types of memes and institutions producing a delicately balanced tension between originality and criticism, we have a mechanism for evolutionary change which could transform economics, politics, sociology etc. In particular, technological evolution becomes a much more interesting concept than has been the case. Many people have considered the evolution of technology as hardware. For example Basalla (1988) has a four component evolutionary system which is worth looking at. These four concepts are diversity, continuity, novelty and selection (applied to things rather than to ideas). An essential part of Basalla's case is the continuity of all technology right back to the use of the first stones and flints. Basalla's continuity is of the strong variety. He is not talking about use or concepts. He claims, "Any new thing that appears in the made world is based on some object already in existence". In support of this he tackles some tricky cases. The electric motor, for example, is based on two existing devices- the magnetic compass and the steam engine.

Basalla also discusses the transistor and shows that the first commercial transistors were regarded as improvements on the old style crystals which had preceded the thermionic valve. Literature from Bell in 1948 announced, "*In the Transistor, two point contacts of the 'cat's whisker' or detector type, familiar to radio amateurs, are made to the semiconductor*". The point contact transistor was soon replaced by the junction type but it provided an evolutionary link. In Basalla's four component system, the concept of continuity is the weak link.

Just what does it mean to say that something is based on something else? This is why we need memes. If we take the Medawars' point that it is ideas that evolve then we are looking for a system which allows ideas to compete, propagate and change.

The first question has to be, why do we need memes to have an evolutionary theory of technology? Why not just stick with ideas and their history? You can have a history of ideas without using memes. It has been claimed that Arthur Lovejoy invented the history of ideas as an academic discipline (he certainly founded a Journal) but the introduction to his 'Great Chain of

'Being' (1936) clearly states that he was looking for atomistic 'units'. Since advances may lie in the direction of scrapping the notion of idea units, Lovejoy's approach will not help.

Ideas change over time; they evolve. Selectemes (ideas about what is best) evolve and other memes have to adapt to their changing environment. Thus the planners of the Concorde shared the selecteme that faster meant better. So the world's first supersonic civil air travel became possible. However, the selecteme for fast lost out to other selectemes to do with both economy and noise. Different selectemes compete for the mental space marked 'the idea of the best' and different recipemes compete to supply ways of achieving the desired result. Some ideas are more competitive in association with a reason; so we find that explanemes are required to assist.

To summarise, memes are needed for evolutionary explanations of human activities. They compete, replicate and vary. They are involved in much more fascinating activities than those suggested by epidemiology. The concept of memes becomes more sophisticated and powerful when broken down into different types of memes with different ways of competing and being replicated. Lack of space prevents discussion of where new ideas come from but it should be obvious that the different memes whilst sharing the common features of an evolutionary system have different variety production mechanisms. New explanemes sometimes result in a Nobel Prize. New recipemes can be patented in the name of their inventor. New selectemes are rarely associated with a named individual (there are some notable exceptions such as Schumacher's Small is Beautiful) but many new ideas seem to crop up simultaneously. There is a need for studies of meme mutation; we don't need any more epidemiology. Evolution and memes need a 'B' view of the world. The spreading of identical units through a cultural 'field' is far too 'P' to be much help in the complex adaptive world of humans, their artifacts and their changing ideas.

## References

- Abu-Risha, M. (1999). *Purposive Pattern Recognition*. PhD, DeMontfort University.
- Ball, J. A. (1984). Memes as Replicators, *Ethology and Sociobiology*, 5:145.
- Basalla, G. (1988). *The Evolution of Technology*. Cambridge: Cambridge University Press.
- Blackmore, S. (1999). *The Meme Machine*. Oxford: Oxford University Press.
- Blum, H. F. (1955). *Times Arrow and Evolution*. Princeton: Princeton University Press.
- Bowker, J. (1995). *Is God A Virus? Genes, Culture and Religion*. London: SPCK.
- Burhoe, R W. (1967). Five steps in the Evolution of Mans Knowledge of Good and Evil. *Zygon*, 2:77-96.
- Butler, S. , (1872). *Erewhon: or Over the Range*. London: Jonathan Cape.
- Dawkins, R. 1976. *The Selfish Gene*. Oxford: Oxford University Press.

- Dawkins, R. 1993. Viruses of the Mind, in Dahlbom B. (ed.) *Dennett and his Critics: Demystifying mind*. Oxford: Blackwell, 13-27.
- Dennett, D. C. 1995. *Darwins Dangerous Idea: Evolution and the Meanings of Life*. London: Allen Lane/Penguin.
- Durham, W. H. 1991. *Coevolution: Genes, Culture and Human Diversity*, Stanford University Press.
- Eldredge, N. 1995. *Reinventing Darwin: The Great Evolutionary Debate*. New York: John Wiley.
- Gamst, C. G. and Norbeck, E. 1976. *Ideas of Culture: Sources and Uses* p6. New York: Holt, Rinehart and Winston.
- Grant, G. 1990. *Memetic Lexicon*. Share-Right, PO Box 36, Montreal H3C.
- Hull, D. L. 1988. *Science as a Process: An Evolutionary Account of the Social and Conceptual Development of Science*. Chicago: University of Chicago Press.
- Locke, J. 1690. *An essay concerning human understanding Vol 1*. Oxford.
- Lovejoy, A. O. 1936 (Harper and Row edition 1960 New York. ) *The Great Chain of Being: A study of the history of an idea*. Harvard UP.
- Lumsden, C. J. and Wilson E. O., 1981. *Genes, Mind and Culture*. Harvard University Press.
- Lumsden, C. J. and Wilson, E. O., 1983, *Promethean Fire: Reflections on the Origin of Mind*. Harvard University Press.
- Lynch, A. 1996. *Thought Contagion: How Belief Spreads through a Society*. New York: Basic Books.
- Mayr, E. 1982. *The Growth of Biological Thought: Diversity, Evolution and Inheritance*. Cambridge: Harvard University Press.
- Medawar P. B. and J. S, 1977. *The Life Science*, Wildwood House, London, Chapter 6. (See also Medawar P. B.,1973, *Technology and Evolution*, The Frank Nelson Doubleday Lectures, New York. )
- Packard, V. 1957. *The Hidden Persuaders*. David McKay. English ed. 1960.
- Rose, S. 1997. *Lifelines: Biology, Freedom, Determinism*. London: Penguin.
- Sargant, W. 1957. *Battle for the Mind*. London: Heinemann. Pan edition 1959.

Steadman, P. 1979. *The Evolution of Designs: Biological analogy in architecture and the applied arts*. Cambridge UP.

Swanson, C. P. 1973. *The Natural History of Man*. Prentice Hall.

Vajk, J P. 1989. Memetics: The Nascent Science of Ideas and their Transmission.  
[<http://www.notam.uio.no/~mariusw/memetics/memetics.txt>](http://www.notam.uio.no/~mariusw/memetics/memetics.txt)

Westoby, A. 1994. The Ecology of Intentions: How to make Memes and Influence People.  
Unpublished MS, available from Centre for Cognitive Studies, Tufts University and on www.

Ziman, J. 1998. A Limited Excaliber, Times Higher Education Supplement, 24th April p25. See also Public Knowledge.

## A NOTE ON THE ORIGIN OF “MEMES/MNEMES”

By

John Laurent

School of Science

Griffith University

In the *Selfish Gene* Richard Dawkins says that he had originated the term 'meme', a cultural equivalent of 'gene', by shortening 'mimeme' which he says he derived from the Greek mimeisthai, to imitate. This may be, but another, and more straightforward source for the term, I would suggest, is 'mneme', referring to a unit of memory and taken from the Greek mimneskesthai, to remember (and ultimately from Mnemosyne, the Greek goddess of memory [Note 1]), and which appears in the book *The Soul of the White Ant*, by the Belgian dramatist, essayist and amateur entomologist, Maurice Maeterlinck, first published in 1927. Nowhere does Dawkins refer to 'mneme', but it is difficult to believe that he had not come across the term, especially since he has much about termites in both *The Selfish Gene* and *The Blind Watchmaker* and cites a number of works about these creatures (though not Maeterlinck's) in both books. Also, in an endnote in the second (1989) edition of *The Selfish Gene*, where he elaborates on a sentence in the main text of the book which reads "*Memes should be regarded as living structures, not just metaphorically but technically*".

Dawkins writes as follows:

*"DNA is a self-replicating piece of hardware. Each piece has a particular structure, which is different from rival pieces of DNA. If memes in brains are analogous to genes they must be self-replicating brain structures, actual patterns of neuronal wiring-up that reconstitute themselves in one brain after another. I had always felt uneasy spelling this out aloud, because we know far less about brains than about genes, and are therefore necessarily vague about what such a brain structure might actually be. So I was relieved to receive recently a very interesting paper by Juan Delius of the University of Konstanz in Germany, [who] is bold enough to ram home the point by actually publishing a detailed picture of what the neuronal hardware of a meme might look like" (Dawkins, 1989, p.323).*

Now, the actual phrase that Maeterlinck uses - where he is discussing various theories which attempt to explain 'memory' in termites as well as the other 'social' insects (ants, bees etc.) - is "*engrammata upon the individual mneme*" (Maeterlinck, 1927, p.198), and according to my dictionary (Webster's Collegiate), an engram is "*a memory trace; specif.: a protoplasmic change in neural tissue hypothesized to account for persistence of memory.*" For what it is worth, Maeterlinck explains that he obtained his phrase from the "*German philosopher*" Richard Semon. The only book that I have been able to locate by this author (who describes himself as a naturalist, though he could have said 'natural philosopher' in the parlance of the time) is In the Australian Bush and on the Coast of the Coral Sea (Semon, 1899), but this book is nevertheless interesting in the present context in that Semon does indeed devote some pages to a discussion of animal intelligence and 'instinct' in it, together with possible correlations with brain structure (in Echidna, for example, whose brain "*in proportion to the size of the body...is more voluminous*

*than that of marsupials, and is further remarkable for its degree of convolution and the fissures on its surface".*

To repeat, it is hard to believe that Dawkins was not at least aware of Maeterlinck (who also wrote *The Life of the Bee*, which my copy shows went into 34 printings [in English] between 1901 and 1948), and this author's use of the term 'mneme'. In any event, Semon's earlier specific use of the term in an identical context to that in which Dawkin's correspondent Delius - another German writer - uses 'meme' is a remarkable coincidence.

## Notes

1. I am grateful to Peter Pegg for this information.

## References

- Dawkins, R. (1989). *The Selfish Gene*. Oxford: Oxford University Press.
- Maeterlinck, M. (1927). *The Life of the White Ant*. London: George Allen & Unwin.
- Semon, R. (1899). *In the Australian Bush and on the Coast of the Coral Sea: Being the Experiences and Observations of a Naturalist in Australia, New Guinea and the Moluccas*. London: MacMillan and Co.

# THE REDEFINITION OF MEMES: ASCRIBING MEANING TO AN EMPTY CLICHE

By

**Michael R. Lissack**

Institute for the Study of Coherence and Emergence

Naples, FL

[lissack@lissack.com](mailto:lissack@lissack.com)

## Abstract

By redefining memes as efficient tools for evoking particular affordances to be attended to in situ, this article argues that the challenges Edmonds (2002) issued to memetics can be met. Such a definition is consistent with theories of niche construction. The indexical Dawkins created (meme=gene) has exceeded its carrying capacity and thus lost its efficacy. Worse, that indexical is evoking images and affordances which stand in the way of the memetics field making true progress. It is time to recognize that ontic status has been misplaced. Memes needs a new meme: meme as catalytic indexical.

## 1. Introduction

Memetics has reached a crunch point. If, in the near future, it does not demonstrate that it can be more than merely a conceptual framework, it will be selected out. While it is true that many successful paradigms started out as such a framework and later moved on to become pivotal theories, it also true that many more have simply faded away. A framework for thinking about phenomena can be useful if it delivers new insights but, ultimately, if there are no usable results academics will look elsewhere. Such frameworks have considerable power over those that hold them for these people will see the world through these "theoretical spectacles" (Kuhn, 1969)—to the converted the framework appears necessary. The converted are ambitious to demonstrate the universality of their way of seeing things; more mundane but demonstrable examples seem to them as simply obvious. However such frameworks will not continue to persuade new academics if it does not provide them with any substantial explanatory or predictive "leverage." Memetics is no exception to this pattern. (Edmonds, 2002)

While the popular media has had several rounds of fascination with the concept of memes, the application of memes to management and to complexity has been negligible. The Edmonds quote above ascribes the problem to the stage of development of memetics: that memetics needs to provide explanatory leverage to get past the "crunch point."

This article suggests that the answer to memetics' "crunch point" lies in turning the concept of memes inside out. If memes are "units of cultural transmission which propagate themselves" (Dawkins, 1976) or "the least unit of sociocultural information relative to a selection process that has favorable or unfavorable selection bias that exceeds its endogenous tendency to change" (Wilkins, 1998), then the failure of the field of memetics to meet the three challenges outlined by Edmonds (a conclusive case study; a theory for when memetic models are appropriate; and a simulation of the emergence of a memetic process) is problematic and perhaps indicative of

"irrelevance." Indeed, there have been few managerial examples of the potency of a meme to explain or cause anything — and in the absence of explanatory or causal power; it is difficult to find the relevance of a concept for managers.

If, on the other hand, memes are redefined such that the evolutionary selection process is no longer an aspect of the ontology of memes but rather of the environmental niche (cf. Laland & Odling-Smee, 2000; Laland et al., 1999; Odling-Smee et al., 2003) of which the memes are evidence, then the field may have other avenues of advancement and a potential relevance to managers. Such a redefinition would entail recognition of the relationship between a given meme and the context of the social and ideational environment of which it is an affordance and which it demands be attended to. Memes in this casting are a label for successful boundary object indexicals and lose their privileged status as replicators. Instead, the replicator status is ascribed to the environmental niches and the memes are their representatives, symbols, or semantic indexicals. [Note 1]

With this definition, memes are repackaged as symbols and their impact on management is not that of a viral contagion but rather as an indicator of success and change in environmental niches. If an environmental niche has an important managerial role, then paying attention to its symbols and affordances can also be important. Memes are stripped of their causal role and instead become semantic tokens capable of evoking ascribed meanings. It is the process of evoking and the efficacy of the meme as the trigger for attention, recall, and repetition of the ascribed meaning that give memes relevance to managers.

The argument herein [note 2] thus assumes that memes exist, but that their definition is not that of replicator but rather that of indexical token. The meme tokens are representatives of the environmental niche in which they flourish and about which they offer efficient communicative potential. Memes, it is argued, succeed when they are accepted and used as tools for the accomplishment of a communicative purpose. Memes fail when their ontic status itself becomes a focus. Memes have longevity only if they both succeed and serve as a useful tool for a successful environmental niche. Memes can be short-lived due to the failure of their communicative efficacy or the failure of the niche they represent or both.

## 2. Memes as Indexicals

Indexicals are concepts that we make use of nearly every day but, for most of us, they are unknown and unthought about. The dictionary or encyclopedia entries are actually of little help. Take this entry from the *Stanford Encyclopedia of Philosophy*:

Indexicals are linguistic expressions whose reference shifts from utterance to utterance. "I," "here," "now," "he," "she," and "that" are classic examples of indexicals. Two people who utter a sentence containing an indexical may say different things, even if the sentence itself has a single linguistic meaning. For instance, the sentence "I am female" has a single linguistic meaning, but Fred and Wilma say different things when they utter it, as shown by the fact that Fred says something false, while Wilma says something true. (Braun 2001)

Indexicals are words used to stand for a set of other words; that is, they function like an index on the stock market. The Dow Jones Industrial average, for example, stands for a basket of particular stocks and stands for many of us as an indicator of the market as a whole. Pronouns as described above are indexicals in that they stand for the noun and take on different meanings in different situations. "Where a word acquires its sense from the context in which it appears; in different contexts, it changes its sense" (Vygotsky, 1986).

In American society the most commonly heard indexical is the mythical "they" who do things to us or others. "Look at what they are doing now." Perhaps the second most popular indexical is "You know..., you know who I mean." Indexicals are often distinguished by the fact that their reference systematically varies with the context of usage. Indexicals offer a simple means of making, expressing, and communicating our references, and they are particularly useful when proper names or descriptions are either cumbersome or unavailable. Similarly, in interpreting someone's "that way" in response to a request for direction, one must be able to determine independently what direction the person is indicating.

One must discover what relations one bears to the indexical referents in order to locate and act on them, but there is nothing indexical about these relations themselves (Millikan, 1993). To interpret an indexical, therefore, is to establish what other items, entities, and representations it coincides with. By itself, it tells us neither about its contents — what it bears its adapting relation to — nor about its contexts; context determines the indexical's content, but context is not what content is about. As David Kaplan (1989) puts it:

What is common to [indexicals] is that the referent is dependent on the context of use and that the meaning of the word provides a rule which determines the referent in terms of certain aspects of the context. [Note 3]

## 2.1 Efficiency

Indexical language possesses a quality that Barwise and Perry (1983) call efficiency. This refers to the capacity of an indexical representation to refer to different individuals on different occasions. The efficiency of an indexical relates to the engineering quality of efficiency in the notion that the same symbol is capable of standing for a multiplicity of meanings — the greater the number of potentially stood-for meanings, the greater the efficiency of the indexical. In organizations, management frequently refers to the company, a team, a symbol, or mission, vision, and values as absolutes without being aware of their indexical content. Thus, despite the likelihood that an organizational symbol carries a multiplicity of meanings, managers might make use of such symbols as if the only possible referents are the ones conceived of by the manager him- or herself. Yet, it is possible by ethnographic observation for a careful outsider to discern at least some of the various referents summoned up by the use of these "unintended" indexicals *in situ*. Indeed, by such processes of observation a corporate ethnographer or anthropologist can reach tentative conclusions about the corporate culture and climate of the organization being observed.

Indexicals are situated. The use of an indexical succeeds when the combination of context and symbol evokes an intended meaning. The indexical provides a locating space into which many

variants of personalized and situated meaning can be ascribed, attributed, or devolved. This space is the container of which Prigogine first spoke when describing self-organizing systems. In the absence of such containers, self-organization is nearly if not totally impossible. In the arena of culture, the meme as indexical is a locator or referent affording the evocation of situated meaning. Efficacy of memes is determined by that evoking and the relevant situating.

The efficiency of indexicals means that they can break, that the context and situatedness that help to afford meaning to the indexical can stretch it beyond breaking points. When indexicals break or are challenged, they raise questions of boundaries, frames, and identities. While the challenges to indexicals can be very subtle, when they break they break the entire frame of how a situation is understood. Americans were made aware of this when Senator Jim Jeffords of Vermont switched his party affiliation from Republican to Independent (and thereby altered control of the US Senate). His speech of explanation was a discussion of his perceived limits to the indexical "republican." If memes are understood to be indexicals, then their success or failure is not marked by evolutionary inheritance but rather by the longevity of their efficacy. That longevity is in many ways determined by the situation and the meaning "carrying capacity" of the meme indexical.

## 2.2 Frames

Donati (1992) observes that people frame an object around which an issue revolves rather than the issue itself, and that the study of frames involves identifying how people understand an issue, rather than determining if they are "for" or "against" a proposition. Gratton (2000) notes, working to create a shared and coherent meaning in line with corporate aspirations demands an understanding of the current meanings within the organization. We strive to interpret our world, not simply by imposing structure but by translating events and developing frameworks for understanding.

Frames are patterns of organized information by which people make sense of the world. These "patterns," "schemas," or "frames" form part of the "discursive universe" in which people interact with each other. People learn frames as they learn to use a language fluently and as they learn the narrative structures and ideologies present in the cultures that use that language. When people encounter new information or a new experience, they make sense of that information or experience by fitting it into an existing frame. Nevertheless, people will generally be able to fit any given collection of information into multiple frames; though, at the same time, they will also tend to perceive information selectively, focusing on details that most readily fit into the frames they know.

As humans, we seek to solve problems as presented; we acquiesce in their frames. Indeed, we become prisoners of the frame. Shira White (2002) tells an illustrative story of such prisons: Scientists have done some fascinating and suggestive experiments with ordinary houseflies. If you capture and keep houseflies in a jar and then remove the lid after a few days, most of them will not fly away. In fact, they stay right where they are — inside the jar — even though they could escape if only they could see their way to freedom. But they seem "committed" to a lid that is no longer there. Psychologists have identified this phenomenon as "premature cognitive commitment." It is premature cognition in the sense that it occurs, more or less automatically,

before we are aware of or fully understand the stimulus. It is "commitment" because we are locked into a specific set of thoughts. Like the houseflies, we give up the freedom to choose once we become committed to the nonexistent lid. The first step in challenging a commitment is recognizing that you have made it in the first place.

The American philosopher John Dewey prefigured this situation in his 1934 book, *Art and Experience* (Dewey 1934):

No matter how ardently the artist might desire it, he cannot divest himself, in his new perception, of meanings funded from his past intercourse with his surroundings, nor can he free himself from the influence they exert upon the substance and manner of his present being. If he could and did there would be nothing left in the way of an object for him to see.

### 2.3 Boundaries

The symbols and signs that we use to express meaning or hope will evoke meaning in others not only as communication devices but also as boundary setters. The words and the meanings that they evoke set up boundaries with regard to our ability to attend to, cognize, or be aware of aspects of our situation.

According to what can loosely be described as "boundary theory" (*Michaelsen & Johnson, 1997; Nippert-Eng, 1996a, b; Zerubavel, 1991*), individuals create and maintain boundaries as a means of simplifying and ordering the environment. "Mental fences" (*Zerubavel, 1991: 2*) are erected around geographical areas, historical events, people, ideas, and so on that appear to be contiguous, similar, functionally related, or otherwise associated. The process results in the creation of slices of reality domains that have particular meaning for the individual(s) creating and maintaining the boundaries. "Home," "work," and "church" are examples of the social domains created by boundaries (*Nippert-Eng, 1996a*). The boundaries are real in the sense that the individual perceives them as such and acts as though they are real (cf. *Weick, 1979*).

Although a given domain may be socially constructed and more or less institutionalized (e.g., people share a general consensus on what home means), Nippert-Eng (1996a, b) has shown that the boundaries around that domain are somewhat idiosyncratically constructed (e.g., one person allows home to cross over into work, whereas another keeps them separated). Further, by circumscribing domains, boundaries enable one to concentrate more on whatever domain is currently salient and less on other domains. (Ashforth *et al.*, 2000)

These boundaries can be triggered by repetition and word choice. Gould (2000) suggested that they were triggered by "canonical stories" — the shorthand for which are often labeled by the media as "memes."

The vertebrate brain seems to operate as a device tuned to the recognition of patterns. When evolution grafted consciousness in human form upon this organ in a single species, the old inherent search for patterns developed into a propensity for organizing these patterns as stories, and then for explaining the surrounding world in terms of the narratives expressed in such tales. As for mind, even when we can attribute a pattern to conventional nonrandom reasons, we often

fail to apprehend both the richness and the nature of those causes because the lure of canonical stories leads us to entertain only a small subset among legitimate hypotheses for explaining the recorded events. Even worse, since we cannot observe everything in the blooming and buzzing confusion of the world's surrounding richness, the organizing power of canonical stories leads us to ignore important facts readily within our potential sight, and to twist or misread the information that we do manage to record. In other words, and to summarize my principal theme in a phrase, canonical stories predictably "drive" facts into definite and distorted pathways that validate the outlines and necessary components of these archetypal tales. We therefore fail to note important items in plain sight, while we misread other facts by forcing them into preset mental channels, even when we retain a buried memory of actual events.

Word choices afford the possibility of new meanings, new analogies, and new insights, which, in turn, can lead to new or next activity. As people share framed information, they need not refer to all aspects of a frame directly to communicate which frame they have adopted to make sense of the information. Instead, they need only make reference to one dimension of a pattern to enable hearers or readers of their text to recall the whole frame. This evocative "power" is one of the attractive aspects of the meme concept. By viewing signifiers of meaning such as memes as mediums with a context dependence, we can see how the frames that result from word choice can work to limit or expand the very possibilities that we recognize as being afforded by our current situation. Word choice matters as a delimiter of possibility space as well as a means of communication. This provides the context in which mechanisms of memes operate.

### **3. Mechanisms [note 4]**

A word in context means both more and less than the same word in isolation: more, because it acquires new context; less, because its meaning is limited and narrowed by the context. The sense of a word... changes in different minds and situations and is almost unlimited. It is not merely the content of a word that changes, but the way reality is generated and reflected in a word. A complex is a word which does not function as a carrier of a concept but rather as a family name for a group of objects belonging together not logically but factually. (Vygotsky, 1986, page 245)

It is through language that we construct reality. With words we define, shape, and experience. Without the words to think, communicate, experience, or understand our lives would be very different from what they are. Words expand our consciousness but also limit us as we can only fully experience those things that we have the words for. Language provides the framework through which we perceive, experience, and act. As language constructs reality, so symbolization constitutes objects. Symbolization constitutes objects not conceptualized before, objects which would not exist except for the context of social relationships wherein symbolization occurs. Language does not simply symbolize a situation or object which is already there in advance; it makes possible the existence or the appearance of the situation or object, for it is a part of the mechanism whereby that situation or object is created. (Mead, 1934)

One example of this is the word "set," which has more than 100 meanings. The multiplicity of such meanings is the substrate for the mechanisms of imitation, transmission, and evolution that are "normally" ascribed to memes. Words evoke families of meanings. In Lissack and Letiche

(2004), these families of meanings are referred to as a glom. (Vygotsky, in his work, used a word that is usually translated as "complex." Lissack and Letiche opted for "glom" so as to avoid confusion.) The multiplicity of meanings implicit in a glom allows, when each such meaning is viewed as a medium, new possibilities for action. Vygotsky distinguishes between more primitive gloms — a word that does not function as a carrier of a concept, rather as a family name for a group of objects belonging together not logically but factually — and higher-level concepts. First come the gloms, and it is when abstracted traits are synthesized anew and the resulting abstract synthesis becomes the main instrument of thought that a concept emerges (Vygotsky, 1986).

### **3.1 Gloms versus Indexicals**

Gloms differ from indexicals. Gloms are primitive collections of families of meaning. Thus, when a child is learning about daddy going to the office, an entire realm of experience is built into the glom of "daddy's office," "office," and "going to the office." Only later will the child be able to separate the primitive wealth of experiences into distinguishable parts and associate a socially acceptable label with some of those parts to better "bind" the concept of daddy's office (the subway trip and its associated people and smells may be in the glom but will have been removed from the concept). By contrast, indexicals have no meaning independent from the situated context in which they are evoked.

Because indexicals have no inherent meaning independent from context, the use of an indexical is constrained by the variety of contexts in which it is deployed and the multitude of meanings from which the interplay between context and indexical is required to distinguish. The effective indexical serves as a medium to evoke meaning. The ineffective indexical will instead call attention to itself with the demand for further clarification. In essence, its ability to carry meaning will have been compromised. The overloaded indexical reveals itself via a lack of transparency to its medium-serving (medionic) functions and the implicit question of "this or that?" When something new is encountered (a perturbation) or emerges at another level, the prior sense of clarity in the fundierung between indexical and situation can break down, much like the tragedy of the commons as described by economists. When the context does not evoke a clarity of meaning and multiple meanings are possible, evoked, and present, the indexical is broken and what has been called a glom has been evoked instead.

The important observation is what Vygotsky says occurs when there is dissonance between the understood meaning of a concept and new input, whatever it might be. When a concept breaks down there is reversion back to the glom. That reversion allows for change. The dissonance produced thereby forces a reversion in the perceived meaning of the word. Context dependence takes over. "It is not merely the content of a word that changes, but the way reality is generated and reflected in a word" (Vygotsky, 1986, page 213).

### **3.2 Multiplicity of Meanings**

Inherent in the multiplicity of meanings is the recognition that only one meaning will be primary within the context of a given situated activity. That primary meaning will not be the solely representative meaning, but will take its primacy from the context. When there is coherence

between the situation and the meaning, the word choice will display a transparency with regard to medionic function. When that coherence is weak or absent, the very act of picking a label will demand some amount of attention. What coherence there is about the meaning, if it is to exist, would be forced to overcome or overwhelm such attentional demands.

The multiplicity of meanings undercuts the effective use of analogy as the word tokens of memes. Metaphors and analogies create constraints by focusing attention on that which is like and the resulting tendency by the user to attempt to justify the analogy. These constraints may not be readily apparent when the weakness of the analogy or the affinity is being exposed. Analogy involves inexact likeness. Butchvarov (1970, 1979) distinguishes between conceptual clarity and conceptual distinctness. Via analogy we can see the relative position as far as distinctness but can never achieve clarity. The former is a location in conceptual space and can be determined by noting similarities and differences between the entity being compared and other entities in conceptual space. However, clarity involves the content of the entity itself. It can be modeled, but the limitations of the model must be noted. Understanding is the desire for clarity, not merely distinctiveness. There is a remainder between the two. While categorization can suffice on distinctiveness, understanding cannot.

When we use analogy we are calling attention to some "like aspect" of two entities (call them source and object). If we were to dialogue with full disclosure about the analogy — this is similar, this is different; notice how the similars might react in situation x and contrast that with the differences, and so on — we would lose the shorthand and efficiency evoked by the analogy. Thus, we tend to allow our use of analogy to emphasize similarity over difference and substitution over care. In practice, when we assert that a is analogous to b, we often then make use of b as a label or category into which a falls. Affording such primacy to the similarities is to grant supervenience to the characteristics of the source at the expense of a fuller description of the object. When the similarities of a metaphor or analogy are allowed to supervene such that the analogy source is substituted in meaning for the description of the object, mistakes happen, possibility spaces are misconstrued, retrospective sensemaking might not make sense, and taken-for-granted fundierung relations may hide nasty surprises. To the extent that coherence is perceived, it may be based on fantasy. Indeed, much of the Internet/telecom bubble of the late 1990s seems to have been fueled by the supervenience of the characteristics of an "insatiable appetite" associated with convenience and newness over the demands for infrastructure, use, and value. The same pattern has been displayed by many of the bubbles documented in the history of economics.

### 3.3 Conceptual Slippage

Word choice and metaphor use allow for the emergence of new memes, the replacement of memes, and the death of memes via a concept that Douglas Hofstadter (1995) has labeled "conceptual slippage." In essence, the use of a metaphor or analogy evokes a glom of meanings. Each such use of metaphor is a perturbation to the existing self-referencing system (be it an individual, the organization, or some part thereof). The perturbations (please notice the plural) caused by the glom or gloms interact in multiple dimensions with the self-referenced core. As this series of interactions and resultant emergent behavior self-organizes, the principle of "least action" takes over. The basin of attraction that is the least demanding of energy is likely to

determine the "winning" meaning. The least-action principle suggests that the energy demands of attention or of the carrying of a full description are likely to be supervened by the efficacy of using an analogy, a label, or a name, even if incorrectly. Thus, one concept can slip to another via the energy demands of the least-action principle. The "whatever" of the current teen does not mean permission, tolerance, or inclusiveness, it means indifference — though most over-40s would not recognize that except after a series of painful experiences.

Thagard and Nerb (2002) make a similar claim to Hofstadter's conceptual slippage in describing emotional gestalts:

Thagard (1996: *Ch. 11*) described how dynamical systems theory can be applied to psychological phenomena by means of the following explanation schema: Human thought is describable by a set of variables. These variables are governed by a set of nonlinear equations.

These equations establish a state space that has attractors. The system described by the equations is chaotic. The existence of the attractors explains stable patterns of behavior. Multiple attractors explain abrupt phase transitions. The chaotic nature of the system explains why behavior is unpredictable. In the language of dynamical systems theory, the perceptual system has two attractor states, and the gestalt shift involves a phase transition from one attractor to the other. Analogously, we might think of an emotional state as a gestalt that emerges from a complex of interacting environmental, bodily, and cognitive variables, and think of emotional change as a kind of gestalt shift... Emotional gestalt shifts occur when changes in representations and their valences generate a new array of acceptances and valences that maximize constraint satisfaction differently from before. Through parallel constraint satisfaction, this shift may alter the acceptance status of other propositions.

When an emotional gestalt occurs, so too might conceptual slippage. Both undercut the effectiveness of a meme set in a new context.

Fauconnier and Turner (2002) go further in that they not only look for a slippage in conceptual meaning, but also for the activation of a new meaning. This is an extension of the emotional gestalt argument.

In any theory of meaning, activation does not come for free. The existence of frames, knowledge, experience, scenarios, and memories does not come for free. Ease of activation and degree of entrenchment by themselves impose very strong constraints on the imagination and the use of language. Linguists, logicians, and, for the most part, even psychologists tend to focus on the entrenched cases, which are already built and usually easy to activate. When only the rigid and entrenched patterns are used, meaning becomes predictable based on the mapping schemes and those patterns... Blends arise in networks of mental spaces which they call conceptual integration networks. Conceptual integration networks can have several input spaces and even multiple blended spaces. In conceptual integration, there is partial matching between input spaces of many kinds: connections between frames and roles in frames, connections of identity or transformation or representation, analogical connections, and metaphoric connections. In blending, structure from two input mental spaces is projected to a new space, the blend. Generic spaces and blended spaces are related: Blends contain generic structure captured in the generic

space but also contain more specific structure, and they can contain structure that is impossible for either of the inputs. Similarly, not all elements and relations from the inputs are projected to the blend. Thus, emergent structure can arise in the blend that is not copied there directly from any input. (Fauconnier and Turner 2002, page 168).

Blends, emotional gestalts, and conceptual slippages are all evidence of the least-action principle (lower energy expenditure) at work.

### 3.4 Least-Action Principle

Lower energy expenditure is the driving pursuit in the information space world (cf. Boisot, 1995). In Vygotskian terms, a group and its members begin with some existing set of concepts and they encounter change. The encounter reduces some of the concepts to the status of gloms, and in such a status, the possibility arises for new conceptual understanding to emerge. This understanding will be influenced by the metaphors and analogies available to label the gloms, for in the adjacent meanings implicit in the metaphors is the potential synthesis represented by the new concept. The premises of least action suggests that a context-dependent glom is an efficient vehicle (in the same manner that Perry and Barwise suggest that indexicals are efficient), provided that supervenience is possible. This is because we use words as tokens and allow context to evoke meaning from among the gloms represented thereby. If supervenience is possible, then such evoked meanings are triggered by the situated activity in which they occur. By contrast, gloms will not work well in a system that is dependent on representations, reductions, and causality. In such a world, evoked meanings become reified and are carried across new situated activities. Dissonance from the mismatch is the likely result.

To a group member, context includes ongoing change — which then disrupts the shared-context content of existing codification and disturbs the agreed meanings of abstractions. A key least-action observation is that personal coding of meaning is transformed within an organization into institutionalized codification, so as to both maximize the value of shared meaning and minimize the need for the energy expended to transmit shared context. Emergent change erodes the ability of codification to hold. In the absence of an offsetting response to this erosion, institutional codification recedes to personalized coding, and the ability of common abstractions to transmit shared meaning deteriorates. Concepts become gloms. Such disturbances can have an emergent character that itself is disturbing, because the cumulative effects thereof cannot be predicted or planned for. This lack of prediction or planning poses a threat to coherence. And coherence preservation is another energy-conserving action within the information space.

Thus, we have a mechanism for meme success and failure. Emergent change occurs in the environment. In Vygotskian terms, the dissonance introduced by emergent change forces previously accepted concepts to recede to gloms. Uncertainty of meaning is introduced. For our purposes, uncertainty can be regarded as a label better defined as the inverse of one's propensity to act (Dretske, 1981; Fransman, 1994). Given uncertainty's threat to coherence, organizations must find a way to combat its increase, for uncertainty is a significant energy drain running counter to the principle of least action. Increases in uncertainty can be attributed to loss of identity, to a perceived need for more and "better data," and to an increase in the perceived threat from taking an incorrect action.

This translates into the lack of a well-understood model of the possibility space and thus the substitution of a need to search for a willingness to act. If identity is to be preserved, then there must be an offsetting emergent response to rebuild context so as to replace the content lost to uncertainty (i.e., that which was contained in the institutional codifications and abstractions that have now encountered disconfirming notions and been forced to revert to the more primitive gloms of meaning). Success is related to the evolution of the ideational niche for which the meme is a token. If that niche has failed, so too will the meaning-evocation powers of the token.

The successful meme is one whose indexical quality can bridge both the old context and the new, such that the users of the meme token can dialogue about the meanings evoked by that token without asserting incommensurability. The unsuccessful meme is one whose indexical quality cannot bridge the gap between contexts and thus cannot make the transition to new context and new situation.

#### **4. Implications**

We need to do memetics to demonstrate when, where and how memetics has a relative and relevant advantage over social science devoid of memetics. The future of memetics will not be decided by those talking about memetics, whether grand theorising or armchair philosophy about the evolution of culture, history, consciousness or how we think, but will be decided by those doing memetics and demonstrating its relevance. (Hales & Marsden, 2002)

A memetics that accepts memes as indexical catalysts and tools can demonstrate the advantage that Hales and Marsden seek. Such a memetics allows for study of the content of information and its use, with a focus on processes and mechanisms vastly different from what passes today as information science, knowledge management, or linguistics. This is not a memetics that studies the evolution of memes per se, for the ontological status of memes is changed within it.

However, such a memetics can demonstrate relevance, advantage, and application.

For example, if this approach were adapted to an extension of Salingaros and Mikiten's (2002) exploration of modernism as an architectural meme, the discussion would explore the environmental niche in which the qualities laid out for the success of the modernism thrive. This would be followed by an exploration of what potential risks for the success of the meme lie within and without that niche, and what factors of the meme and/or the niche contribute to its ongoing resilience. Once the risks and resilience factors have been so identified, they can be mapped to other domains and compared with the success/failures of other memes both within and without the architectural domain. This seems far more fruitful an approach to making social science advances than the mere mapping of modernism in architecture as a meme (a mapping that allows critics to reply "So what?").

Edmonds' (2002) three challenges can be answered by this revised form of memetics. For example, his first challenge argues that a "conclusive case study" would "clearly demonstrate that there is at least one cultural process that is of an evolutionary nature, where 'evolutionary' is taken in a narrow sense." If the requirement that it is the meme that must be of an evolutionary nature is dropped, then Edmonds' challenge is easily fulfilled. Anthropology and sociology can

document hundreds of cultural processes that are evolutionary and many of these will have a history of successful memes associated with them. What is difficult for meme as replicator is much easier for meme as indexical catalyst.

Lissack and Letiche's forthcoming *Coherence Emerges: A Complexity Theory of Organization* (2004) is an example of work that meets Edmonds' second challenge. So too does much of the case-study work on organizational symbolism. Edmonds' third challenge is perhaps incommensurate with the revisions suggested above. Memetic processes of the catalytic indexical variety are easily found and documented in the "real" world and are not in need of "simulation." Such "real" examples should, in any case, be considered as a firmer foundation for an applied theory than simulations could provide. Memetic processes of the catalytic indexical variety also seem to address many of the concerns raised in Bloch (2000) and Kuper (2000).

This redefinition of memes recognizes that they are efficient tools for evoking particular affordances to be attended to in situ. Such a definition is consistent with theories of niche construction. This article suggests that Dawkins created an indexical (meme=gene) and it has exceeded its carrying capacity and thus lost its efficacy. Worse, that indexical is evoking images and affordances that stand in the way of the memetics field making true progress. It is time to recognize that ontic status has been misplaced. Memes need a new meme: meme as catalytic indexical.

For managers, memes defined as catalytic indexicals raise the potentialities offered by other catalysts — the provisioning of an environment with a catalyst can afford the possibility of a transformation that is much more difficult than without the catalyst's presence. Memes would be studied for their catalytic roles and managers would be taught sensitivity to the conditions that aid and hinder the evolution of such catalysts.

As catalytic indexicals, memes can be meaningfully assigned explanatory and causal roles — the very ingredients that Edmunds claims memetics needs, and the qualities that managers are often seeking.

## Notes

1. This notion does not conflict with the definition of meme in the Oxford English Dictionary: An element of a culture that may be considered to be passed on by non-genetic means.
2. An application of this argument is forthcoming in Lissack & Letiche (2004).
3. Cf. T. Kapitan's "Autonomy of indexical reference" and related works can be found at: <http://www.soci.niu.edu/~phildept/> and G. Nunberg's "Indexicality and deixis" (1993) and related work at: <http://www-csli.stanford.edu/~nunberg/linguistics.html>.
4. Much of the mechanism argument was first developed in Lissack & Roos (1999) and has been expanded in Lissack & Letiche (2004).

## References

- Ashforth, B., Kreiner, G., & Fugate, M. (2000) All in a day's work: Boundaries and micro role transitions, *Academy of Management Review*, **35**:472–91.
- Barwise, J. & Perry, J. (1983) *Situations and Attitudes*, Cambridge, MA: MIT Press.
- Bloch, M. (2000) A well-disposed social anthropologist's problem with memes, in R. Aunger (ed.) *Darwinizing Culture: The Status of Memetics as a Science*, Oxford, UK: Oxford University Press.
- Boisot, M. (1995) *Information Space*, Oxford, UK: Oxford University Press.
- Braun, D. (2001) Indexicals, In the Edward N. Zalta (ed.) *Stanford Encyclopedia of Philosophy* (Fall 2001 Edition), <http://plato.stanford.edu/archives/fall2001/entries/indexicals/>
- Butchvarov, P. (1970) *The Concept of Knowledge*, Evanston, IL: Northwestern University Press.
- Butchvarov, P. (1979) *Being qua Being: A Theory of Identity, Existence and Predication*, Bloomington, IL/London, UK: Indiana University Press.
- Dawkins, R. (1976) *The Selfish Gene*, Oxford, UK: Oxford University Press.
- Dewey, J. (1934) *Art as Experience*, New York: Perigree.
- Donati, P. (1992) Political discourse analysis, in M. Diani & R. Eyerman (eds), *Studying Collective Action*, London: Sage.
- Dretske, F. (1981) *Knowledge and the Flow of Information*, Oxford, UK: Blackwell.
- Edmonds, B. (2002) Three challenges for the survival of memetics, *Journal of Memetics: Evolutionary Models of Information Transmission*, **6**(2). [http://jom-emit.cfpm.org/2002/vol6/edmonds\\_b\\_letter.html](http://jom-emit.cfpm.org/2002/vol6/edmonds_b_letter.html)
- Fauconnier, G. & Turner, M. (2002) *The Way We Think: Conceptual Blending and The Mind's Hidden Complexities*, New York, Basic Books.
- Fransman, M. (1994) Information, knowledge, vision and theories of the firm, *Industrial and Corporate Change*, **3**(3):713–57.
- Gould, S. J. (2000) Jim Bowie's letter & Bill Buckner's legs, *Natural History*, **109**:26–40.
- Gratton, L. (2000) *Living Strategy: Putting People at the Heart of Corporate Purpose*, London: Financial Times/Prentice Hall.
- Greeno, J. & Moore, J. (1993) Situativity and symbols, response to Vera and Simon, *Cognitive Science*, **17**:49–59.

Hales, D. & Marsden, P. (2002) Editorial: Memetics malaise?, *Journal of Memetics: Evolutionary Models of Information Transmission*, **6**(2). <http://jom-emit.cfpm.org/2002/vol6/editorial.html>

Hargadon, A. & Fanelli, A. (2002) Action and possibility: Reconciling dual perspectives of knowledge in organizations, *Organization Science*, **13**(3):290–302.

Hofstadter, D. (1995) *Fluid Concepts and Creative Analogies*, New York, Basic Books.

Kapitan, T. (2003) Autonomy of indexical reference,  
[http://www.soci.niu.edu/~phildept/speakers/Kapitan\\_Indexicals.html](http://www.soci.niu.edu/~phildept/speakers/Kapitan_Indexicals.html)

Kaplan, D. (1989) Demonstratives, in J. Almog, J. Perry, & H. Wettstein (eds), *Themes from Kaplan*, New York: Oxford University Press: 481–563.

Kuhn, T. (1969) *The Structure of Scientific Revolutions*, Chicago: University of Chicago Press.

Kuper, A. (2000) If memes are the answer, what is the question?, in R. Aunger (ed.), *Darwinizing Culture: The Status of Memetics as a Science*, Oxford, UK: Oxford University Press.

Laland, K. N., Odling-Smee, J., & Feldman, M. W. (1999) Niche construction, biological evolution and cultural change, *Behavioral and Brain Sciences*, **23**(1):131-175 .

Laland, K. & Odling-Smee, J. (2000) The evolution of the meme, in R. Aunger (ed.), *Darwinizing Culture: The Status of Memetics as a Science*, Oxford, UK: Oxford University Press.

Lissack, M. & Letiche, H. (2004, forthcoming) *Coherence Emerges: A Complexity Theory of Organization*, Cambridge, MA: MIT Press.

Lissack, M. & Roos, J. (1999) Words count: Viewing organizations as emerging systems of languaging, <http://lissack.com/writings/AMRfinal.pdf>.

McGrenere, J. & Ho, W. (2000) Affordances: Clarifying and evolving a concept, *Proceedings of Graphics Interface 2000*: 179–86. <http://www.graphicsinterface.org/proceedings/2000/177/>

Mead, G. H. (1934) *Mind, Self and Society*, Chicago: University of Chicago Press.

Michaelsen, S. & Johnson, D. (1997) *Border Theory: The Limits of Cultural Politics*, Minneapolis: University of Minnesota Press.

Millikan, R. (1993) *White Queen Psychology and Other Essays for Alice*, Cambridge, MA: MIT Press.

- Nippert-Eng, C. (1996a) Calendars and keys: The classification of ‘home’ and ‘work,’ *Sociological Forum*, **11**(3):563–82.
- Nippert-Eng, C. (1996b) *Home and Work: Negotiating Boundaries through Everyday Life*, Chicago: University of Chicago Press.
- Odling-Smeel, F. J., Laland, K., & Feldman, M. (2003) *Niche Construction: The Neglected Process in Evolution*, Princeton, Princeton University Press.
- Prigogine, I. (1997) *The End of Certainty, Time, Chaos and the New Laws of Nature*, New York: Free Press.
- Prigogine, I. & Stengers, I. (1984) *Order out of Chaos: Man’s New Dialogue with Nature*, New York: Bantam.
- Salingaros, N. A. & Mikiten, T. M. (2002) Darwinian processes and memes in architecture: A memetic theory of modernism, *Journal of Memetics: Evolutionary Models of Information Transmission*, **6**(1). [http://jom-emit.cfpm.org/2002/vol6/salingaros\\_na&mikiten\\_tm.html](http://jom-emit.cfpm.org/2002/vol6/salingaros_na&mikiten_tm.html)
- Simon, H. A. (1994) Literary criticism: A cognitive approach, in G. Guzeldere & S. Franchi (eds), *Bridging the Gap*, special supplement to *Stanford Humanities Review*, **4**(1):1–26.
- Stanford Encyclopedia of Philosophy, <http://plato.stanford.edu/>.
- Thagard, P. (1996) *Mind: Introduction to Cognitive Science*, Cambridge, MA: MIT Press.
- Thagard, P. & Nerb, J. (2002) Emotional gestalts: Appraisal, change, and emotional coherence, *Personality and Social Psychology Review*, **6**:274–82.
- Vygotsky, L. (1986) *Thought and Language*, A. Kozulin (ed.), Cambridge, MA: MIT Press.
- Weick, K. (1979) *The Social Psychology of Organizing*, Reading, MA: Addison-Wesley.
- White, S. P. (2002) *New Ideas about New Ideas: Insights on Creativity from the World’s Leading Innovators*, New York: Perseus.
- Wilkins, J. S. (1998) What’s in a meme? Reflections from the perspective of the history and philosophy of evolutionary biology, *Journal of Memetics: Evolutionary Models of Information Transmission*, **2**(1). [http://jom-emit.cfpm.org/1998/vol2/wilkins\\_js.html](http://jom-emit.cfpm.org/1998/vol2/wilkins_js.html)
- Zerubavel, E. (1991) The great divide. The social lens, in E. Zerubavel (ed.), *The Fine Line: Making Distinctions in Everyday Life*, Chicago: University of Chicago Press.

# UNITS, EVENTS, AND DYNAMICS IN MEMETIC EVOLUTION

By  
**Aaron Lynch**  
aaron@mcs.net

[NOTE: equations reproduced flawed]

## Abstract

An evolutionary recursive replicator theory of mental/brain information is presented. With all replicator theories resting at least tacitly upon the fundamental notions of causation and of calling two or more entities "the same" with respect to an abstraction, the concept is rendered explicit in defining the terms "mnemon" and "meme". It is argued that memetics may have no "absolute" system of memory abstractions much as physics has no absolute coordinate system (framework of space-time abstractions). A symbolic calculus of mnemonic conjugations and replication events is introduced. The term "meme" is given a technical definition, and reasons are offered for avoiding more expansive definitions. Arguments that meme sets are generally only partially ordered then provide a formal reason for rejecting mnemonic "size" as a crucial element in defining the word "meme". Differential equations are developed for meme host population versus time in a two-meme system, modeling the dynamics whereby events at the individual level give rise to trends at the population level. This lays a foundation for computerized simulations and the falsification or validation of specific memetic hypotheses, and for testing population memetics theory with animal experiments. As memetic hypotheses generally involve observable communication events, they are found to have stronger empirical standing than hypotheses involving unidentified genes. Mechanisms of creativity as a population phenomenon are examined, with memetic analysis yielding a novel explanation for the temporal clustering of independent co-creations. Creation and propagation are integrated into a theory of evolution by variation and natural selection of memes.

**KEYWORDS:** Meme, mnemonic, evolution, replication, abstraction, transmissivity, receptivity, longevity, recursive algorithm, differential equation.

## 1 Introduction

In the last chapter of his 1976 book *The Selfish Gene* [3], Richard Dawkins introduced the term "meme" for lay audiences as "a unit of cultural transmission or a unit of imitation". He went on to explain that this covers ideas, tunes, fashions, and other information stored in brains. In *The Extended Phenotype* [4] Dawkins more explicitly clarified that "a meme should be regarded as a unit of information residing in the brain"<sup>1</sup>. I have since taken this meaning of the word "meme" and refined it into a more technical definition suitable for symbolic and mathematical analysis [11]. In arguing that meme theory is not only sound but also broadly unifying, my book *Thought Contagion* [12] went on to use a great variety of non-technically presented examples to show the relevance of meme theory to important human phenomena. Yet the further development of strong theoretical treatments, formal definitions of terms, mathematical analyses, discussions of empirical methodology, and criteria for falsifiability remain subjects of interest among hard core

scientists. This paper revisits the theoretical issues and the underlying philosophy of science, and elaborates on topics of empirical methodology and falsifiability.

## 2 Non-Metaphoric Memetics

In memetics, organic evolution, biological contagions, plasmids, and computer viruses have served as metaphors for the spread of ideas. In principle, however, someone could have offered a treatise on the replication, proliferation, and evolution of ideas long before Darwin's work on evolution, Pasteur's work on microbiology, Mendel's work on genetics, or Turing's work on automata. Indeed, that treatise might even have afforded metaphors to these other disciplines had it appeared by the early 1800's. On this planet, though, the history of science has its social and information fields drawing metaphors from biology rather than the other way around.

Like all metaphors, the ones linking beliefs to organic and software contagions have their limitations. Metaphors make powerful literary and explanatory devices, and can render a potentially dry and abstruse subject interesting and readable. They also serve to spark important new scientific insights. But relying too heavily on them can lead to bad science. It makes sense, then, to recast the core concepts of memetics theory in language that does not depend on analogies to the biological or computer sciences. The analysis is drier, but it does formally illuminate some ways that memes depart from strict analogy to other replicators. It also provides some philosophy of science behind memetics theory while offering a stronger platform for mathematical and empirical investigations.

## 3 Units of Memory Replication

Defining the "unit of imitation" is as critical to memetic evolution theory as defining the units of heredity was to genetic evolution theory. The gene's definition developed as empirical investigations led away from the hypothesized inheritance of acquired characteristics to the biochemical sequencing of DNA strands. The universal code of these strands constituted a natural "language" upon which scientists based their own more symbolic and abstract language. Yet for the evolution of ideas, no equally understood concrete language has been discovered. Science has achieved no direct observation of the neural encoding of ideas, which might have provided us a precise language for discussing ideas. Indeed, even if we knew in principle how to express the ideas of a single person in terms of neurons, synapses, etc., the description would likely be prohibitively complex. Moreover, as Dennett points out in *Darwin's Dangerous Idea* [5], it would be flabbergasting to find that the brain-cell complex that stores an idea in one person is the same as or very similar to the brain-cell complex that stores the idea in all who have that idea<sup>2</sup>. So instead of language based on a concrete mechanism of information storage, we must settle for an abstract representation of the information stored. Thus, memory abstractions form the basis for memetic evolution theory.

Most people use abstract representations of memory content on a daily basis to discuss ideas. When we say that two people have "the same" idea, we do not use "sameness" to mean equality in every concrete detail, or else we could never correctly say that two people have "the same" idea. As Dawkins [3] put it, "If this were not so, then almost any statement about two people agreeing with each other would be meaningless". What we mean by saying that two people have

"the same" idea is that one person's idea has at very least one quality in common with the other's idea. Perceiving two people to have "the same" idea involves abstracting out a set of common qualities. So saying that two people's ideas are "the same" only means that they are in some way "of the same kind".

This ability to say that two people have the same idea lies at the foundation of the notion of a "replicating idea". When an idea "replicates", it acts to produce or preserve ideas that we call "the same idea". The resultant ideas can for now be called "self-replicated" ideas. To be a self-replicated idea means not only to have resulted from a given idea, but also to be "the same" as that idea. Specifically, it means meeting some abstract, observer-defined criterion for sameness. Self-replicated ideas are not all exact replicas of their originals. A wide range of ideas may result from each self-propagating original. An observer just selectively lumps these proliferated ideas along with their original(s) into a set, using an abstract inclusion criterion.

As an example, the belief that "abortion is wrong" has a wide range of different meanings to different people. The range includes people who regard abortion as merely unethical to those who see the "morning after pill" as a high felony. So, although the belief varies greatly from person to person, its occurrences are all "the same" in the limited respect of fitting the above definition. Encountering a range of such beliefs in actual people, one "abstracts out" a common element running through all of them. On writing out a definition of this abstracted common element, or sameness criterion, one can proceed to use it for natural selection reasoning.

The abstraction could just as well be more general or more specific, depending upon our interests. A memeticist might, for instance, choose the more specific belief "abortion is a mortal sin". Then many quantitative variables involved in its natural selection would differ from those associated with the broader definition. First, the more restrictive definition would almost certainly identify a smaller host population. Second, when the hosts communicate their belief to friends and family, a likely smaller fraction of listeners will become new hosts per exposure. This is because any listeners who go away newly convinced that abortion is merely unethical no longer count as new hosts. Becoming a host of the more restrictively defined belief takes more of what Dawkins [3] calls *copying fidelity*. Third, once a host, one may do less "ideological wavering" before being counted as a drop out. That is, remaining the host of a more restrictively defined belief takes more preservation fidelity. So changing an idea's definition can make a big difference in the evolutionary phenomena identified with it.

## 4 Other Propagating Items

Many psychological phenomena other than ideas can be observed to self- replicate. These include certain habits, attitudes, class identities, cognitive associations, education, emotional dispositions, addictions, and even neurotic and psychotic symptoms.

All these traits may be broadly classified as human memory content. This category is more general than the word "idea" or even "memory" usually connote. It includes everything in the fairly broad meaning of "memory", as defined in Merriam Webster's Collegiate Dictionary, 10th edition: "the store of things learned and retained from an organism's activity or experience as evidenced by modification of structure or behavior or by recall and recognition". Thus, the

principle abstractions manipulated with memetics theory are memory abstractions, or *mnenoms*<sup>3</sup>. Mnemos do not include inanimate propagating items such as chain letters, Bibles, etc. Nor do they include traits considered genetically instinctual.

Using mnemos helps to standardize the measurement of propagation in terms of a host count. Thus, a chain letter or the copying machine duplicating it do not count as hosts, but the person photocopying the letter does. The relationship of artifacts to mnemos is discussed later. If a mnemon resides very redundantly in someone's brain, that person still counts as only one host and one mnemon instantiation. The number of duplicates of a memory item in one brain is not currently measurable, so it escapes further treatment in this article.

## 5 Representing Mnemos Symbolically

Mnemos can be represented conveniently with symbols such as "*A*", "*B*", etc. Thus, we can call the hell belief mnemon *A*, the imminent doomsday belief mnemon *B*, and a combination of mnemos such as the hell/imminent doomsday combination "*A\*B*". The "\*" indicates that *A* and *B* are instantiated in the same host. Extending this concept, one can represent a whole system of mnemos as "*A\*B\*C\*..*".

Symbolic expressions can also represent mnemon replication much the way chemists represent chemical reactions. Thus, the hell belief's non-parental conversion propagation may be represented as *A+~A2A*. This expression is read as "Host of *A* together with a non-host of *A* yields two hosts of *A*". (The "~" indicates only non-host of *A* status, and does not of itself imply hostship of a contrary belief.) The two hosts on the arrow's right are the same two people as on the left side, one of whom is converted from non-host to host status. (The word "horizontal" may also be used, to describe this kind of non-parental event, as long as it is not taken to refer only to transmissions between people of similar social status.) The mnemos on the left side of the arrow are called the input mnemos, and those on the right, the output mnemos.

A transition such as *A+~A2A*, realized in particular people at a particular time, constitutes an evolutionary event in the ideosphere<sup>4</sup>. Other evolutionary events include: *A~A* (host of *A* drops out), *~AA* (non- host independently forms *A*), *A0A* (host of *A* dies), and *2A2A+~A* (two hosts of *A* have a baby non-host of *A*).

Even multistage evolutionary events are handily represented by this system. For instance, the childbirth event can be extended to the form *2A2A+~A3A*. (Two hosts of *A* have a baby and then spread their *A*-mnemon to the child.) A more complicated possibility is *A+~AA+2~A 2A+~A*. (Host plus non-host of *A* have a baby who later adopts the *A*-mnemon from the *A*-host parent.)

Some multistage events are best represented by two or more diagrams. For example, if two hosts of *A* have a baby non-host and then have their local Sunday school teacher instill the *A*-mnemon in the child, one would represent it as two events: *2A2A+~A* (the birth) and *3A+~A4A* (the catechism). In the second event, the three input *A*-hosts are the teacher and the two parents (who select the teacher). Representing the events separately (instead of as *3A3A+~A4A*) conveys with greater clarity that the teacher's role happened only after the birth. Many two stage events pertain

to the parental replication because that mode generally involves both having and training children.

## 6 Complementary Mnemons

Treating  $\sim A$  as a mnemon along with  $A$  may seem rather strange. People labeled " $\sim A$ " may be called  $\sim A$ -mnemon hosts or  $A$ -mnemon non-hosts. ( $A$  and  $\sim A$  are called complements of each other.) But can a person's lack of the  $A$ -mnemon justifiably be called a mnemon itself? It sounds like saying that nothing is something.

The meaning of the term "mnemon" provides an answer to this question. Mnemons are merely memory abstractions. As such, a negatively defined mnemon which only states what a person does not have is just as much of a memory abstraction as is a positively defined mnemon. Saying that someone "has" or "hosts" a mnemon like  $\sim A$ , really means that the person satisfies the definition of the particular memory abstraction (i.e., the person instantiates the abstraction). This is what we mean when we say that a person "hosts" a positively defined mnemon. The only difference is that the term "memory" is expanded in a way that somewhat resembles its usage in computer science: it now includes some (partly) "empty memory" states that are not *specifically* determined by genetically based instinct.

Negatively defined mnemons can also self-propagate, as with the childbirth event  $2\sim A 3 \sim A$ , where  $A$  is a knowledge-of-birth-control mnemon. (One might define this mnemon functionally as knowing how to use at least one of some particular list of methods.) Because  $\sim A$ -hosts have fewer birth control options than do  $A$ -hosts, the  $\sim A$ -mnemon enjoys a greater quantity-parental replication advantage. Nonetheless, the  $A$ -mnemon has done very well in modern times, largely since people can be taught about birth control far more easily than they can be made to forget about it. Since proselytizing it is impossible, the  $\sim A$  mnemon depends on parental events that occur just a few times per generation. The  $A$ -mnemon, on the other hand, proliferates by non-parental conversion for various reasons, including both the sexual and humanitarian motives of its hosts. So both  $A$  and  $\sim A$  self-propagate, leaving us no choice but to consider the propagation of both a positively defined mnemon and a negatively defined mnemon in investigating this example.

Incidentally,  $\sim A$ 's parental propagation does not conform to the tendency, mentioned earlier, for the parental replication mode to involve multistage evolutionary events. People are born with  $\sim A$ , so that giving birth and "imparting" the mnemon are actually the same event. After birth, the parents can at most act to preserve  $\sim A$  in their child.

## 7 Mnemon Combinations

Multi-mnemon events describe phenomena like idea alteration or recombination too complex to represent with one complementary mnemon pair alone. For example, forming the hell/imminent doomsday pair by recombination can be represented as  $A * \sim B + \sim A * B A * \sim B + A * B$ , where  $A$  = the hell mnemon and  $B$  = "doomsday is imminent". In this event, a person with only the  $A$ -mnemon spreads it to someone with only  $B$ , so that the latter person now has both  $A$  and  $B$ . This new combination may very well spark some new ideas in the  $A * B$ -host. He might experience the

event  $A*B*\sim CA*B*C$ , where  $C$  is the belief that "I must urgently spread my faith to others because it may soon be too late!" Consequently, he may repeatedly play the leading role in the event  $A*B*C+\sim A*\sim B*\sim C2A*B*C$ .  $A$ ,  $B$ , and  $C$  act cooperatively here to bring about their collective propagation.

Mnemons  $A$ ,  $B$ , and  $C$  are all rather "unpleasant" ideas, so it does not seem too surprising to find them propagating cooperatively. Yet in the real world, these three propagate cooperatively with a very "pleasant" mnemonic  $D$ : "Love your neighbor as yourself".  $A*B*C*\sim D$  motivates its hosts to spread their faith only to those "unbelievers" that they really care to see saved from hell. In marked contrast,  $A*B*C*D$  hosts should be motivated to spread their faith to any unbelievers they should happen to meet. So the combination of mnemons probably spreads more vigorously due to the inclusion of the  $D$ -mnemon.

When mnemons propagate synergistically as do  $A$ ,  $B$ ,  $C$ , and  $D$ , they may propagate primarily as a set rather than individually. In such cases, the set may be usefully defined as one mnemonic. Thus, one might define  $E=A*B*C*D$  and consider  $E$  to be a stable, propagating mnemonic in its own right. This approach can sometimes be used to study very large ideological systems, such as religious and political doctrines, by treating them as single (but lengthily defined) memory abstractions. The propagation events, however, often contain many intricate stages.

Moreover, in specifying mnemonic " $X$ " by an equation like  $X=A*B*C*...$ , one runs the risk of specifying so many "little" constituent mnemons that no one person ever actually has all of them. Mnemon  $X$  would then be a useless abstraction. This symbolic method allows its own hosts enormous freedom to specify their preferred abstractions, yet writing down symbols does not guarantee all those abstractions to be useful in studying the real world.

## 8 Competing Mnemons

Competition rather than cooperation characterizes many interacting mnemons. A mnemonic competes against its complement in the knowledge-of- birth-control case mentioned above. The two complementary mnemons each "armed" with distinct replication advantages "struggle" for host population.

Yet more heated than this competition, at least in some nations, is the competition between moral stances on the subject. Personal morality, after all, has a greater bearing on an adult's use or non-use of various methods in a society where birth control information is highly available. Moreover, morality strongly influences the parental decision to actively instruct children on birth control or to actively obstruct their learning on the subject.

So in addition to the knowledge-of-birth-control mnemonic (mnemon  $A$ ), we should also consider mnemonic  $B$ , the moral acceptance of practicing birth control, and mnemonic  $C$ , the belief that birth control is immoral. Here people are born with the  $\sim A*\sim B*\sim C$  combination, but they never end up simultaneously having  $B$  and  $C$  later on. This is because  $B$  and  $C$  are *contrary*, as distinct from complementary mnemons. Since they are mutually exclusive, they are destined to have a competitive rather than a cooperative relationship in the population. The  $B$  mnemonic is favored by a high occurrence of the non-parental conversion event  $B+\sim B2B$ , while  $C$  is favored by a high

rate of the parental event  $2C \rightarrow 2C + \sim C$ .  $B$  benefits spreads due to the same kind of sexual and humanitarian motives for spreading it as those shown by the  $A$  mnemonic.  $C$  spreads "parentally" because people who believe that birth control is immoral do tend to have more children.

Economic motives for spreading and adopting the two mnemonics also exist, and vary among situations. Additionally, the drop out rate of children raised as  $C$ -hosts varies with population homogeneity, media exposure to  $B$ , etc.

## 9 Homogenic and Heterogenic Events

The birth control example as treated so far fails to acknowledge people's ability to independently invent or re-invent moral decisions on their own. People are portrayed as acquiring mnemonics only by copying them from others or by being born with them (as in the case of negatively defined mnemonics).

To remedy this omission, consider the cases of two students who learn about birth control methods and then make moral judgments based on what they have learned. The first decides the practice is morally acceptable ( $A^* \sim BA^*B$ ). The second decides it is not ( $A^* \sim CA^*C$ ). Of course, people can make moral judgments without knowing birth control methods, but in this case our two hypothetical students judged only upon gaining the knowledge.

So one mnemonic precipitates the formation of another mnemonic, instead of just a new copy of itself. Mnemonics that do this are called *heterogenic*, or "other-forming". The corresponding events, such as  $A^* \sim BA^*B$ , are called *heterogenic events*. The other type of event, which produces copies of input mnemonics, is appropriately called a *homogenic*, or "same-forming" event. The input mnemonic that gets copied is called a *homogenic mnemonic*.

Mnemonic event diagrams and terminology now acknowledge that people can form opinions without simply copying others' ideas. Yet on many topics, people copy more ideas than they either originate or "re-originate". For many beliefs, such as the birth control taboo, homogenic formation far outweighs heterogenic formation in its overall occurrence rate. The reason is that homogenic events have the tendency toward self iteration while heterogenic events do not. Generally speaking, anything that increases the availability of the input mnemonics will increase the expected frequency of the event. Homogenic events, such as  $2A \rightarrow A + \sim A$ , increase their own input mnemonics. So they often tend to increase their own likelihood of recurring. When they do, they exhibit the *recursion principle*, which can lead to vast increases in mnemonic prevalence. In contrast, heterogenic events, such as  $A^* \sim BA^*B$ , actually decrease their own input mnemonics by one with each occurrence. This tends to decrease the heterogenic event's recurrence rate. Yet the independent thinker can pass the new, *heteroderivative* mnemonic along, resulting in homogenically formed, or *homoderivative* mnemonics. So after the first few heterogenic formations of the birth control taboo, the formation of new taboo hosts tends to become rapidly predominated by recursive, homogenic events.

One mnemonic whose host population accumulates many members through both homogenesis and heterogenesis is the belief that "It is best not to go to church on Sunday". Many of its hosts received it (homogenically) from parents or friends who already had the idea. This makes their mnemonics homoderivative.

However, many people got the idea because as children their parents made them go to church every Sunday, even when they did not feel like going. The resulting aversive experiences often lead the children to conclude that it is best not to go to church on Sunday. Ironically, their belief results from their parents' strict adherence to exactly the opposite belief. The pro-churchgoing mnemonic influences some parents to generate something radically different in some of their children. So those parents' mnemonic is heterogenetic while the children's mnemonic is heteroderivative. Any parent who raises some children to be church going and some to be church avoiding has a churchgoing mnemonic that is both homogenetic and heterogenetic.

The concepts of replicator evolution and epidemiology directly pertain only to the homoderivative sector of a mnemonic's host population. In this sector, we can properly refer to mnemonics as the "units of imitation" discussed by Dawkins. This allows a very restricted denotation in which a meme is *defined* as a *homoderivative mnemonic*. The definition of a particular meme contains an abstract sameness criterion like the ones defining mnemonics, but also includes a criterion of causality, namely, that it is homoderivative. So each meme has a corresponding mnemonic: the memory abstraction defined without reference to homoderivative causation.

Although the church-avoiding mnemonic does not occur primarily as a meme, one can still understand its proliferation in terms of memes. Simply divide the hosts into two groups: hosts of the church avoiding meme and dropouts of the church going meme (meme and meme-derived mnemonic hosts). The first group can be studied directly as meme hosts. The second group's growth rate can be studied as a function of the church-going meme frequency. So the study of meme proliferation can be valuable to understanding the growth or decline of both of these constituent groups, whose combined numbers include nearly all the church-avoiding host population.

Some mnemonics are neither memes nor meme-derived in most of their instances. For these mnemonics, the study of meme proliferation is of little use. Consider what happens when an earthquake of harmless but noticeable intensity strikes Los Angeles. Millions of people suddenly have the idea that an earthquake has struck on that particular day. So at first the host population does not result primarily from mnemonic copying. In fact, the people who directly experience the quake may remain a majority of the idea's hosts, especially if the quake is too mild to get much news coverage. Replicator theory has extremely limited relevance in studying this kind of host population growth. The theory best applies to the limited but still vast realm of memes and meme-derived mnemonics, i.e., the realm of memetics.

The ability to unambiguously identify mnemonics as homoderivative now becomes crucial to memetic theory. For instance, when someone receives the hell mnemonic from one person and the imminent doomsday mnemonic from another, is the resulting  $A*B$  combination homoderivative or heteroderivative? The  $A*B$  host has copied both mnemonics from pre-existing hosts. But the event that actually forms  $A*B$ , namely  $\sim A*B+A*\sim B\sim A*B+A*B$ , does not actually contain  $A*B$  as an input mnemonic. The mnemonic appears not to be distinctly homoderivative or heteroderivative, a problem in deciding how to proceed.

The problem can be resolved by recalling that mnemons propagate only with respect to an abstraction. Although the  $A^*B$  host is homoderivative with respect to abstraction A, and homoderivative with respect to abstraction B, the instance of  $A^*B$  is clearly heteroderivative with respect to abstraction  $A^*B$ . So  $A^*B$  formed by the event  $\sim A^*B+A^*\sim B\sim A^*B+A^*B$  is not a meme, although it is meme-derived.

Yet as mentioned earlier,  $A^*B$  can also propagate as a set by the non-parental conversion event  $A^*B+\sim A^*\sim B=2A^*B$ . Formed this way,  $A^*B$  is a meme. Thus, the host population of  $A^*B$  is yet another mixture of memes and meme-derived mnemons.

Disingenuous propagation raises similar questions. Consider the following mnemons:

Mnemon  $P$  is the *belief* that "bee pollen invigorates".

Mnemon  $Q$  is *awareness* of the "bee pollen invigorates" proposition.

Mnemon  $R$  is the belief that "it is profitable for bee pollen merchants to tell customers that bee pollen invigorates".

If a merchant holds the  $\sim P^*Q^*R$  combination, then a disingenuous propagation event might then be represented as

$$\sim P^*Q^*R+\sim P^*\sim Q^*\sim R\sim P^*Q^*R+\sim P^*Q^*R\sim P^*Q^*R+P^*Q^*R.$$

This event clearly involves replication with respect to the awareness meme  $Q$ , and the profitability meme  $R$ . Yet this event alone is not homogenic with respect to  $P$ , because the merchant is a non-host of  $P$ . Still, there is a mechanism of indirect replication in this case. If the  $P^*Q^*R$  host goes on to pay the merchant, it helps preserve the meme  $R$  in the merchant and also pays helps pay the bills that the merchant needs paid in order to continue instilling  $P^*Q$  in new hosts. This form of recursion is already recognized in such fields as marketing science and consumer education, and hence is not a distinctly memetic insight. Also, the actual profitability of transmitting  $P^*Q$  involves numerous business and marketing science considerations, such as how many people are persuadable, how much they are willing to spend, how much the bee pollen costs to produce and distribute, etc. Understanding the disingenuous spread of  $P^*Q^*R$  is largely covered by older disciplines, so the propagation event here serves mainly to illustrate how it is diagrammed and in what sense it can involve replication of  $P$ . Note that the event  $\sim P^*Q^*R+\sim P^*\sim Q^*\sim R2\sim P^*Q^*R$  may also happen recursively when bee pollen is incorporated into multi-level marketing schemes.

Simple non-commercial lies also qualify as events in which the transmitting party holds an awareness of a proposition but not the belief in that proposition. If the lie is believed by the recipient, it would take the form  $\sim S*T+\sim S^*\sim T\sim S*T+\sim S*T\sim S*T+S*T$ , where  $S$  is the belief in a given statement and  $T$  is mere awareness of it as a proposition. There may be reasons for the  $\sim S*T$  host to repeat the lie many times to many people, and of course the new  $S*T$  hosts may have reasons for honest retransmission as well (by events such as  $S*T+\sim S^*\sim TS*T$ ).

## 10 *Meme*, Concisely Defined

Defining the word *meme* concisely but technically without reference to the other neologisms, we have:

### **MEME**

*A memory item, or portion of an organism's neurally-stored information, identified using the abstraction system of the observer, whose instantiation depended critically on causation by prior instantiation of the same memory item in one or more other organisms' nervous systems.*  
*("Sameness" of memory items is determined with respect to the above-mentioned abstraction system of the observer.)*

The causation of a new instantiation can happen by a great many routes. The role a *meme* plays in causing a new instantiation can seem rather passive, as when parents spend days trying to decide which *meme* they will use for naming a newborn. Or the *meme* can play a strong role in manipulating host's communication behaviors toward retransmission, as when the idea of astrological compatibility motivates believers to retransmit to all potential mates. The latter kind of *meme* qualifies for the more restrictive term *thought contagion*.

Although my *meme* definition refines Dawkins's original informal definition, it still covers a vast range of phenomena. There are, however, those who want the term "*meme*" to refer to all sorts of cultural phenomena that might be regarded as replicators. Yet this can lead to great confusion, especially over how to measure the extent of propagation.

Still, one can study either a chain letter (non-*meme*) or the thought propagated by the chain letter (*a meme*) as replicators. Artifacts and numerous other cultural phenomena might be investigated as instantiations of informational replicators, yet they are not *memes* by the above definition<sup>5</sup>. The above definition does not have *memes* "residing" in anything but an organism's nervous system, though it might be generalized to include artificial minds.

Yet clothing zippers, co-dependencies, chain letters, groups, behaviors, institutions, entire societies, and numerous other social or cultural phenomena have been discussed using various kinds of replicator and vehicle models. The strength of the replicator concept varies across such a wide range of phenomena. Yet Pocklington and Best [1, 14], for instance, have published good evidence of culturally replicating information in the medium of usenet postings. They quantify instantiation in terms of distinct postings rather than copies downloaded, messages read by people, or mnemonics retained by people. The distinctions between postings and human hosts matter, especially because people on the usenet often create new copies of material with which they *disagree* by clicking the "post reply" button in their Internet software. Usenet postings would qualify as *meme relics*, or from Dawkins [4] *meme phenotypes*, or even *text contagions*, but not *memes* under the above definition. They are valuable indirect evidence of memetic replicators, much as sea shells washed ashore are valuable evidence of reproducing organisms in the water.

When viewed as replicators, most artifacts involve brains at some point in the causal pathway to forming new "copies". A computer virus, however, can be seen as an artifactual replicator that

usually spreads from machine to machine (or disk to disk, etc.) without being hosted by a new brain for every replication. The instantiations of the computer virus are not memes, but the algorithm of the virus can be a meme: the prankster who initially wrote the code may find his code being learned and copied by other pranksters. We thus have artifactual and brain-stored replicators.

Similarly, one can investigate the role an enzyme plays in causing new instantiations of "the same" enzyme in one or many cells. The investigation will, of course, always find that nucleic acids play a crucial ("central") role in causing new "copies" of the enzyme — a fact that makes it not scientifically *necessary* to treat ordinary enzymes as replicators. But the enzyme can still be viewed as a very indirect, multistage replicator. Cytochrome C, for instance, helps produce ATP that in turn helps produce new "copies" of cytochrome C. Yet much as one does not say that a protein enzyme is or contains a "gene", so too do I not say that an artifact is or contains a "meme". Still, an artifact can function as a medium of communicating knowledge of how to make new copies of "the same" artifact, and this difference from the role of proteins in biology suggests that there should not be a "central dogma" in memetics. It is a situation that Dawkins [4] notes in his remark that "there may be 'Lamarckian' causal arrows leading from phenotype to replicator"<sup>6</sup>.

A more general term than *meme* might nevertheless help clarify these distinctions and provide a way of discussing other kinds of cultural replicators. Noting that biologists took the word *code* and invented *codon*, we can take the word *replicate* and coin the word *replicon*. The adjectival form is then *repliconic*. The associated discipline name would be *repliconics*. Its central theory would be the generalized evolutionary replicator theory or just generalized replicator theory. The subfield dealing with culture would be *cultural repliconics*. Text contagion theory, computer virology, and memetics would be some of the sub-subfields.

The various preliminary distinctions that go into defining the "meme" may appear to be another example of how memetics departs from analogy to genetics. Yet chemistry also has many distinctions that place nucleic acids in a very small subclass. Molecules can be catalysts or not. They can also be autocatalysts or heterocatalysts. The early oceans may have contained oligonucleotides formed by polymerization of smaller components. Some would have formed by non-autocatalytic polymerization, even in outer space. But if any did behave autocatalytically, they would have proliferated into predominance. They would have been a subclass of molecules generally, much as memes are a special subclass of memory items. The "new replicators" that Dawkins introduced are in an environment with similar non-replicators, much as the earliest biomolecules were.

## 11 Stemming the Tide of Expanded Definitions

Because Dawkins never offered an explicitly formal definition of the word *meme*, a great many writers have taken the liberty of expanding its meaning to include nearly any social or informational phenomena they wished to see treated with seemingly the same clarity and precision found in a molecular genetics lab. Taking this to extremes, some have said that since "the meme" is just propagating information without regard to medium, that even genes are just a special case of memes. One writer even posted to the Internet that memetics was the science of

information in all sectors of the universe. Another sees "proto-memes" in the infinitesimally early stages of the Big Bang. There are even writers who insist that memes can be instantiated only in artifacts, never in brains. The problem is that if the word *meme* means just about whatever a writer wants it to mean, then it means practically nothing at all. The resulting bedlam has led some critics to issue blanket indictments against all of memetics, insisting that memeticists have not clearly defined their subject matter. These are generally people who have not yet read a formal and more restrictive definition.

Still, some expansions of Dawkins's original usage need to be addressed carefully. Among these is the expansion offered by Dennett, in which a "meme" can be instantiated by a wide range of "meme vehicles" other than brain-based or artificial minds. Thus, a "meme" for spoked wagon wheels is instantiated not only in the brain of the carpenter building it, but also in the finished wheel itself. The wheel does, after all, contain information that a new carpenter can use to copy the wheel. This departs from Dawkins's usage, in which the meme resides only inside the brain and an artifact such as a wheel would presumably be part of its extended phenotype.

Dennett does, however, correctly recognize the wheel as an artifactual medium of communication. This can seem synonymous with the phrase "vehicle of communication", a connection that tempts us to consider the wheel as a meme vehicle and a locus of meme instantiation. Dennett recognizes moreover that communication by artifact is not always intended, as when the enemies of those traveling by spoked wheel copied the visible and mobile artifact for their own benefit. The actual communication event generally involves more than just a stray wheel resting in a field, but includes a human being demonstrating its usefulness by riding around with two or more wheels, an axle, a horse, and the rest of the wagon.

Expanding the term "meme" in Dennett's manner has a further appeal, in that it seems to offer a continuous chain of material substrates from one instantiation to the next. Yet when a carpenter builds a wheel from memory, there is actually no instantiating medium in between the brain and the finished wheel. So any worry about instantiation gaps applies for both definitions. That worry is unfounded anyway, because an instantiation gap does not imply any gap in physical causation. An instantiation gap exists between one autocatalyst molecule and the next, too, as various intermediates are formed and the initial molecule always has at least a tiny physical separation from the new copy. Likewise for a spreading computer virus, where the cable connecting one computer to the next is often too short (and transfer rate too slow) to contain an entire copy of the viral software: a long physical gap may exist between one instantiation of the virus and the next. Yet Dennett's examples show that useful insights are gained by treating artifactual information as replicators, and it is for just such cases that I offer the term *cultural repliconics*.

A more serious problem with allowing for all sorts of "meme vehicles" as meme instantiation loci is that it tends to confuse the measurement of propagation. Even when the "vehicles" are discrete, there are questions of whether they should all be enumerated equally. Copies of an old advertising logo may still exist in landfills a century after the product disappeared. The fact that the human host population of a meme is zero certainly tempts us to call the meme extinct, even if it can be brought back from extinction by an inter-century communication event to some future excavator. And if supernova radiation from a few light-years away leaves an earth full of unused

artifacts until the Sun's demise, we would tend to say that the memes went extinct with the end of humanity rather than with the end of the planet.

The problems of enumerating the instantiations become even murkier with non-discrete "meme vehicles". How many copies of the distress signal from the Titanic now exist at a radius of 86 light-years? What about a still local message crammed into a 300 kilometer thick (1 millisecond) burst of microwaves from a satellite? In such cases we tend not to enumerate the theoretically feasible number of receptions of the message as "memes", but instead only those which actually *are* decoded and stored in minds.

The consideration of replicating cultural phenomena not instantiated by any single mind has led to interesting and worthwhile insights. Yet sayings such as "a scholar is just a library's way of making another library"<sup>7</sup> bring on questions of whether even Dennett is trying to pack too many possible meanings into the one word "meme". Scientists reading these statements often get the impression that while "cute", it is not specific enough to be part of a real science. Not surprisingly, Dennett himself, after venturing down this road filled with every variety of "meme vehicle", ends up questioning whether there can be a science of memetics at all and answering in a doubtful voice<sup>8</sup>. The challenges of developing a unified replicator theory for so many kinds of vehicles are perhaps big enough to become terrifying. Better, in my view, to consider the scientific merits of a more restrictively defined "memetics" and take up the question of other classes of cultural replicators separately.

The substrate-specific definition of "meme" stands in some contrast to the apparent substrate neutrality of the term "software". Yet this is actually a result of the many kinds of artifacts that serve as computer "memory". The term "software" does not really refer to information in *any* medium. Suppose a robot using Cartesian coordinates builds a screw. A different model of robot using polar coordinates analyzes the screw and builds copies using completely different measuring units, and digitizing in finer detail. We do not say that the screw contains the "software" for making screws, or any kind of "data" or "software", even though it does contain replicating information. So the word "software" is actually medium-specific, but only with respect to an ever-expanding *class* of information media. The screw can, however, be viewed as a *software relic* much as the spoked wheel can be viewed as a meme relic. Both can also be viewed as instantiations of generalized (non-meme) information replicators. In this latter perspective, a robot is the screw's way of making another screw and a carpenter is the wheel's way of making a new wheel. It is a repliconic perspective that has some of the flavor of memetics even though memes as I define them are not the direct topic of discussion.

In biology, the term "gene" has progressed toward more specific and technical definitions. Scientists have not attempted to expand its meaning to include the information stored in all sorts of vehicles, even though many types of vehicles contain biological information. Even other replicating entities such as prions are not considered to be or contain instantiations of genes. And we certainly do not see anyone attempting to expand the term "gene" to include information in all sectors of the universe. In order to be taken seriously in science, the term "meme" needs to be refined toward greater, not lesser, specificity than its original usage. The definition I offer is more formal and specific than what Dawkins presented in 1976 [3], yet it still covers the vast territory that he intended.

## 12 Meme Sizes

Some efforts have been made to incorporate the idea of "size" into the definition of the word *meme*. This includes those who favor the "smallest" possible units, as well as those who favor the "largest" possible units. Unlike those who expand the definition of *meme*, these scientists are indeed looking for ways to render the meaning more specific. Yet not all attempts at specificity are equally useful in the study of real phenomena.

Size is a fairly easy thing to measure for segments of DNA or their corresponding proteins, as well as for pieces of text. However, the concept of "size" becomes troublesome with beliefs and other information stored in the brain: general methods of "size" measurement are not currently available, and even if they were, they might register different "sizes" for "the same" belief in different brains. The most we can really say is that a set of memes is, in general, a partially ordered set (as defined in set theory). One rather arbitrary system of partial ordering is based on the size of conjunctions of the mnemonics identified by a particular abstraction system. Take the mnemonics expressed by the following 3 statements, for instance:

Mnemon A: "There is only one true God".

Mnemon B: "Christ is Lord".

Mnemon C: "Unbelievers are damned".

Mnemon D: "Earthly life is better among believers".

We have no basis for saying if  $A > B$ ,  $A < C$ , etc.

We can say  $A * B > A$ , and  $A * B > B$ .

But you cannot say  $A * B > C$  or  $A * B < B * C$ , etc.

We can, however, say  $A * B * C > A * B$ ,  $A * B * C > A * C$ , etc.

In other words, if the hypothetical faith only says "There is only one true God", and "Christ is Lord", we cannot, for instance, say that it is "bigger" (has more "size") than the faith that only says "Christ is Lord" and "Unbelievers are damned". Nor can we compare the "sizes" of these component beliefs. Moreover, completely unordered sets such as  $S = \{A * B * C, A * B * D, A * C * D, B * C * D\}$  demonstrate clearly that size cannot be a universal criterion in defining which mnemonics are memes.

Mnemon size can, however, become an empirical consideration when studying the special case of a fully ordered set of mnemonics, such as  $S' = \{A, A * B, A * B * C, A * B * C * D\}$ . We might find that element A does not induce much replication. Then we might find that  $A * B$  induces more replication, but that  $A * B * C$  induces still more. Yet we may also find that  $A * B * C * D$  achieves less replication, perhaps by requiring too much information transfer to happen reliably. One can still decide based on this to choose  $A * B * C$  as the main subject of investigation without requiring size to be a criterion in the definition of all memes. Since A may have propagated long before the first occurrence of  $A * B$ , and  $A * B$  long before  $A * B * C$ , it is important to be able to consider all the elements of set  $S'$  as memes.

## 13 Massively Cooperative Propagation

Though religious meme sets such as those above spread by inducing a wide variety of one-to-one transmission events, some can also induce events involving large numbers of participants. In multistage events, there can also be large numbers of participants in some stages and few in others. Thus, (retaining the symbol meanings of the preceding paragraph) an individual might spread  $A*B*C$  through the event  $A*B*C*D+\sim A*\sim B*\sim C*\sim DA*B*C*D+A*B*C*\sim D$ . But then it may take the 1000 other believers in a community to achieve the next event:

$$1000A*B*C*D+A*B*C*\sim D1001A*B*C*D.$$

Now consider what happens when adding a few more memes:

Meme E: "Love your neighbor as yourself"

Meme F: "The Spirit lives in the community of believers"

Meme G: "Church X is God's Church".

As I noted in *Thought Contagion* [12], memes which cause greater cooperation among their own hosts can thereby achieve propagation advantages by raising prosperity and suppressing the sort of personal conflicts that may lead to dropouts. The "love your neighbor" meme  $E$  even motivates more one-on-one evangelism in some settings while promoting collective meme propagation in other settings. The meme expands the range of unbelievers that adherents wish to save from damnation: instead of just wanting to "save" a few friends and relatives, they want to "save" even strangers. Yet the meme also works in many ways to improve *community* life among believers, an effect I discussed in *Thought Contagion* [12] and Stark documents for early Christians in *The Rise of Christianity* [16]. So when an individual Christian attempts to spread memes  $A$  through  $G$  to a pagan, the pagan may well remain unconvinced. Yet the pagan realizes that meme  $D$  is testable merely by spending some time among believers, and without waiting for an afterlife. In ancient times, they would have found a Christian community thriving from cooperation, and taking better care of each other's misfortunes than did the pagans. That would have impressed them that at least the meme  $D$  part of the Christian message was true. This in turn would have made it cognitively easier to believe the "Spirit" meme  $F$ , and the "Spirit" meme  $F$  would make it easier to believe the "God's Church" meme  $G$ . Finally, that would make it easier to believe other Christian memes such as memes  $A$ ,  $B$ , and  $C$ .

In addition to these cooperative conversion effects, a community can also achieve greater rates of preventing the dropout event  $A*B*C*D*E*F*G\sim(A*B*C*D*E*F*G)$  and the mortality event  $A*B*C*D*E*F*G0(A*B*C*D*E*F*G)$ . The community role in dropout prevention would result in part from conformity pressure, and in part providing social, emotional, material, and health care benefits. It would augment such dropout prevention effects as the threat of hellfire discussed by Dawkins [3], and the efforts of individuals to shore up the faith of anyone expressing renewed doubts. The mortality prevention effect, resulting from such things as the community feeding the poor and nursing the sick, is documented by Rodney Stark in *The Rise of Christianity* [16].

For some memes, the effect on rates of cooperative propagation is particularly striking. With Hutterites (denoted here with meme symbol  $H$ ), for instance, we observe many events of the form  $2H \rightarrow 2H+~H$  (Two Hutterites have a baby non-Hutterite), followed by an event such as  $20H+~H \rightarrow 21H$  (Twenty Hutterite adults impart the faith to the child). Then we see events such as  $165H \rightarrow 83H+82H$  (One Hutterite colony of 165 splits into two colonies of 83 and 82)<sup>9</sup>. Crucial to setting the rate of this event, however, are the family structure memes that raise childbirth rates to 10 per couple while keeping the (permanent) parent to child inculcation rate above 90 percent. A distinctly Hutterite meme distributing the chores and costs of child-raising among the whole colony's adults probably plays a strong role here, as it dilutes the usual pragmatic motives for regulating fertility. The idea of splitting the colony when it grows to a certain size range keeps colonies always in a phase where they need a constant source of new young hands to build and run the collectives. Since they do not recruit outsiders, it effectively augments the pressure to keep having the numerous children.

The phenomenon of large numbers of humans propagating a movement by acting in concert has been cited as evidence for the group selection of human genes. Wilson and Sober [17], for instance, take the colony behavior of Hutterites as evidence of innate hive-like social dispositions arising from group-selected genes. Yet the ideological imperatives discussed above can explain the evolution of a Hutterite meme set in terms of memetic selection alone – without requiring any new innate factors beyond a very general capacity to learn and comply with cultural mores<sup>10</sup>. Likewise, the Christian propagation advantages that arise through community behavior do not suggest a special genetic mechanism. As important as they are, massively cooperative meme replication events remain a subclass of memetic events in general. The fact that they happen for a variety of memes in various populations suggests that communities can achieve certain meme transmission effects beyond the capacities of single individuals. Yet it does not of itself show that the memes must be relying on an innate imperative for highly collective action.

## 14 Centralized Communication

Centralized meme transmission stands in contrast to massively cooperative propagation. Here, a very few people act to spread a meme to vast numbers by events such as  $A+1000000\sim A$   $10001A+990000\sim A$ . (Host of  $A$  communicates to 1000000 non-hosts of  $A$ , yielding 10000 new hosts of  $A$  plus 990000 continued non-hosts of  $A$ .)

Mass media often play an important role in such replication events. When they do, most of the recipients cannot go on to play the role of meme retransmitter by the same kind of centralized event. This may at times make centralized communication less suitable for study with *recursive* replicator theory. It can also at times limit the role of distributed replication: much as humans can acquire legionella bacteria centrally from central air conditioning systems, so too can we acquire memes by way of intentionally central meme spreading systems.

Another feature of centralized communication is that not all of the people transmitting the message are actually hosts of the meme. The idea that "Coke is the real thing", for instance, has probably been broadcast by many celebrities who prefer to drink something entirely different in the privacy of their homes. This again limits the usefulness of the recursion principle, as sender and receiver do not always have "the same" meme. The sender, for instance, may only have the

idea that it is *profitable* to say "Coke is the real thing" rather than actually *believing* in the message. Yet as with non-centralized profitable communications, people who adopt the meme often spend money that pays merchants and advertisers to send the message again and again. This is a phenomenon recognized and applied long before the advent of memetics, and already studied in marketing science.

Mature recipients of centralized messages often understand that the sender may not believe what she says, thus limiting receptivity to the message. Communications from friends and family generally do not have an apparent profit motive, too, and can thus enjoy a more privileged reception. Desires to maintain harmonious relationships also cause people to pay closer attention to messages from direct acquaintances. This helps maintain the importance of non-centralized communications from immediate family and acquaintances in the age of mass media. In turn, it prevents all of memetics from being reducible to existing work in media studies or marketing science.

The continued importance of distributed communications even manifests itself in the recursive propagation of memes specifying which centralized sources to consume. The Bible, for instance, may look like a centralized communication, coming from mass production plants ever since moveable type was invented. But the idea of buying and reading it propagates person to person in the population, along with various memes specifying what passages to emphasize and what they mean. Another example that I discussed in *Thought Contagion* [12] was the localized propagation memes for listening to centralized political talk-radio shows. Listeners to shows with pro-business, anti-tax messages play the radio openly in the workplace, allowing co-workers to learn about the show and its message. But listeners to shows that might politically offend the management feel inclined to use headphones, which blocks the local retransmission of their political and show-preference memes. Even that most centralized of modern media, television, exhibits natural selection processes, as when the remote control caused a trend toward greater violence, sex, and melodrama — the types of content that happen to stop the passing "channel surfer"<sup>11</sup>. Far from replacing distributed memetic evolution, the centralized media have instead opened up new realms for recursive propagation and natural selection in competing meme bundles.

## 15 The Fundamental Role of Abstraction in Science

The great potential for confusion in discussing cultural phenomena forces explicit reference to systems of abstraction. Yet the axiom of abstraction, fundamental to all of science, is generally invoked on a tacit level in most disciplines. Indeed, the axiom of abstraction is also known as the axiom of comprehension. It merely states that for any attribute, one can define a set of all the entities that have that attribute. Saying that two entities are "the same", then, is tacit shorthand for saying that they are elements of the same set, whose inclusion criterion is the attribute. The phrase "the same with respect to an abstraction" serves to remind readers of the underlying mental operations involved in calling any two entities "the same".

In the physical sciences, abstraction manifests itself every time someone refers to an entity as having the attribute of being a "water molecule". "Water molecule", is an abstraction which scientists have found useful for discussing phenomena. What they really mean by "water

"molecule" is a particular *pattern* of matter and energy. It is a way of calling two or more patterns of matter and energy "the same", ignoring such differences as location, velocity, rotational states, vibrational states, oxygen 18, Hydrogen 2, nuclear spin states, electron spin states, electron excitation states, quark states, etc.

Is it ever legitimate to identify *patterns* amid all this seething matter and energy? A philosophy of science question. Yet the grouping of vast categories of objects into sets defined by attributes seems to be a fundamental feature of human thought. Without formally justifying the practice, one can note that this can be an efficient way of storing vast amounts of knowledge in a finite brain. One can also theorize that evolution has favored brains that do this because it provides sufficiently functional representations of reality to give an animal a survival and reproductive advantage. But these arguments themselves depend on abstractions.

Regardless of why we rely so much on abstractions, science as we know it is largely a project to develop and test ever stronger systems of abstractions with which to describe and comprehend reality. The "stronger" systems of abstractions are defined as those which explain more of reality with reliance on fewer essential abstractions. Yet deep questions remain on whether nature will someday indulge the physicist's quest for the kinds of "ultimate" abstractions of "grand unification theories". Will our desire for ever more unifying abstractions instead be frustrated by some fundamental complexities of a reality not designed specifically to please the human mind? This we have yet to learn.

Despite advances in fundamental knowledge, we cannot run through the equations of quantum mechanics to explain something as complicated as a cell or a DNA molecule. So we invoke further layers of abstractions for these vastly complicated systems. Instead of representing "adenosine" in terms of a chemical formula, a single letter is used. And "little things" like the dissociation of H<sup>+</sup> from phosphate groups and particular serpentine shape traced out by the molecule are ignored. Compared to the level of elementary particles, molecular biologists focus on patterns of patterns of patterns, layers of abstraction removed from what physicists call "fundamental".

A great deal of science went into demonstrating that "genetic material" is a useful abstraction, and that "nucleic acid" can be used as an equivalent abstraction here on Earth. Likewise for the discovery that "nucleotides" are useful abstractions for describing nucleic acids. Yet the pre-molecular concept of "gene" was a useful abstraction, and its very utility provided impetus to the research that lead to more powerful molecular abstractions. More recent arguments over different molecular definitions of "the gene" amount to arguments about which abstractions are stronger. Yet the "the gene" started as an abstraction based on studies of entire organisms, and has progressed into an abstraction based on studies of molecules.

Materials can also be *manipulated* to fit a set of desired abstractions, as has been done with digital electronics. One voltage interval is chosen to represent the abstraction "1", and another the abstraction "0", with all sorts of "low level noise" to be ignored. Analog vacuum tubes, transistors, and other devices are then wired to produce abstractly bimodal outputs corresponding to abstract operations on abstractly bimodal inputs. Hence the "NOR gate", for instance. Still higher levels of abstraction arise when referring to "software", where specific voltages and even

patterns of "1's" and "0's" can be ignored in deciding whether two "programs" are "the same". Science again builds layers of abstractions, this time working from the low level up rather than the high level down.

Because "replication" depends on "sameness", and the "sameness" of any two "water molecules", "DNA molecules", etc. exists only with respect to an abstraction, it follows that replication only happens with respect to an abstraction. Abstraction systems and sameness criteria are important in all evolutionary replicator theories. Yet "replication" also involves the notion of causation. Specifically, the word "replicator", can be replaced with the cumbersome phrase "instantiation of an abstraction that causes a new instantiation of the same abstraction". This is the tacit meaning in use for all kinds of evolutionary replicator theories ranging from the "artificial life" to memes to autocatalytic molecules that might evolve in the oceans of Europa.

Memes are abstractions currently based on macroscopic observation. Whether it will stay at this level remains to be seen. There is, however, a school of thought that seems to insist that memes must be "detected" by some kind of molecular or microscopic means in order to be valid abstractions. Yet the same people may routinely treat "Nazism" or "monotheism" as "real" phenomena without providing any molecular or cellular basis. Many even go on to attempt to get others to hold "the same" idea as theirs about memetics. In doing so, they unwittingly affirm the underlying premise of memetics, that one person can cause another person to have "the same" idea by way of "communication". They need only admit further that different ideas can spread recursively at different rates before seeing the evolutionary arguments of memetics follow mathematically. There too, many of them affirm a principle of memetics by avidly trying to make their own beliefs spread at a greater rate than competing beliefs. This is not the same thing as accepting specific memetic hypotheses, but it does show tacit acceptance for the underlying tenets of memetics.

Nevertheless, some of the discomfort that many voice over memetics arises almost subconsciously from the fact that science has not discovered an absolute, fundamental, or privileged system of abstractions with which to discuss socially transmitted information. To worsen the situation, some popular writers seem convinced that they have actually discovered a set of fundamental "units" of mass culture, that is, a strongly preferred or "privileged" system of abstractions resembling the nucleotides and genes of biology. But the "sameness" of any two people's learned information about a topic depends on the abstraction system of the observer. With genes, nature has rendered a few very similar abstraction systems far more powerful than the alternatives, so that scientists do not notice the role of the observer's abstraction systems. But there are many ways to define "anti-abortionism", for instance — which is another way of saying that the abstraction system and survey questions can be constructed in many different ways. In reality, memeticists will probably have to abandon the "fundamental unit" idea much as physicists have now abandoned the idea of an absolute coordinate system (framework of location and time abstractions) for the universe. Not only are meters and seconds arbitrary units of measure, but measurement results depend on the framework in which measurement is taken. More fundamentally, just as physicists now accept that the "simultaneity of events" exists only with respect to an observer's coordinate system, so too must memeticists accept that the "sameness of ideas" exists only with respect to an observer's abstraction system. While this is not a "special theory of cultural relativity", it does show scientific precedent for admitting that

multiple systems of abstraction can apply to a given class of phenomena. It does not mean that all abstraction systems are equally useful, but it does suggest that science can proceed without the assumption that there must be one abstraction system that is universally right for memetics.

## 16 Population Memetics

Evolution is fundamentally a quantitative as well as a qualitative theory. The abstractions that allow us to discuss memes and memetic events set the stage for discussing quantitative *event rates*. Differences between memes in the rates of replication and termination events are what give rise to the natural selection of memes. They are hence essential to the evolution by natural selection in memes.. From the fact that event rate parameters differ between memes, it follows *mathematically* that natural selection must happen. Yet this statement can be rendered much more specific. The mathematically *necessary* consequences of differing event rate parameters can be *characterized*, using systems of differential equations for meme host population versus time. Defining the units and quantifying the events thus allows us to analyze the host population *dynamics* of memetic evolution.

The particular terms needed for a system of equations to model the population dynamics of a specific set of memes will depend upon what kinds of events occur at non-negligible rates for the memes being modeled. A vast range of event forms can be conceived, but the rarest of them do not need to be modeled in order to gain acceptable accuracy — much as the rarest side reactions in a chemical process can be ignored in quantitative chemistry. Models will in general have to account for at least two population groups, namely, those for hosts and non-hosts of a meme. Accounting for additional population groups is accomplished by expansion to a system of more than two equations. Most models will also have account for some very common types of memetic events, such as simple parent to child transmission, non-offspring conversion, spontaneous dropout, and mortality. Two-equation systems to model these basic kinds of events therefore constitute a starting point for mathematical memetics, allowing further elaborations to be added as required.

The following two differential equations pertain to two memes whose host populations are represented by  $N_1(a,t)$  or  $N_2(a,t)$  — population age profiles. All members of the total population  $N(t)$  are assumed to be counted somewhere in either  $N_1(a,t)$  or  $N_2(a,t)$ , indicating that the two memes are complements of each other. (In actual practice, one would often want to divide the population into more subgroups, such as the host populations of an idea, its opposite, and hosts of neither. Some subgroups might correspond non-replicated mnemons rather than memes. The present discussion is limited to two groups in order to illustrate quantitative methods as simply as possible.)

The other parameters in the equations have the following meanings:  $t$  is time in years,  $a$  is host age,  $p$  is the age of a second person — the idea propagator — used in places where two people's ages are involved.

$R_1(a)$  is the fertility rate for meme-1, in children per host of age  $a$  per year — the quantity parental parameter.  $K_{11}(p,a)$  is the fraction per year of children of age  $a$  who learn meme-1 from an age  $p$  parent who hosts meme-1—the efficiency parental rate.  $K_{21}(p,a)$  is the fraction per year of

children of age  $a$  who learn meme-1 from an age  $p$  parent who hosts meme-2, a kind of parental "failure rate" for meme-2. Again, for the sake of "simplicity", the different  $R$  and  $K$  values that may occur when one's parents come from different host populations are not modeled here. (More than just  $R$  and  $K$  parameters are involved, since the occurrence rates of "mixed" versus "unmixed" couples change with changing host populations.)

$\gamma_{12}(p,a)$  is the average annual net number of non-parental converts a meme-1 host of age  $p$  makes per unit meme-2 host population-age density at age  $a$  in his society.  $\gamma_{12}(p,a)$  is the average annual net number of non-parental converts a meme-1 host of age  $p$  makes per *percentage-year* of meme-2 hosts of age  $a$  in his/her society. Non-parental conversion rates represented by  $\gamma_{12}(p,a)$  are sensitive to how crowded the society is as a whole while  $\gamma_{12}(p,a)$  rates per meme-1 host are purely sensitive to the *fraction* of meme-2 hosts in the society. The latter reflects the non-parental conversion between, for instance, spouses: people do not generally double the number of spouses as the population doubles. Yet the number of people one encounters on the street might well double as the population doubles. If so, then non-parental conversations on street corners would be modeled using  $\gamma_{12}(p,a)$ .

$\gamma_1$  is the fraction per year of meme-1 hosts who convert to meme-2 without any prior meme-2 hosts teaching them, or who at least convert by mechanisms that do not depend on the size of the meme-2 host population<sup>12</sup>. Preventing such "dropouts" is one form of preservational advantage (information longevity advantage) for meme-1. Finally,  $M_1(a)$  is the rate of mortality per age  $a$  meme-1 host per year.

Swapping "2" subscripts for "1" subscripts in the above sentences gives the parameter definitions corresponding to changes in meme-2 prevalence<sup>13</sup>.

$$\begin{aligned}
 \frac{dN_1(a,t)}{dt} = & \int_a^{\infty} R_1(p-a)K_{11}(p,a)N_1(p,t)dp \\
 & + \int_a^{\infty} R_2(p-a)K_{21}(p,a)N_2(p,t)dp \\
 & + \int_0^{\infty} \gamma_{12}(p,a)N_1(p,t)N_2(a,t)dp \\
 & + \int_0^{\infty} \frac{\beta_{12}(p,a)N_1(p,t)N_2(a,t)}{N(t)} dp \\
 & - \alpha_1 N_1(a,t) + \alpha_2 N_2(a,t) \\
 (1) \quad & + \frac{\partial}{\partial a} N_1(a,t) - M_1(a)N_1(a,t)
 \end{aligned}$$

$$\begin{aligned}
\frac{dN_2(a,t)}{dt} = & \int_a^\infty R_2(p-a)K_{22}(p,a)N_2(p,t)dp \\
& + \int_a^\infty R_1(p-a)K_{12}(p,a)N_1(p,t)dp \\
& + \int_0^\infty \gamma_{21}(p,a)N_2(p,t)N_1(a,t)dp \\
& + \int_0^\infty \frac{\beta_{21}(p,a)N_2(p,t)N_1(a,t)}{N(t)} dp \\
& - \alpha_2 N_2(a,t) + \alpha_1 N_1(a,t) \\
& + \frac{\partial}{\partial a} N_2(a,t) - M_2(a)N_2(a,t)
\end{aligned} \tag{2}$$

The first two terms in equation 1 are the parental terms. In the first term, the group of  $N_1(p,t)$  parents of age  $p$  is multiplied by the average number  $R_1(p-a)$  of children per adult that they had  $a$  years ago to get the number of children of age  $a$  having parents of age  $p$ . This then is multiplied by  $K_{11}(p,a)$ , the fraction per year of children in this latter group having meme-1 passed down to them (while they are of age  $a$  and the teaching parent is of age  $p$ ). This is then integrated over the entire range of parents' ages ( $p=a$  to  $p=\infty$ ) to get the total rate at which meme-1 hosts are passing the meme down to children of age  $a$ . The second term gives the rate at which the meme-2 host population parentally produces meme-1 hosts of age  $a$ .

In practice, there may be many cases where a useful mathematical model can be attained by treating the parent to child meme transmission as if it all happened when the children reached the single age  $c_1$ , the *average* age at which they pick up meme-1 from a parent. Also, one might find that the transmission rate per child depends very little on parent age differences within the mainstream host child raising years. If this is true, then one can replace the  $K_{11}(p,a)$  function with the very simple function  $k_{11}(a-c_1)$ , where denotes the delta function, and  $k_{11}$  is simply the overall fraction of children who acquire meme-1 from their parents a much easier thing to measure than transmission versus age. The remaining functions in the parental terms are just fertility versus age and the population age profile the sort of data that demographers and census-takers have already measured for some groups.

The next two terms in the equation concern the non-parental conversion mode of transmission. The first of these is the one that is sensitive to the total number of potential converts,  $N_2(a,t)$ , rather than merely their proportion to the total population. If both  $N_1(a,t)$  and  $N_2(a,t)$  doubled, each individual meme-1 host would be winning twice as many converts and the meme-1 host population (at double size) would be winning four times as many converts per year. On the other

hand, the second term on line two would only double, and not quadruple, if each group doubled. In reality, the dependence of non-parental conversation rates on host population sizes is more complicated than the two terms suggest, and the nature of the dependency would need to be studied empirically as part of any detailed mathematical modeling effort for memes with significant non-parental transmission. As with the parental terms, the non-parental conversion terms may be simplified in some cases by replacing the age dependencies with "lumped" effective propagations at certain effective ages.

The following two terms (beginning of line 3) express the "spontaneous" dropout rate for meme-1 and meme-2, respectively. "Spontaneous" dropout rates are assumed to be proportional simply to the number of hosts capable of dropping out. Such rates could also be modeled as age-dependent rates by replacing  $\gamma_1$  with  $\gamma_1(a)$  and  $\gamma_2$  with  $\gamma_2(a)$ .

The next term is the partial derivative of  $N_1(a, t)$  with respect to  $a$ . This term indicates that part of the changing population age profile of meme-1 is due to simple aging of its host population. The final term expresses the mortality rate as a function of age among meme-1 hosts. Mortality per host per year at age  $a$  (a kind of actuarial data) is simply multiplied by the number of hosts at age  $a$  to give the overall rate.

Equation 2 above models the same kinds of propagation processes for meme-2 as are modeled for meme-1. Equation 1 and equation 2 form a system of non-linear differential equations modeling the interdependent propagation of meme-1 and meme-2.

What follows are five more equations that go with equation 1 and equation 2, defining relationships between the propagation parameters. Equation 3 states that all offspring of meme-1 parents end up holding either meme-1 or meme-2. Equation 4 says the same thing for meme-2 parents. Equation 5 and equation 6 state that one group's non-parental conversion gains are the other group's non-parental conversion losses, so that the net non-parental gain to the whole population is 0. The last equation ( equation 7) merely defines the function  $N(t)$ , the total population versus time, as the sum of the as the sum of the two meme host populations, all ages included.

$$(3) \int_{\tau=0}^{\tau=a} [K_{11}(p, \tau) + K_{12}(p, \tau)] d\tau = 1, \quad a > 0$$

$$(4) \int_{\tau=0}^{\tau=a} [K_{22}(p, \tau) + K_{21}(p, \tau)] d\tau = 1, \quad a > 0$$

$$(5) \gamma_{12}(p, a) + \gamma_{21}(a, p) = 0 \quad \text{for all } a, p$$

$$(6) \beta_{12}(p, a) + \beta_{21}(a, p) = 0 \quad \text{for all } a, p$$

$$(7) N(t) = \int_0^{\infty} [N_1(a, t) + N_2(a, t)] da$$

Not all meme propagation events are given their own separate terms in equation 1 and equation 2. For instance, if it frequently happened that meme-1 hosts produced meme-2 offspring who then converted their parents to meme-2, then the rate at which meme-1 hosts were non-parentally converted to meme-2 might depend greatly on how many children they had. A new term might have to be added to the equations to make this phenomenon adequately modeled. The model can, in fact, be made arbitrarily complex, but it is obviously desirable to keep it as simple as the application permits.

Another elaboration of the equations, and one that might interest mathematical sociobiologists, is the explicit inclusion of specific genes and their propagation parameters into the picture. Such a model would analyze "host populations" of memes, genes, and meme-gene combinations all in the same system of equations. Of course, one sets the  $K$ 's,  $\alpha$ 's, and  $\beta$ 's to zero wherever the propagation of a gene is modeled. The parental terms become more numerous and take on a diversity of forms corresponding to all the combinations of genes and memes that can occur in two mating adults. Keeping track of the homozygous and heterozygous combinations would also cause more population parameters and equations. Each possible outcome of each parent combination must have its own term in one of the equations. Such equations embody no *a priori* assertions about the relative importance of either genetic change or cultural change over a modeled time span, nor any assertions of how strongly or weakly prior genetic and cultural evolution constrains the course of change over that time span. Instead, they allow for considering these matters on a case by case basis once the requisite starting data are fed into the models. Although emotional and cognitive receptivity factors are not readily conspicuous in equation 1 and equation 2, they are in fact represented. The reason is the  $K$ 's,  $\alpha$ 's, and  $\beta$ 's are measures of *successful* meme transfer events. As such, they are composites of both the rates at which propagation is attempted and the rates at which it is cognitively and emotionally well-received. Likewise, the "spontaneous" dropout rates ( $\gamma$ 's) include their own products of cognition and emotion.

Anti-competitor propagation advantage, on the other hand, is not fully represented by the  $K$ 's,  $\alpha$ 's, and  $\beta$ 's. Part of the reason is that this mode can occur in quite a wide range of ways. It makes a big difference, for instance, whether the meme-1 group merely bans meme-2 proselytizing or launches a meme-2 extermination campaign. Moreover, the effectiveness of such measures does not vary as a simple function of the meme host populations. The Nazis, for instance, became dramatically more harmful to competitors after they became numerous enough to gain political power. Such phenomena may well defy mathematical modeling techniques based solely on predicting host populations versus time, and require the detailed modeling of political processes as well.

The  $K$ 's,  $\alpha$ 's, and  $\beta$ 's are each modeled as overall *effective* rates of meme transmission. The  $K$ 's, for instance, do not indicate how many times a parent needs to repeat a message to her children before it is effectively learned. The  $\alpha$ 's, and  $\beta$ 's likewise do not reflect how many a message was voiced from hosts to a non-host before that non-host converted. A more detailed model might therefore break down these parameters into the subfactors of *transmissivity*, a measure of how often each host attempts to transmit a meme, and *receptivity*, a measure of the likelihood

each host to non-host transmission attempt has of actually imparting the meme to a new person. Much research has been done on how various components of receptivity affect the diffusion rates of innovations [15]. Receptivity parameters can also be broken down to reflect different probabilities of meme acceptance on first, second, third, etc. exposure. Modeling meme-based *differences* in receptivity to meme-1 or meme-2 that result from the presence or absence of third or fourth (etc.) memes requires expansion to more than two main equations, but may be necessary in cases where the empirical evidence suggests that non-host populations have both "susceptible" and "non-susceptible" subsectors. Such elaborations generate more complicated mathematics.

All meme propagation involves some level of transmissivity, receptivity, and longevity of the memes. Memetic evolution arises from differences in transmissivity, receptivity, and longevity<sup>14</sup>. One of the differences between memetics and "classical" social science is that memetics looks at all three of these general factors, while the classical social sciences often neglect transmissivity and longevity in favor of receptivity. Though pre-memetic social sciences do acknowledge transmissivity differences in mass media, questions of why a whole society came to believe something are often posed as questions of "why did the people want to believe". Investigating such questions does provide useful information, but not a complete picture. Equation 1 and equation 2 model fairly ideal cases of the two-idea propagation problem, but also serve as the kernel of a multi-equation system for cases involving three or more memes with host populations  $N_1, N_2, N_3$ , etc. The equations offer a sample of the kinds of terms that can arise in realistic applications, and they illustrate that a unified quantitative analysis can be given to qualitatively dissimilar modes of propagation. They also illustrate the concept that once the main mechanisms of an idea's propagation have been empirically discovered and measured, the degree to which each mechanism contributes to instantaneous rates of propagation can be mathematically modeled. The resulting systems of differential equations govern host populations as a function of time, and so can be used to generate limited predictions of what will happen if the equation parameters remain reasonably constant in non-chaotic intervals.

The model presented above treats each meme as being either instantiated or not instantiated in a given person. An alternative is to model ideas as a continuum, in which there are an infinite number of mutually exclusive degrees of "hostship" that are instantiated or not. In other words, each individual is considered to occupy a single point on an ideological continuum. Strictly speaking, it would no longer model "replication". Instead, each point on the continuum would exhibit its own birth and mortality rate. Parents at each degree of the trait would have a distribution of rates of imparting all possible degrees of the trait to offspring. People communicating with non-offspring would have distributions of second-party transfer rates denoting how often they move people from one level of the trait to any other level of the trait. Yet in measuring such continuum traits, investigators generally end up assigning subjects to a finite number of trait levels such as "strongly agree", "agree", "neutral", "disagree", and "strongly disagree", designated here as mnemons  $A, B, C, D$ , and  $E$ , respectively. This returns us to dividing the total population into five (or alternately, seven, ten, etc.) mnemonic host populations with a corresponding number of equations. Still, horizontal communication can involve heterogenic events in which, for instance, someone in the "strongly agree" category gets someone in the "disagree" category to transfer to the "neutral" category. As noted earlier, not all memory items are memes with respect to the abstraction system in use, and an abstraction system

using finitely stratified categories would include both memetic and non-memetic (heterogenic) events. However, certain unidirectional multistage heterogenic events can add up to form homogenic (memetic) events with multiple participants. If, for instance, four distinct  $E$ -hosts convert an  $A$ -host into an  $E$ -host with in stages (a separate  $E$ -host participating in each stage), then the heterogenic stages still add up to a homogenic event. The first three events taken individually are heterogenic:  $A+EB+E$ ,  $B+EC+E$ , and  $C+ED+E$ . Only the fourth event  $D+E2E$  is homogenic by itself. Yet the overall sequence adds up to the net event  $A+4E5E$ , which is homogenic with respect to  $E$ , and hence potentially recursive. Other events, including ones going in the opposite direction (e.g.,  $E+4A5A$ ), are also potentially recursive. The potential for recursively shifting large segments of population to higher levels of agreement or disagreement would be inherent in the system of five equations for the host populations of  $A$ ,  $B$ ,  $C$ ,  $D$ , and  $E$ . The question of which, if any, of the potential recursive events predominate would be determined by the measured event rate functions (generalized 's, and 's) used in that system of five equations<sup>15</sup>.

Returning to non-stratified memes, a great many phenomena that may look like two-meme problems are in fact applications calling for three or more equations. This includes cases that enumerate not only the hosts of a meme, but also the exposed and unexposed non-hosts. For instance, if we define mnemonic  $P$  as *belief* in proposition  $X$ , then non-belief is  $\sim P$ . If *awareness* of proposition  $X$  is designated as  $Q$ , then unawareness is  $\sim Q$ . An exposed non-host of the proposition is designated as  $P^*Q$ . Exposed hosts are  $P^*Q$ . Unexposed non-hosts are  $\sim P^*\sim Q$ . The  $P^*\sim Q$  combination presumably has a host population of zero. The resulting mathematical model therefore has three equations modeling  $N_1$  hosts of  $\sim P^*\sim Q$ ,  $N_2$  hosts of  $\sim P^*Q$ , and  $N_3$  hosts of  $P^*Q$ .

Host populations modeled as distributions in age and time in equation 1 and equation 2 can also be measured and modeled using a variety of additional parameters such as those of geography (e.g., longitude and latitude), duration of hostship (as distinct from the age of the host), family income, years of schooling, etc. The  $R$ 's,  $K$ 's, 's, 's, 's, and  $M$ 's could be modeled with the additional parameters as well. The  $R$ 's,  $K$ 's, 's, 's, 's, and  $M$ 's would then quantify a range of events that include individuals changing not only their belief their host status, but also individuals changing their locations, incomes, etc. The corresponding event diagrams would include non-memetic attributes, such as location coordinates  $x$ ,  $y$ , and  $z$ , so that the simple migration of a meme  $P^*Q$  host from location 1 to location 2 would become  $P^*Q^*(x_1, y_1, z_1)$   $P^*Q^*(x_2, y_2, z_2)$ . The modified a's could then quantify this migration event in addition to memetic events such as  $P^*Q^*(x_1, y_1, z_1)$   $P^*\sim Q^*(x_1, y_1, z_1)$ .

As an alternative to working with systems of difficult equations, one can also use quantified propagation mechanisms to run predictive computer simulations of memetic evolution without the intermediate step of writing down differential equations. The latter method can run particularly well on systems based on generalized cellular automata, where each "cell" simulates the memetic state and behavior of a person. Such an individual-based computer simulation is now feasible using the *SWARM* program developed by the Santa Fe Institute [8].

## 17 Qualitative and Quantitative Evolution

As mentioned earlier, occurrence rates of heterogenic events often depend on the prevalence of precursor memes. When vigorous precursor memes proliferate, they achieve substantial odds of causing the creation of memes that only a rare host can form. Additionally, when the new meme is simply *defined* as a combination of two or more precursor memes, the vigorous propagation of the precursors greatly hastens the arrival of the combination. So the independent spread of meme  $X$  and meme  $Y$  hastens the formation of the combination  $X*Y$ . Indeed, if  $X$  and  $Y$  spread vigorously,  $X*Y$  can pop up explosively among many widely separated individuals. So if  $X*Y$  tends to inspire hosts to create mnemonic  $Z$ , then  $Z$ , too can pop up explosively among widely separated individuals. This much acclaimed feature in the innovation of ideas shows that creativity is largely a population phenomenon.

The most vigorous precursor memes tend to recombine with more varieties of new ideas, some of which form an even more vigorously propagating meme package in combination with the precursor set. Mnemonic variation thus feeds new operands into the quantitative processes of natural selection while the quantitative processes give many subsequent qualitative variations an appreciable chance to occur. The two kinds of change continuously feedback on each other to form a genuine process of evolution. A recursive process therefore happens at a higher level than that of individual mnemonic replications. This is the iterating cycle of variation and selection that Dennett [5] describes as an *algorithm*. It is, more specifically, a *recursive algorithm*, a central feature of all evolutionary replicator theories.

## 18 Falsifiability

The sort of predictions generated by mathematical models and computer simulations are falsifiable. This provides a general method for attempting to falsify specific memetic hypotheses. In general, propagation parameters can be measured over the duration of the time interval being modeled, as can the initial host populations of the memes under study. If the propagation parameters remain in some specified interval, then the final host population plus or minus a calculable error margin should be measured at the end of that interval. A host population measurement outside those error margins would then falsify the model being used for the specific memes under study. The procedure can be quite elaborate, as it involves conducting surveys and performing difficult computations.

Yet particular aspects of a specific memetic hypothesis can be falsified by less elaborate means as well. If a proposed model for the spread of anti-abortionism hypothesizes that the belief "abortion as wrong" leads adherents to raise more children than non-adherents, then the whole model could be falsified by, for instance, showing that non-adherents raise equal or greater numbers of children. If such data were gathered, and proven representative of whole societies over long time spans, then there would be no need to begin the more elaborate procedures of measuring additional parameters and performing computations.

In some cases, this method can be performed using existing data sets. For example, the National Health and Social Life Survey [9] contains data about respondents' attitudes toward various sex acts. It also has data on reproductive histories. So, after making suitable functional definitions

based on the questionnaire, one can statistically reanalyze the raw data set to see how much of the variance in reproductive history is attributable to variance in a particular attitude response. Applying the method to the older respondents could provide results that tend to either corroborate or falsify a hypothesis that a particular sex taboo, for instance, increases its hosts' average number of offspring<sup>16</sup>.

The overall theory would conceivably be falsified by refuting its premises or the logic and mathematics based on those premises. The assertion that memetics is a *useful* theory is also falsifiable, for instance by showing that differences in propagation parameters are *never* great enough to account for large-scale ideological shifts, or that the only parameters that ever account for large-scale ideological shift are the "classical" receptivity parameters.

## 19 Other Empirical Issues

At present, there is no brain scanning device that tells unambiguously whether someone holds a particular belief or not. Nor is there a memory probe that recovers all the communications that lead someone to adopt a particular idea. This means that measuring host populations and propagation parameters must depend on survey methodology. Questionnaires and interview protocols may thus form functional definitions of the abstractions used in memetics.

The rates of occurrence of specific events must also be measured for use in a quantitative model, in order to determine the propagation parameters. Surveys already exist which ask how many children people have. Parent to child meme transfer rates would have to be measured by new survey designs that interview both parents and offspring. Human subjects would in general have to be asked to remember what roles their friends or parents may have played in imparting specific beliefs. Respondents "crediting" their parents or people the parents designate on the basis of memetic similarity would be counted toward a meme's parental transmission. Respondents "crediting" other people would be counted toward the meme's non-offspring conversion rate. Much work and proper funding are needed for such surveys.

In a few cases, raw data may have been incidentally gathered for entirely different purposes. The Mormons, for instance, have kept unusually extensive long-term records not only of member births and deaths, but also of how many children followed their parent's faith and how many dropped out, as well as which members arrived through non-parental conversions [10]. A project to re-tabulate such data by types of events should enable the calculation of parameters for use in the above differential equations or computer simulations. The resulting host population versus time figures can then be compared to the records of total membership also culled from church records.

With non-human animals, behavior and communication can in principle be monitored with tiny, individual electronic recording devices and studied in detail later. The method is ethically encumbered for humans, but some useful propagation data may still emerge in limited consensual monitoring studies. Though it gets past some of the limitations of surveys, it is still an enormous project.

Small recording devices may also be used in animal experiments in memetics. Bonner has documented the imitational spread in British titmice of the learned skill of pecking through foil caps of milk bottles delivered outside of houses [2]. A more deliberate experiment could place a large number of smaller artificial food containers throughout an extended (perhaps closed) environment and monitor them. They would be designed so as not to resemble any previous food source, but would be easily opened by the species under study. A container that looks like a distinctive pebble but can be pecked open in a particular spot might do. (To study multiple modes of transmission, care may have to be taken to insure that a single bird cannot have too many imitators in a single feeding, or else the meme will spread to saturate the population under study in less than one generation. Limiting the food per container and choosing a solitary-foraging bird may be necessary.) A large display screen with computer-simulated images of birds opening the containers and eating from them could then be used to release a synthetic meme into a small initial host population. Subsequent monitoring of the birds for reproductive rates, death rates, parent to offspring imitation rates, non-offspring imitation rates, and "forgetting" rates could yield useful parameters that differ between host and non-host populations. Observed host and non-host population sizes versus time can then be compared to those predicted using the quantitative theory and measured parameters.

Human society also affords some specialized environments in which to investigate memes by investigating meme relics. Modern libraries, for instance, allow us to track the proliferating citations of scholarly papers indexed in a computer database. Pocklington and Best [1, 14] have also provided strong evidence of meme proliferation by demonstrating the corresponding proliferation of meme relics on the Internet.

Some look at the lack of a strictly physical definition of meme instantiation as reason for dismissing memetic evolution entirely in favor of theories that describe the same phenomena in terms of genes. Genes, after all, have reliable molecular definitions and are detected using impartial technologies such as nucleotide sequencing. One problem with this argument is that even genetic psychology theories still require functional definitions of psychological features as abstract as a meme. Thus, if an evolutionary psychologist hypothesizes that the valuation of female premarital chastity is genetically based, the empirical investigation must include a method of deciding which females value premarital chastity and which females do not. Then some specific genetic factor must be linked to the presence or absence of this trait. It should, moreover, not just be responsible for learning in general, as the same gene would then account for the ability to learn either the valuation of chastity or the valuation of promiscuity. Rather, it should be a genetic factor whose phenotype and evolutionary adaptation can somehow be more specifically identified with chastity preferences than with strong sexual cravings or cerebral functioning in general.

In many such cases, memetic hypotheses are on empirically *stronger* footing than alternative genetic psychology hypotheses. We can, for instance, *observe* the hypothesized memetic replication event of parents telling sexually eager daughters to "say no". Yet no one has identified and *observed* a gene that makes daughters want to repress the sexual urges brought on presumably by other genes. The memetic hypothesis involves parents feeling economically motivated to impart the meme to daughters and having more grandchildren when they do so [12]. It is, in essence, as much Darwinian as the genetic hypothesis is. Yet the memetic hypothesis is

more parsimonious: a competing genetic explanation must still account for the *observable* communication of sexual repression messages from parents to daughters.

When a communication event is *not* observed, however, it can serve as indirect evidence of genetic influences. For instance, if parents and others are consistently observed *not* teaching their daughters any preferences for male lower body physique, but daughters develop their preferences anyway, then a genetic hypothesis for such preferences is on stronger footing than a memetic hypothesis. Specific genetic factors still need to be identified to completely validate this hypothesis, but it at least does not face a stronger and more parsimonious memetic alternative. Memetics does face methodological challenges in measuring prevalence and propagation parameters that are more complicated than the functional issues in polling. Yet these should not be seized as arguments for the universal superiority of genetic theories.

## Conclusion

Even physical scientists routinely take the existence of human "memory items" and "ideas" as a given. They even accept the concept that there is such a thing as "communication", in which one person can "cause", by complex means, the occurrence of an "idea" or "memory item" in another person. Sometimes the newly caused "memory item" in the other person is even considered to be "the same" as one in the former person, an event called "replication" by Dawkins and others. Replication events include many cases where participants assume the ideas to be playing very passive roles in causing new instantiations of "the same" ideas.

Yet even physical scientists again routinely assume that "ideas" can influence "behavior", including "communication behavior". When an "idea" causes "communication behaviors" that result in a new instantiation of "the same" idea in a different person, we have a generalized kind of autocatalysis, or "replication", in a stronger sense of the term. It is a stronger sense of the term "replication" primarily because these special-case ideas are seen as playing a stronger role in the causation of "the same" idea in new individuals. Ideas which play such strong causal roles in their own replication can thereby gain extreme prevalence and influence in a population. Moreover, their propagation results in the formation of new varieties of strongly replicating ideas. These observations cause dismay in many who have become accustomed to viewing themselves as controlling their ideas rather than the other way around, but major advances in scientific theory have often caused initial discomfort. The new theory offers a strong and parsimonious understanding of diverse phenomena in the areas of religion, sexual ideology, family structure, health beliefs, war, fascism, and numerous others. The theory of evolution by natural selection of memes is thus a solid and broadly unifying theory whose time has come.

## Acknowledgments

My gratitude goes to Carl Wegner for his early support of my memetics work.

Thanks also to Daniel Dennett, Valla Pishva, and Walter Cuirle for valuable comments on earlier drafts of this paper, and to Curt Hicks and Michael Crowley for contributing their time and talent to the 1991 precursor of this paper. Thanks also to Richard Pocklington, Kieth Henson, and Mario Vaneechoutte for engaging me in on-line discussions that in 1997 sharpened my thinking

for this expanded treatment, and to the anonymous reviewers who likewise contributed their insights. Special thanks to publisher Bruce Edmonds for working patiently and diligently with this long paper.

## Notes

1. From Dawkins [4], p. 109.
2. See the discussion of the meme for bifocals in Dennett [5], p. 353.
3. "Mnemon" is being introduced as a neologism here, and may differ in meaning from the word as coined elsewhere for use in other fields.
4. The term *ideosphere* was co-invented by myself and Douglas Hofstadter. See *Metamagical Themas* by Douglas Hofstadter [6], p. 66.
5. Dawkins [4], uses the term *phenotypes* in connection with memes to describe the "outward and visible (audible, etc.) manifestations of the memes within the brain".
6. See Dawkins [4] p. 112.
7. See Dennett [5], p. 346. He uses this statement as a pithy summation of the memetic perspective, whereas I regard it as an example of what memetics is *not*. Libraries do indeed help scholars to write new books, and the production of new books plays a crucial role in causing the formation of new libraries. In this sense, libraries can be treated as cultural replicators that use scholars (among other things) in the process of replication—especially as the abstraction "library" does not call for even a single shared volume between "parent" and "daughter" libraries. But the phenomenon captures only the *flavor* of memetics, as the library itself is not a meme.
8. See Dennett [5], p. 369 for the summary of chapter 12.
9. Colony sizes here come from Hostetler [7], p. 188, Table 1, which presents demographic charts of a colony branching event in 1969.
10. Dennett [5] argues effectively on pp. 473-475 that the Hutterites exhibit a striking *clash* between memes and genetic proclivities.
11. This argument is spelled out more completely in [13].
12. If there is a non-zero rate of completely "spontaneous" dropouts in the model, then it is technically more accurate to use the terms "mnemon-1" and "mnemon-2" in place of "meme-1" and "meme-2".
13. The population age-time profiles [ $N_1(a,t)$  &  $N_2(a,t)$ ], fertility rates [ $R$ 's], child inculcation rates [ $K$ 's], non-parental conversion rates [ 's, and 's], and mortality rates [ $M$ 's] are represented as distribution functions, but this does not mean that smooth continuous functions

will ever emerge from the data gathered for an actual application. Rather, age and time will be divided into discreet intervals such as a month, a year, or a decade. The population age-time profiles, fertility rates, etc. will then be the kind of functions often represented with 2- or 3-dimensional bar graphs. The specific time and age resolution ("granularity") used in data gathering will in general have some effect on the accuracy of computational solutions to equation 1 and equation 2. This is because differing ways of subdividing time and age into discrete intervals amount to making different approximations of the functions in these parameters. Conversely, the required accuracy plays a key role in determining just how finely time and age must be partitioned into intervals to gain useful data for a specific application. Once such a partitioning is chosen for a given application, the distribution functions in time and/or age take on a finite number of measured [ $R$ 's,  $K$ 's,  $\alpha$ 's and  $M$ 's] or computed [ $N_1(a,t)$  &  $N_2(a,t)$ ] values and the integrals may be reduced to finite summations for computational purposes. The present treatment does not use any specific time/age partitioning scheme, because the optimal scheme varies across applications. Hence, functions are left in their short-hand form rather than as data sets for fertility rates, child inculcation rates, non-parental conversion rates, and mortality or computed result sets for population age-time profiles. (The  $a$ 's are treated as constants, but they too could be measured as functions of age  $a$ .)

14. The term "longevity" might be replaced by "durability" to better signify the resistance to dropout and mortality.

15. Recursive multistage events such as the ones that add up to  $A+4E - 5E$  or  $E+4A - 5A$  are especially likely to predominate if people at the extremes ("strongly agree" or "strongly disagree") play a preponderant role in persuading others to take those small steps of unidirectional mnemonic change. The circumstance is probably common. Sequences of heterogenic events adding up to smaller recursive homogenic events such as  $C+2E - 3E$  and  $C+2A - 3A$  might happen most frequently, because they do not require shifting an individual from one extreme to another-although there are individuals who go from one extreme to the opposite on a topic. To model all possible two-party conversion events for  $A, B, C, D$ , and  $E$  where only one participant changes mnemonics requires 80 event rate functions (160 if both b and g types are used), but measurement could determine that many of them are negligible in a given application.

16. I say "*tend to* either corroborate or falsify" because there remains the possibility that a taboo that increased reproduction for thousands of years no longer does so, or that a taboo that had no effect for thousands of years suddenly made a difference in the 1994 data.

## References

[1] Best, M., L., 1997. Models for Interacting Populations of Memes: Competition and Niche Behavior. *Journal of Memetics - Evolutionary Models of Information Transmission*, 1.  
[http://jom-emit.cfpm.org/1997/vol1/best\\_ml.html](http://jom-emit.cfpm.org/1997/vol1/best_ml.html)

[2] Bonner, J. T. 1980. *The Evolution of Culture in Animals*. Princeton: Princeton University Press

- [3] Dawkins, R. 1976. *The Selfish Gene*. Oxford: Oxford University Press.
- [4] Dawkins, R. 1982. *The Extended Phenotype*. San Francisco: W. H. Freeman and Company.
- [5] Dennett, D. C. 1995. *Darwin's Dangerous Idea*. New York: Simon & Schuster.
- [6] Hofstadter, D. R. 1985. *Metamagical Themas: Questing for the Essence of Mind and Pattern*. New York: Basic Books.
- [7] Hostetler, 1974. *Hutterite Society*. Baltimore: The Johns Hopkins University Press.
- [8] Langton, C., Burkhart, R.r, Lee, I., Daniels, M., Lancaster, A., 1998. *The Swarm Simulation System*. <http://www.santafe.edu/projects/swarm>
- [9] Laumann, E. O., Gagnon, J. H., Michael, R. T. and Michaels, S. 1994. *The Social Organization of Sexuality*. Chicago: The University of Chicago Press
- [10] Ludlow, D. H. ed. 1992. *Encyclopedia of Mormonism*. New York: Macmillan Publishing Company.
- [11] Lynch, A. 1991. Thought Contagion as Abstract Evolution. *Journal of Ideas*, 2,:3-10. Republished with revisions at <http://www.mcs.net/~aaron/mememath.html>. Scanned unrevised reprint at <http://www.mcs.net/~aaron/Lynch1991.htm>.
- [12] Lynch, A. 1996. *Thought Contagion: How Belief Spreads Through Society The New Science of Memes*. New York: Basic Books.
- [13] Lynch, A. 1997. *Thought Contagion and Mass Belief*. <http://www.mcs.net/~aaron/tcamb.htm>
- [14] Pocklington, R., and Best, M. L. 1997. Cultural Evolution and Units of Selection in Replicating Text. *Journal of Theoretical Biology*, 188:79-87.
- [15] Rogers, E. M. 1983. *The Diffusion of Innovations*. New York: The Free Press.
- [16] Stark, R. 1996. *The Rise of Christianity*. Princeton: Princeton University Press.
- [17] Wilson, D. Sl., and Sober, E. 1994. Re-introducing Group Selection to the Human Behavior Sciences. *Behavioral and Brain Sciences*, 17:585-608.

## A LETTER ON: RATIONALES FOR MATHEMATICAL MODELING AND POINTS OF TERMINOLOGY

By  
**Aaron Lynch**  
aaron@mcs.net

Michael Best opens his *Critique of Memetic Models* letter (Best, 1998) with what might be taken as a broad swipe at new science being done under the heading of 'memetics'. Yet Best has himself published in this journal (Best, 1997), using a very recent definition of meme and very recent methods. So it might seem surprising that he should now proclaim a sweeping 'weakness' in memetics models or cast doubts over thoughts of 'new science' in the field.

The letter does not mention, however, that my *Units, Events, and Dynamics* paper (Lynch, 1998) finds a wide range of meme definitions, including his, functionally encumbered and contrary to Dawkins (1982). In particular, the considerations in section 12 of my paper find an unrecoverable problem in the broader usefulness of definitions based on the embedded theoretical construct of 'unit size', because sets of proposed cultural units can in general only be considered partially ordered in the set theoretic sense when applied to human memory items. This affects both 'smallest units' and 'largest units' definitions of the word 'meme'. My paper also maintained that to avoid terminological confusion, we should remain consistent with Dawkins's clarification that memes are informational replicators *residing in the brain*. These were matters of terminology, rather than criticisms of Best's empirical netnews findings or the findings of many others who have used vastly divergent definitions of the word *meme* or different words altogether. I even proposed a more general term than memetics, namely, *cultural repliconics* to encompass the study of divergent kinds of non-brain based replicators. I also favor inserting the word *and* into the title of this journal to emphasize its inclusion of the broader scope of evolutionary cultural replicator theory: *Journal of Memetics and Evolutionary Models of Information Transmission*. Still, it is possible for one to misinterpret a merely terminological discussion as a declaration that researchers along the lines given by Best are 'memetics outsiders'. One might further conclude, then, that papers such as Best's cannot, by definition, be considered exemplars of 'weaknesses' that he presently claims to see in memetics. This possibility diminishes the surprise we might feel about broadly condemnatory tone in Best's opening remarks. The rest of his essay, however, shows a variety of serious flaws which I address severally.

Best's first specific complaint was that the *Units, Events, and Dynamics* paper should have cited Campbell (1974) for first proposing the term 'mnemon', even though Campbell does not use this term. Campbell (1974) does use the phrases "*nonmnemonic problem solving*" and "*mnemonically supported thought*", but does not use or introduce the word "*mnemon*". Nor does Campbell indicate that the substring 'mnemonic' means anything other than 'pertaining to memory'. The term 'mnemonic' is sufficiently old that derivatives such as 'nonmnemonic' and 'mnemonically' do not warrant special citation in the context of my paper. Nevertheless, a few words about the derivation and history of the word 'mnemon' are in order. Although I formed the word from its Greek roots, this should be recorded as a reinvention of a word that was once coined before in a physiology article about engrams (Cherkin, 1966). That 1966 term, which does not seem to have

gained wide use, was originally defined as "*the minimum physical change in the nervous system that encodes one memory*". It is not defined as something that is 'the same' from one organism to another. Because I derived the term "mnemon" from its Greek roots and gave it a different meaning, it would have been inaccurate to attribute my usage to Cherkin. Nevertheless, my paper might have benefited from some mention of the prior use of 'mnemon' by Cherkin, especially as this may have prevented any suggestion that I was plagiarizing Campbell.

It would also have been interesting and informative to include a paragraph calling attention to the mathematical work of Cavalli-Sforza, Feldman, Boyd and Richerson, and Lumsden and Wilson in the preface to section 16 of my own paper. This is not because the equations I present are derivable from equations by these other authors, but rather, to point out that a range of models exists and to make it clear to readers that the present model differs from the others. Still, my own paper is not in any way offered as a review of literature, and should not be critiqued as such. There is much that can be said about the relative merits of different mathematical models, and why my own model is warranted in light of earlier work, but proposing a new model does not of itself require taking on the lengthy task of sorting through older models. The equations in section 16 are quite different from those of the other authors that Best has cited. They were also developed 'from scratch', rather than by borrowing from the other authors. Had it been otherwise, then failing to cite the other authors would have been a serious error indeed.

I chose to use the language of mathematics to discuss the quantitative consequences of phenomena discussed earlier in the paper. However, the transition into the mathematical language does not represent any kind of 'magic moment' at which one must always pause to give recognition to other cultural scientists who have spoken in mathematical language, when what is presently being said is not based on the previous works. Likewise, if one voices original arguments in prose, the author should likewise not be faulted for not citing non-source works that happen to use English prose for substantially different arguments. The scholarly courtesy of citing other mathematical models of cultural evolution is a good suggestion, but not mandatory in the present case. Original arguments in the language of mathematics do not have a different standard for "mandatory" citations than do original arguments expressed in prose. Unlike the physical sciences, where the use of mathematics is routine, there may be an irrational mystique about mathematics in the social sciences, but it is unwarranted.

Michael Best complains that my model fails to "*describe empirical data (present or sought)*", to which I invite readers to read the original article and consider whether host populations of memes should be considered empirical data or not. He also asserts that the equations shed no explanatory light on phenomena despite the fact that they express changes in meme prevalence in terms of rate parameters for such constituent processes as having and teaching children. As for using the equations to model a specific meme propagation phenomenon, this is indeed a more extensive project than merely publishing equations. Nevertheless, the equations are far more usable when published than when unpublished. Best fails to see the publication of equations as part of a larger process. For this, we can examine the field of theoretical physics, where equations are frequently published without yet having a specific plan to test them. We may also wish to reexamine the other works of mathematical culture modeling to see numerous examples of equations published without specific testing plans. I will not do so here, however, as I am still not offering a review of literature.

Best has further faulted the equations for modeling everything as continuous functions. Best goes on: "...even within his formalisms, Lynch is compelled to discretise time." This is confusion. I have chosen the year as an arbitrary unit of time, but I have nowhere indicated that time is an integer or an integer multiple of any quantity. Nor have I imposed any kind of discrete reproductive generations on a population. In an actual study, it will generally be necessary to choose a finite number of times at which to measure host populations, and so forth, but these details are not built into the equations themselves. They are left to the discretion of potential users of the equations. The corresponding approximate evaluation of integrals as finite summations and derivatives as ratios of differences is also left to potential users of the equations. Why Best chooses to demand formal inclusion of a scheme to "discretise time" for my equations in advance of specific application eludes me. Would he demand the same from certain other mathematical models of culture or indeed, mathematical models of physical phenomena?

Best has asked why the age of a parent (more exactly, a meme host of a given age who may have zero to many children) is relevant. The reason is that human reproduction is closely tied to the phases of our species' life cycle, so I have chosen age as the variable with which to describe phase in the life cycle. A 5-year old meme host will generally have about zero children per year, as will a 95-year old meme host. Mortality is also modeled as age-dependent, largely for biological reasons beyond the scope of this journal. Yet once the number of hosts at a given age needs to be known for one term in the equation, it must be modeled in all terms of the equation. As it happens, the rate at which many memes are learned from parents also depends on the age of the child, even if the parents' age matters little. As I have pointed out in the paper, the model can be simplified to treat each meme as being learned at just one effective childhood age and at a rate such as  $k_{11}$  that is invariant with respect to parent age. Treating  $k_{11}$  as invariant with respect to time, location, socioeconomic status, etc. is of course a simplification too.

Still another complaint that Best lodges against  $K_{11}(p, a)$  regards its definition as a ratio rather than a probability. It is true that I could have defined it as the probability per unit time of meme transmission from a parent of age  $p$  to their child of age  $a$ . Correspondingly,  $N_1(a, t)$  and  $N_2(a, t)$  would have been defined as expectation values of host population-age profiles. This is a matter of taste. Probabilities, however, are empirically measured by taking ratios. The probability per unit time of a child of age  $a$  acquiring a meme from an age  $p$  parent, for instance, is measured by taking the ratio in a representative sample of population of those age  $a$  children with age  $p$  parents acquiring the meme in a representative time interval to the total number of age  $a$  children with age  $p$  parents. There is, however, a branch of "*arm-chair theorizing*" known as statistics that concerns itself with such questions as how likely it is that a parameter measurement taken on a 'representative sample' matches the parameter for the whole population to a given degree of accuracy. I have deliberately left this sort of analysis for future work. Section 16 of my paper serves only as a starting point for a particular line of mathematical modeling in population memetics, and was certainly not offered as a complete and final treatise.

As I stated clearly in the paper, "*equation 1 and equation 2 model fairly ideal cases of the two-idea propagation problem.*" Thus, I do not include certain complications that can arise from biparental transmission, such as memetically mixed marriages. So the present model works better when the proportion of meme-discordant marriages is either quite low (e.g., with certain religions) or remains constant, with reproductive rates of mixed families well modeled as the

average of the two meme-concordant cases. In the constant intermarriage case, the meme-1 partner is still credited with an  $R_1(a)$  reproductive rate and the meme-2 partner is credited with an  $R_2(a)$  reproductive rate. However, because some meme-1 parents' children are being partly raised by meme-2 spouses, and vice versa, we expect the measurement of  $K_{12}(a, p)$  and  $K_{21}(a, p)$  to show significant effects of crossover inculcation. (The reduction of  $K_{12}$  and  $K_{21}$  figures is one of the evolutionary pressures favoring intermarriage taboos in religions.)

In cases where the proportion of intermarriages changes substantially over time, or where reproduction rates are poorly modeled as a by averaging the two unmixed cases, a more elaborate modeling of parental transmission is called for. For example, a set of four equations could model the population of female meme-1 hosts, male meme-1 hosts, female meme-2 hosts, and male meme-2 hosts. A model for intermarriage rates as a function of (for instance)  $N_1(a, t)$  through  $N_4(a, t)$  is required, as are additional  $K(p, a)$  and  $R(a)$  values for meme-concordant and meme-discordant parents. (More generally, for all possible parental meme combinations being modeled. Again, the simplified  $k_{11}$  etc. parameters may be used.)

Still further elaborations can be pursued for increasingly sophisticated and realistic models. Yet even when the simpler system of equations can be used, the application of the model to a specific meme is an extensive project requiring further resources and collaboration. It is unreasonable to demand that such a project be well underway before a system of equations can even be published, but the question of how to handle two parents and meme discordance is appropriately raised at this stage. Presenting a first-order model of biparental transmission may seem a bit like treating a person as a point mass in a physics problem, but even the point-mass person is a good approximation in certain situations. However, for the most ambitious investigations of meme propagation, one should be prepared to face a degree of mathematical and computational complexity rivaling the modern models of terrestrial weather.

Best apparently demands that I model parent to child meme transmission in terms of a whole "host of factors" without being specific. Does the system of equations need to take account of every conceivable variable, or else nothing at all? If so, this could demand a degree of accuracy that would have prohibited early models of the atom as well. I agree that a whole host of variables can be added to the model, but how can one reconcile demands for a host of added variables while at the same time complaining about the "*sheer complexity*" of the model? I can only conclude from these incongruous demands that Best would have rejected any model I had proposed, regardless of potential utility. And I must ask him if he knows of a mathematical model that actually meets all of his demands, to please identify it.

The purpose of my present reply to Best's essay is not to provide details for the next and subsequent levels of elaboration to the mathematics in section 16 of my paper. At 17,100 words, the *Units, Events, and Dynamics* paper is already quite lengthy for a journal article, and elaborated models require entire new articles at the very least. Another kind of paper is one that simplifies the equations, as by taking lumped, single-event models of reproduction and child inculcation. The value of such simplifications would be to demonstrate that a particular magnitude of reproduction difference, for instance, can in principle lead to significant meme proliferation even if there are also non-parental transmission processes at work. Alternately, if even large reproduction differences never produced much long-term effect, then this may be

taken as a falsification of the hypothesis that reproduction differentials sometimes account for large relative gains in the host populations of memes. The presentation of the population memetics section in mathematical terms should be seen as an invitation to investigate such matters quantitatively, rather than by sheer guesswork. It should also be seen as a demonstration that methods do exist for translating specifically memetic theory into mathematical/quantitative terms, which is widely deemed essential to taking any variant of scientific theory seriously. An absence of such work among those concerned with "memes" could easily have been taken as a signal that in order to see serious mathematical reasoning, one must ignore those using the word "meme" in favor of those using words conceived in association with highly sociobiological tenets, such as 'culturgens' (Lumsden and Wilson, 1981). A deep suspicion exists among many scientists for lines of thought advanced without any hint of possible mathematical expression, for such efforts are widely seen as vague and evasive. In any case, Best falls far short of justifying his assertion that the mathematical model "*does damage to the problem.*" Because such phrases may discourage people from pursuing the mathematical investigations further, any claim of 'damage' should really be withheld unless strongly justified--unlike the present case where the term 'damage' is not even remotely warranted.

Best comes very close to asserting that no new mathematical models should ever be developed directly from a particular combination of concepts including 'transmission event', 'host population', 'reproductive rate', 'child inculcation rate', and so forth. Instead, he seems to assert that one must limit modeling to efforts to building upon past equations developed around substantially different premises and parameters. For this, he gives no justification, theoretical or otherwise.

Best has also seriously misread the transmission of a negatively defined mnemonic (e.g., the lack-of knowledge of birth control knowledge spreading by way of more numerous childbirths) as a "forgetting" event. Individual forgetting events are of course possible, and are a subset of individual dropout events generally (e.g.,  $A \rightarrow \sim A$ ). One can also persuade a peer to drop a belief, for instance. But nowhere in the paper is there a notion of peer to peer 'forgetting events'. Subsequent to his letter, Best explained to me that he got the idea of the transmission of lack of birth control information from his reading of the clause "*people can be taught about birth control far more easily than they can be made to forget about it.*" Apparently he didn't notice that while knowledge of birth control is discussed as transmitted, the lack of such knowledge is not. To *make* someone forget about birth control could involve drastic measures such as concussion or drug induced amnesia, or even lobotomy. Such extreme measures do not constitute 'transmission' of lack of knowledge, because 'transmission' presumes that lack of knowledge existed in the person who performed the lobotomy, drugging, etc., and that such lack of knowledge played a role in causing or directing his actions. Moreover, nothing in my paper suggests that lack of knowledge can be *taught*. Given this wild misreading of part of a sentence in section 5 of my paper, I can understand how Best would expect the worst from the whole rest of the paper. This and other misunderstandings show that before concluding that a refereed paper proposes an outlandish process divorced from reality, a few well-placed queries to the author may help.

Best barely touches upon the original disagreement he had with me, namely, the definition of 'meme'. After praising Richard Dawkins (1976), he asserts merely that 'meme' describes

"replicating cultural units." Yet Best cannot flatter Dawkins for being so nonspecific as this, for Dawkins has distinctly clarified (1982) that the word he coined refers to brain-based replicators only. Ironically, Best, whose paper and letter overlook the specificity in Dawkins (1982), now calls out for "*an appreciation of past art*". Yet this disagreement between me and Best is only a matter of terminology. We may have a deeper disagreement over the value of doing new work on theoretical frameworks in memetics. Yet regardless of the broader nature of our disagreements, more careful readings of my own paper and related works such as Campbell's are in order.

## References

- [1] Best, M.L. (1997). Models for Interacting Populations of Memes: Competition and Niche Behavior. *Journal of Memetics - Evolutionary Models of Information Transmission*, 1. [http://jom-emit.cfpm.org/1997/vol1/best\\_ml.html](http://jom-emit.cfpm.org/1997/vol1/best_ml.html)
- [2] Best, M.L. (1998). A Letter on: Memes on memes - A Critique of Memetic Models. *Journal of Memetics - Evolutionary Models of Information Transmission*, 2. [http://jom-emit.cfpm.org/1998/vol2/best\\_ml.html](http://jom-emit.cfpm.org/1998/vol2/best_ml.html)
- [3] Boyd, R. & Richerson, P.H., (1985). *Culture and the Evolutionary Process*. Chicago: University of Chicago Press.
- [4] Campbell, D.T. (1974). Evolutionary Epistemology. In P.A. Schilpp (Ed.). *The Philosophy of Karl Popper*. La Salle, IL: Open Court Publishing.
- [5] Cavalli-Sforza, L. & Feldman, M. (1973). Models for Cultural Inheritance, I: Group Mean and Within Group Variation. *Theoretical Population Biology* 4: 42-55.
- [6] Cavalli-Sforza, L. & Feldman, M. (1981). *Cultural Transmission and Evolution: A Quantitative Approach*. Princeton, NJ: Princeton University Press.
- [7] Cherkin, A. (1966). Toward a Quantitative View of the Engram. *Proceedings of the National Academy of Sciences* 50: 88-89.
- [8] Dawkins, R. (1976). *The Selfish Gene*. Oxford: Oxford University Press.
- [9] Dawkins, R. (1982). *The Extended Phenotype*. San Francisco: W. H. Freeman and Company.
- [10] Lumsden, C.J. & Wilson, E.O. (1981). *Genes, Mind and Culture: The Coevolutionary Process*. Cambridge, MA: Harvard University Press.
- [11] Lynch, A. (1998). Units, Events and Dynamics in Memetic Evolution. *Journal of Memetics - Evolutionary Models of Information Transmission*, 2. [http://jom-emit.cfpm.org/1998/vol2/lynch\\_a.html](http://jom-emit.cfpm.org/1998/vol2/lynch_a.html)

## **MISLEADING MIX OF RELIGION AND SCIENCE: A COMMENTARY ON GATHERER'S PAPER: *THOUGHT CONTAGION* *METAPHOR IS RETARDING THE PROGRESS OF MEMETICS***

By  
**Aaron Lynch**  
[aaron@mcs.net](mailto:aaron@mcs.net)

In the abstract and the body of his paper, Gatherer shows particular concern with the incorporation thought contagion memetics into critiques of religion. In this, we see continuity with objections expressed in his paper "Meme Pools, World 3, and Averroës's Vision of Immortality" (Gatherer 1998a). There he states that "Richard Dawkins has developed his meme concept as the philosophical basis for a militant atheism," and attempts to demonstrate instead that "acceptance of the meme concept need not necessarily lead to atheism, as Dawkins and his more zealous followers would maintain." The paper also argues for a particular concept of immortality. Elsewhere in that paper, Gatherer clearly considers his religious views as being challenged and berated by those zealous Dawkinsians and "mind virology." Against this background, we can understand Gatherer (1998b) as an extension and intensification of his largely religious agenda of countering "militant atheism." We can also take it as a continuation of his critique of "mind virology," (Gatherer 1998a) which focused attention on the pejorative connotations of the term "mind virus." In both respects, Gatherer's efforts are misguided. First, thought contagion theory is not offered as a basis for atheism or as a critique of religion. It is offered as a theory of how both religious and secular ideas evolve and spread. Second, the term "thought contagion" was explicitly given a neutral connotation in (Lynch 1996a), and was even applied to socially positive ideas.

Gatherer's present method of defending religion from what he perceives as the "thought contagion/mind virus" critique is more drastic than the approach he took in his 1998a paper. He has quickly switched from accepting the memetic analysis of religious belief in (Gatherer 1998a) to treating this subject as forbidden by definition in closing his present paper (Gatherer 1998b). Exactly what happened to change his mind so quickly and sharply is unclear. However, pages 206 to 210 of his 1998a paper suggest that he was recently still unaware of the evolutionary epidemiological ("thought contagion/mind virus") explanation for the origins of prevailing forms of monotheism by natural selection, as advanced in (Lynch 1996a). Having presumably finished reading (Lynch 1996a), Gatherer now walls off all belief from memetic analysis, by redefining meme to exclude all internally stored information. This new "external" memetics then doubles as something whose progress is "retarded" by the "thought contagion metaphor." With its subject thus phrased in terms of the "progress of memetics," the article may appear to have a mostly scientific agenda rather than the religious agenda expressed in his 1998a paper, which appeared in a journal of religion and science.

Gatherer offers so many misrepresentations of Lynch 1996a, Lynch 1996b, and Lynch 1998 that I must advise anyone reading Gatherer's present paper to read my work directly to see what it really says. I do not in any way suggest that the brain has something analogous to RAM memory, for instance. Even material Gatherer presents as quotations must be double checked for context-changing distortions. For example, I do not pretend to know precisely how the brain stores

information. So I used the conditional sentence "If a mnemon resides very redundantly in someone's brain, that person still counts as only one host and one mnemon instantiation." (Lynch, 1998). In (Gatherer 1998b), the antecedent is quoted alone to give the impression that I definitely posit that "a meme - sides very redundantly in someone's brain." This becomes the basis for his assertion that I treat the brain as containing "memory banks" of awarenesses "reminiscent of a computer RAM." In correspondence, Gatherer finds nothing at all wrong with this, so I can only invite readers to check the passages in question themselves and form their own conclusions. Gatherer's present paper has so many other misrepresentations of my work that I cannot discuss them in the limited space allotted to commentaries.

The question of whether ideas, thoughts, beliefs, and opinions are valid subjects of science is yet another topic far too extensive to cover in the space of a commentary. Hard core behaviorists treat these as unobservables that cannot be independently verified. They stand in opposition to cognitive scientists, sociologists of ideology, and many others. Nevertheless, I should clarify that many instances of what I call "mnemons" are highly connected to behavior. If a rat is operantly conditioned to run a maze, then the learned and internally stored skill is a mnemon, while the actual running of the maze is not. If the rat conditions her pups to likewise run the maze, then their internally stored skill is not only a mnemon, but also a meme by virtue of being copied. In any case, memetic analysis of internally stored information does not repudiate empirical science as Gatherer suggests. I refer readers to the discussion of empirical methods and animal experiments in sections 17, 18, and 19 of (Lynch 1998).

Despite the discussion of mathematical and quantitative methods in memetics in section 16 of Lynch 1998, Gatherer claims that "Quantification became impossible." This comes partly from Gatherer's unfounded insistence that quantification requires a one to one correspondence of observed behavior to internal information. By this argument, we might as well throw out all science of operant conditioning as well, for not giving us perfect correlations between conditioning history and behavior. Or we might forbid use of antibody tests to infer the presence of HIV, again because they are imperfect and indirect. Opinion polls and various other indirect tests for internal information can, of course, be deliberately confounded. But indirect tests for internal information will be hard to escape regardless of how the word "meme" is defined. Indeed, universities use imperfect instruments called "exams" to indirectly quantify something they call "knowledge."

A wealth of original, if confused arguments permeate not only the whole body of (Gatherer 1998b), but also particular subsections, such as the section 2.6 argument that forms the basis for his claim of non- quantifiability. Consider just the passage where he attempts to show that mnemons for awareness of a proposition are impossible because having one such mnemon supposedly requires the individual to have an infinite number of them: "However, any individual who has heard of Napoleon at all will be capable of constructing and transmitting an infinite variety of mnemons of the form 'Napoleon died in x'." This demonstrates a casual lack of understanding of the term "infinite." Not only is it physically impossible for humans with finite brains to construct an infinite number of dates for Napoleon's death, but there are single numbers that humans cannot construct. If we could construct 10 digits per second for 100 years, this is still only 32 billion (32 thousand million) digits in a lifetime. Therefore, there are individual dates that cannot even be conceived. I don't claim to know just how many alternative possible

Napoleonic death dates typical people might consider--I have focused on other topics. But there is nothing to my work that requires the number to any greater than the number of years of recorded history, let alone infinity.

Gatherer goes on to claim that treatment of beliefs as memes "lacks explanatory power." While many people who do not wish to see their belief systems explained in terms of evolutionary epidemiology may feel drawn to this position, I must again recommend that scientists read works such as (Lynch 1996a, Sperber 1996, Balkin 1998) and others to decide for themselves. Does internally stored information about numbers of players to recruit explain why baseball out-propagates tennis in the U.S? You have to read (Lynch 1996a) for yourself. Gatherer asserts that "To say that behaviour can be explained in terms of something which cannot be observed (e.g., beliefs) is unhelpful." Has he never told anyone "I wasn't there because I thought the meeting was tomorrow"? Anyone who makes such a statement is explaining their behavior in terms of something that, according to Gatherer, cannot be observed. If one person's belief causes behavior that leads to another person having the same belief, then this is the replication that makes the belief a minimal meme contagion. So if Gatherer ever explains his actions in terms of beliefs acquired from someone else, then he is going back on what he has said in focus item C. Still, if anyone tried to call meme contagion the totality of psychology, then they would indeed "trivialize psychology." But it is a gross misrepresentation to suggest that my work does so.

My initial reaction to Gatherer's 12,000 word paper was to type my numerous rebuttals as red text in the html file. Those comments took up far more space than is allotted for our journal's commentary format. So I instead offer this shorter commentary that barely adheres to the 1500 word limit. Besides, a careful reading or re-reading of (Lynch 1998, 1996a, 1996b) would probably clear up any confusions imparted by reading (Gatherer 1998b) much more efficiently than reading a paragraph by paragraph rebuttal. Suffice it to say that I still consider "internal" memetics highly tenable despite Gatherer's attempt to persuade otherwise.

## References

- Balkin, J.M. (1998). *Cultural Software: A Theory of Ideology*. New Haven: Princeton University Press.
- Gatherer, D. (1998a). Meme pools, World 3 and Averroës's Vision of Immortality. *Zygon: Journal of Religion and Science*, 33:203-219.
- Gatherer, D., (1998b). Why the 'Thought Contagion' Metaphor is Retarding the Progress of Memetics. *Journal of Memetics - Evolutionary Models of Information Transmission*, 2. [http://jom-emit.cfpm.org/1998/vol2/gatherer\\_d.html](http://jom-emit.cfpm.org/1998/vol2/gatherer_d.html)
- Lynch, A. (1996a). *Thought Contagion: How Belief Spreads Through Society: The New Science of Memes*. New York: Basic Books.
- Lynch, A. (1996b). Contribution to 'Memes: Self-replicants or Mysticism.' Wired debate October 1996. <http://www.wired.com/braintennis/96/43/bt0a.html>

Lynch, A. (1998). Units, Events and Dynamics in Memetic Evolution. *Journal of Memetics - Evolutionary Models of Information Transmission*, 2. [http://cfpm.org/jom-mit/1998/vol2/lynch\\_a.html](http://cfpm.org/jom-mit/1998/vol2/lynch_a.html)

Sperber, D. (1996). *Explaining Culture: A Naturalistic Approach*. Oxford: Blackwell Publishers.

## **A RESPONSE TO PAUL MARSDEN OF A REVIEW OF *THOUGHT CONTAGION: HOW BELIEF SPREADS THROUGH SOCIETY***

By  
**Aaron Lynch**

Marsden's (1999b) review of my book *Thought Contagion: How Belief Spreads Through Society* listed numerous supposed flaws in the text, but failed to mention the one real flaw most crucial to readers of JASSS. That flaw, imposed by standards of simplicity in the trade book market, was the complete omission of technical, mathematical and quantitative material I had published earlier and submitted as a pair of appendices. A serious gap in the coverage of the book arose from the fact that the mathematical population memetics analysis (Lynch 1991) could not be included even as an appendix. The system of non-linear, partial differentio-integral difference equations for host population as a function of time would have provided a good method of parameterisation for simulations of interest to JASSS readers. The definition of a "meme" as a "homoderivative mnemonic" (memory item caused by prior instantiation of the same memory item) would also have lent a technical clarity that JASSS readers could utilise. Moreover, the non-metaphoric event diagrams used in that paper could have formed a useful conceptual framework from which to design memetic simulations. Yet Marsden has previously mislabeled the equations (republished in Lynch 1998) as "... arbitrary mathematical manipulations ..." and the event diagrams as "... laws of combination and permutation ..." (Marsden 1999a). So we should not be surprised that he has not flagged their absence as a serious flaw. The 1998 edition of this technical, mathematical and quantitative population memetics (Lynch 1998) even receives further negative commentary in Marsden's review of *Thought Contagion*.

Other aspects of Marsden's review demonstrate that by carefully choosing which sentences and sentence fragments to quote or misquote out of context, and how to mislabel them, one can make any book look totally silly. So I invite JASSS readers to read chapter 1 of *Thought Contagion* online and judge for themselves whether the book's opening is really as silly as Marsden makes it look.

While Marsden's review does mention topics covered in *Thought Contagion*, his portrayal of my analysis is misleading to say the least. For example, there really is a section on the evolution of memes pertaining to homosexuality (p. 79-84). In a nutshell, it says that adherents of the taboo out-procreated more tolerant people over the course of many generations in ancient times, leading to increased prevalence of the taboo. (This does not require anything like a perfect correlation between morality and behaviour, or a perfect child inculcation rate, but only enough to increase taboo prevalence by several percent per generation over hundreds of generations.) Then horizontal transmission kicked in as people maligned homosexuality to "prove" their adherence to the taboo. As the taboo becomes extremely widespread, most homosexuals live heterosexual lives, leading them to reproduce any genes involved. As these genes gain prevalence, the rate of taboo dropout increases. Gene carriers who have dropped the taboo are more sexually and socially motivated to spread acceptance of homosexuality than are non-gene carriers who drop the taboo. So the rising gene prevalence can lead to a self-sustained propagation of pro-gay memes. (Horizontal transmission again, contrary to Marsden's claim that I ignore this mode.) That, in turn, can lead to lower gene prevalence in the next generation, and

even lower prevalence of pro-gay memes. All of this leads to potential fluctuations over long time spans. Incidentally, there is a brief mention of the way that beliefs about anal sex become involved with homosexuality taboos. Yet despite a multi-faceted discussion of homosexuality taboos, Marsden excerpts a sentence on anal sex and presents it as if I had offered it as the answer to the question "What memes deter homosexual behaviour?" He has thus taken a complex argument and misrepresented it as childishly silly. He uses a similar distorting approach in all of his other indented excerpts.

The "appendix" in which Marsden supposedly lists "Lynch's Seven Modes of Memetic Transmission" in the same indented format is not based on excerpts at all. There, the wording of each mode is Marsden's, and phrased so as make a mockery of what is actually said. Finding out what I actually said, however, is as easy as reading chapter 1 online. For example, it would have been simple for Marsden to quote me directly as saying "Any idea influencing its hosts to have more children than they would otherwise have exhibits *quantity parental* transmission." (p. 3) Instead, he misrepresents the label I give to a natural phenomenon as being an inane exhortation to "Have lots of children and use your position of authority to indoctrinate them and increase the prevalence of your beliefs." Marsden likewise misrepresents all of the other transmission modes, mostly as exhortations to the reader.

Under proselytic transmission Marsden even attributes to me (without page reference) a vague and incomplete statement that "A belief that spreads will spread ..." (ellipsis is Marsden's). This directly contradicts what the book actually says: "Proselytic thought contagion becomes self-limiting as host population growth diminishes the supply of nonhosts. Few nonhosts remain by the time the host population is a great majority since most have already converted by then. Without enough nonhosts, especially persuadable ones, the proselytizing cannot win many new adherents. This creates cycles in which successful proselytic movements lose momentum, setting the stage for renewed outbreaks of old movements and initial outbreaks of new movements" (p. 6). In other words, many beliefs that spread will therefore stop spreading or even go into decline. In order to find *Thought Contagion* a "spectacular failure" in putting memetics on a more serious footing in the sciences, Marsden not only misrepresents what it says, but also demands that the book serve a far broader function than it claims to serve. In the first sentence of the preface, I state that the book "... introduces a new branch of science dealing with ideas that program for their own retransmission ..." (p. vii) That is, ideas that gain propagation by influencing or manipulating adherent's idea-propagating behaviours. It is *not* offered as a general theory of cultural evolution, nor a general theory of social contagion. Even within the subject area of memes, the focus is not on the field in general, but rather, on the small but important subclass of memes that play a particularly active role in causing their own retransmission. For example, in my discussion of how my work relates to that of Durham (1991), I state "Yet Durham admits that in special circumstances, memes can depart from the mainline rout to prevalence. Such 'special case' memes are the ones I call *thought contagions*". (p. 25)

Some take a very expansive definition of the word "meme" and envision "memetics" as covering almost all topics in social science and social philosophy, for instance whether the self is an illusion (Blackmore 1999). For them, *Thought Contagion* may indeed be a disappointment. Yet the scientific community is by no means united behind Blackmore's vision of a universal memetics, as exemplified by a scathing review in *Nature* (Coyne 1999). This leaves a valuable

role for my own more modest and scientifically conservative agenda of demonstrating that there are a reasonable number of phenomena that really call out for the evolutionary replicator analysis of self-spreading brain-stored information. Like Rogers (1995), I narrow my focus to discuss only a subset of cultural evolutionary phenomena, and this does not call for a thorough survey of literature in the broader field as Marsden seems to imply.

Contrary to one of Marsden's misquotations, I do not claim that my brand of memetics is "the missing link" in the social sciences. The phrase "the missing link" does not appear anywhere in my book, and for good reason. *Thought Contagion* expressly emphasizes that memetics should *not* be taken as a replacement to existing social science. In the preface, it says "Thought contagion theory considers mainly the question of how ideas program for their own propagation, and the consequences for cultural evolution. This is an important and long overdue addition to social science, but it is hardly offered as a replacement for all existing social science" (p. ix). Thought contagion memetics is only called *a* missing link, and is *absolutely never* treated as the last or only "missing link" in the social sciences.

Attached to Marsden's false claim that I call my work "*the* missing link" comes another false assertion that my second chapter is presented as a "brief misrepresentation of the social sciences." The actual chapter title is "A Missing Link: Memetics and the Social Sciences." Not only does this contain the only use of the term "missing link" in the whole book, but it also clearly labels itself as something other than a survey. An honest reading of the chapter, however, reveals that it is presented as nothing more than an abbreviated overview of the relationship between thought contagion memetics to the social sciences.

As if distorting the overall purpose of the book and Chapter 2 were not enough, Marsden attempts to generate suspicion based on the 3% of that chapter given over to a discussion of "Memetics and Psychohistory". Says Marsden, "Fortunately, suspicion is likely to be aroused in even the most uninitiated reader when Lynch includes a discussion of the science fictional "psychohistory" in his brief misrepresentation of the various social sciences, in the process of claiming that his brand of memetics is 'the missing link.'" What he does not mention, however, is that the material actually emphasizes the *dissimilarities* between memetics and Asimov's popular science fiction. So I invite JASSS readers to see for themselves whether that discussion causes suspicion as Marsden suggests, or whether suspicion comes solely from the effect of Marsden's words on the uninitiated reader. To help JASSS readers decide where to place their suspicions, I copy the entire three short paragraphs alluding to "psychohistory" right here:

## **Memetics and Psychohistory**

Some have likened memetic history to the science fiction account of a theory called "psychohistory" in Isaac Asimov's *Foundation* series. Although psychohistory did not inspire memetics, the two theories do have surprising similarities. Both theories concern *how* history unfolds, and both give more consideration to the cumulative behavior of great masses than to the actions of special individuals. Both theories can cover hundreds of generations, and both can be translated into quantitative, mathematical equations. Like psychohistory, the memetic equations can even predict "future history" given some well-measured parameters and starting conditions.

Yet the analogies start to fade from that point on. In the science fiction story, a psychohistorian can predict most of society's behavior far into the future and quite precisely. Thought contagion theory mainly considers a special subclass of ideological behaviors, and just measuring the variables can raise serious practical challenges. The quantitative translation of the theory also leads to nonlinear equations, which mathematicians and meteorologists now see as a bane to long-range forecasting.

Thought contagion memetics might never amount to the stuff of science fiction, but it can make an important contribution to the understanding of history and the human condition. (*Thought Contagion*, p. 38-39)

There is in fact so much misrepresentation, distortion, misleading assertion and poor scholarship in Marsden's review that it would take an unreasonably long line-by-line critique to expose all of it. Instead, I caution anyone who reads any part of Marsden's review to read my own work directly and see what it really says, no matter how innocent, innocuous or plausible Marsden's statements look. I also invite readers to consider the possibility that if Marsden needed so much misrepresentation in order to produce his negative review, *Thought Contagion* may actually be quite good. Indeed, contrary to Marsden's statement that there is not "the slightest risk that anybody could possibly take it seriously," *Thought Contagion* has endorsements from Richard Dawkins and Douglas Hofstadter on its cover.

It is true that the full title of the book is "Thought Contagion: How Belief Spreads Through Society: The New Science of Memes." While this may seem to claim a far wider territory than the book actually covers, I should explain that the subtitle "How Belief Spreads Through Society" is only intended to identify the subject area in general terms to people who have never heard the word "meme." The line "The New Science of Memes" is only intended to place the book in the subject area of memetics for those who *have* already heard of memes. Yet the body of the book makes clear that thought contagion memetics is a limited subdomain of memetics generally (p. 25). That may trouble those readers who hold a grand vision of memetics as accomplishing everything for culture that Darwinism did for biology, but I remain committed predominantly to the specialised realm of thought contagion memetics.

As Marsden asserts, *Thought Contagion* is indeed filled with hypotheses awaiting empirical investigation and the book warns the reader of this in advance (p. viii-ix). In the preface, I express the hope that publishing hypotheses will stimulate the research needed to test them. Likewise, the publication of hypotheses about evolution, biological contagion, quarks and other phenomena has historically played an essential part in stimulating empirical studies.

In fields ranging from biology to physics, works written for non-specialists must often forego the most technical definitions of terms to achieve broader accessibility. Hence, *Thought Contagion* does not contain my most formal and precise definition of the word "meme." Yet Marsden, despite knowing my formal technical definition, seizes upon the limitations of a book written for general audiences to suggest an underlying imprecision and confusion in my use of the word. For instance, in his eighth paragraph, he refers without page reference to my statement "Because the information resides more at the community than the individual level, ownership qualifies as a kind of 'transcendent meme,' but not as an ordinary meme" (p. 19) to persuade readers that they should be "confused." In keeping with his overall pattern of presentation, Marsden distorts what I

said in a way that makes it look ridiculous: "transcendent meme" in my book becomes "*transcendental* meme" (emphasis added) in Marsden's review and out of context. (In context, the word "transcendent" is not the least bit mysterious: it simply indicates that ownership resides not in the individual, but in social entities such as communities and states that transcend the individual.) Marsden then goes on to announce that he was "no more enlightened" by reading my technical paper "Units, Events, and Dynamics in Memetic Evolution" (Lynch 1998). All of these suggestions of confusion, mystery and unenlightening exposition on my part calls for a presentation of the technical definition of "meme" from Lynch (1998):

"MEME: A memory item, or portion of an organism's neurally-stored information, identified using the abstraction system of the observer, whose instantiation depended critically on causation by prior instantiation of the same memory item in one or more other organisms' nervous systems. ('Sameness' of memory items is determined with respect to the above-mentioned abstraction system of the observer.)"

Removing the philosophy of science about abstractions (explained in the paper), this becomes more simply:

"MEME: A memory item, or portion of an organism's neurally-stored information, whose occurrence depended critically on causation by prior occurrence of the same memory item in one or more other organisms' nervous systems."

This definition identifies the minimum conditions needed to achieve the recursive process (or algorithm) that forms the basis of evolution by natural selection in interpersonally transmitted brain-stored information. It becomes the basis of transmission event diagrams and differential equations (Lynch 1998) useful in developing computer simulations of the meme transmission patterns discussed in the book *Thought Contagion*.

Filled as it is with misinformation about both the book *Thought Contagion* and the technical memetics paper cited in its preface, Marsden (1999b) fails spectacularly as a book review. What Marsden has "reviewed," is, in effect, his own series of misrepresentations. Those looking for fresh methods and topics in quantitative social simulation would therefore do better to ignore Marsden's contentious, misleading "review" and read *Thought Contagion* and its supporting technical papers directly.

## References

- BLACKMORE S. 1999. *The Meme Machine*, Oxford University Press, Oxford. [JASSS review.]
- COYNE J. 1999. The Self-Centred Meme: A Review of *The Meme Machine* by Susan Blackmore, *Nature*, 398:767.
- DURHAM W. 1991. *Coevolution: Genes, Culture, and Human Diversity*, Stanford University Press, Stanford, CA.

LYNCH A. 1991. Thought Contagion as Abstract Evolution, *Journal of Ideas*, 2:3-10. [Scanned reprint available online.]

LYNCH A. 1996. *Thought Contagion: How Belief Spreads Through Society: The New Science of Memes*, Basic Books, New York, NY. [Chapter 1 available online.]

LYNCH A. 1998. Units, Events and Dynamics in Memetic Evolution, *Journal of Memetics - Evolutionary Models of Information Transmission*, 2, <[http://www.cpm.mmu.ac.uk/jom-emit/1998/vol2/lynch\\_a.html](http://www.cpm.mmu.ac.uk/jom-emit/1998/vol2/lynch_a.html)>.

MARSDEN P. 1999a. A Strategy for Memetics: Memes as Strategies, *Journal of Memetics - Evolutionary Models of Information Transmission*, 3, <[http://www.cpm.mmu.ac.uk/jom-emit/1999/vol3/marsden\\_p.html](http://www.cpm.mmu.ac.uk/jom-emit/1999/vol3/marsden_p.html)>.

MARSDEN P. 1999b. Review of *Thought Contagion: How Belief Spreads Through Society* by Aaron Lynch, *Journal of Artificial Societies and Social Simulation*, 2, <<http://www.soc.surrey.ac.uk/JASSS/2/2/review4.html>>.

ROGERS E. M. 1995. *Diffusion of Innovations*, fourth edition, The Free Press, New York, NY.

# **SYNTACTIC STRUCTURE IN BIRDSONG: MEMETIC EVOLUTION OF SONGS OR GRAMMARS?**

**By**

**William Majoros**

Celera Genomics

45 West Gude Drive

Rockville, MD 20850, USA.

## **Abstract**

In order to ascertain whether the units of memetic transmission and recombination in the birdsong of a particular species of finch exist at the level of individual songs or instead at the level of grammar models, a large number of songs were subjected to grammar induction techniques and the resulting grammars analyzed for key structural properties considered indicative of the existence of an underlying grammar. The weight of evidence provided by this approach was judged by applying the same analyses to the results of a genetic algorithm applied to a synthetic repertoire of songs and comparing the results. Memetic evolution applied directly to song elements, as simulated by the genetic algorithm, was found to generate models very similar to those inferred from the House Finch song, thereby demonstrating that the problem of discerning between these two competing hypotheses for explaining the syntactic structure observed in House Finch song is not easily solved using the limited data obtainable in the field. This underscores the importance of experimental rigor when studying memetic systems, and leaves open the question of how to confidently determine the actual level at which memes are operating in a particular system.

## **Introduction**

Beginning with an informal description of House Finch song and a brief review of the studies of that species' vocal behavior to date, we attempt to motivate the question of whether the units of memetic transmission and recombination in this species exist at the level of individual songs, or whether they may instead take a more abstract form, at the level of grammar rules restricting the allowable sequences of syllables within individual songs (section 1). We further point out that selective pressures for the evolutionary emergence of grammar models in bird song are entirely conceivable, based on arguments involving female preference and the advantages of data compression (section 2). We then describe the justification for the experimental technique employed in this study (section 3) and we describe that technique in detail (section 4). After presenting the results of the experiments (section 5), we attempt to interpret those results and provide motivation for further studies (section 6). This is followed by a brief conclusion (section 7).

### **1. Anecdotal Evidence for Syntactic Structure in House Finch Song**

Much of the literature on bird song emphasizes the importance of individual songs, both in the production of acoustic behavior and in the transmission of cultural information via that behavior. Implicit in this song-oriented conception of avian vocal behavior is apparently an assumption

that the underlying representation for that behavior in the bird's brain is in fact some suitable neural encoding for the individual songs which can be produced by that bird; i.e., a simple time-ordered enumeration of the syllables comprising a song. This impression is given by the frequent use of the term "repertoire" by the authors in this field. The usual (but perhaps unfortunate) connotation of this term is of a set of songs each of which has been memorized in completeness and which is performed nearly identically every time.

It is reasonable to inquire, however, whether singing behavior may have a more algorithmic basis than the simple recitation of fixed songs. In particular, linguists have devised a standard hierarchy of so-called "grammars," called the Chomsky hierarchy [8], which classifies increasingly sophisticated methods for the generation of strings of symbols [note 1]. For species exhibiting considerable complexity in the structure of their songs, it is tempting to speculate on the possible existence of grammar-like mechanisms in their behavioral programming [note 2].

The House Finch (*Carpodacus mexicanus*) is clearly one of those species. Various authors have noted the complex and variable syntactic features of the song of this Fringillid. Mundinger [12] observed that in one region of his study area, "each male sang a wide variety of patterns, and individual variation in a given pattern was high." He also pointed out that "the precise mechanism of song learning in the wild is unknown" [note 3]. Bitterbaum and Baptista [1], working in California, noticed that "not only do no two birds sing exactly alike, but even a single bird seldom repeats himself precisely." Pytte [14] likewise observed in Wisconsin that "rarely is the same song repeated exactly" and concluded that "song types reported in this study appear to be largely an artificial construct."

Various authors have reported having the strong impression that many songs had been formed through the recombination of portions of other songs ([1, 14, 16, 18]). Tracy and Baker [18] speculate that "birds may be memorizing song components (syllables or phrases) that are strung together into complete songs at a later time" and Pytte [14], in relating the work of Bitterbaum and Baptista [1], explains that "juveniles learn individual syllables as units, which they segregate and recombine independently during song production."

However, these statements apparently refer only to subsong production by juveniles. As Lynch and Baker [10] stated more explicitly regarding recombination of song elements in the Chaffinch (*Fringilla coelebs*), "As song development proceeds, songs become less variable; by the time they reach 'crystallization,' they are highly stereotyped, and hybrid songs tend to disappear from the birds' repertoire."

Smith [16], on the other hand, offered the following intriguing proposal: "It is possible that house finches do not even choose songs or song elements, but choose a model(s) to imitate. Variation would then come from the 'mutations' of the model." Although he provides no further elaboration on the type of model that he has in mind, the fact that his work involves the construction of Markov chains from House Finch songs is highly suggestive; Markov chains are a simple form of grammar model [note 4]. However, Smith's models were inferred not from the songs of individual birds, but from the songs of all the birds in the local population in a given year, so it is not clear that these models are intended to describe anything other than the overall structure of the local dialect. Furthermore, the models were inferred in a highly subjective way; phrases do

not appear to have been selected systematically, and "hybrid" phrases were excluded from the model.

## 2. Evolutionary Advantages of Grammar-Based Song Generation

It is unfortunate that the hypothesis of grammar-based song generation has received only modest attention in the literature. In addition to providing a proximate explanation for the complexity observed in the song of some species, such models also suggest an ultimate explanation in those species where females have been shown to prefer males with larger repertoires; a grammar model with an appropriate cyclic structure can generate a potentially infinite number of distinct strings, though the model itself may have a fairly compact representation [note 6]. Thus, they provide a form of data compression, which natural selection may favor were such a mechanism to arise [20]. Such models also have interesting implications for the form and rate of cultural evolution in systems utilizing them, because the accurate induction of grammars from only positive example strings is known to be combinatorically difficult [4] [note 7].

## 3. Overview of the Experimental Approach

For the above reasons, I decided to apply grammar induction techniques to a large sample of House Finch song in order to investigate systematically the syntactic properties of this species' song, and to test the hypothesis that adult House Finches actually utilize such models while singing.

In order to test this hypothesis, I further posited that memetic evolution applied directly to songs represented as fixed syllable sequences would be incapable of generating certain statistical properties expected to occur in actual House Finch song. In particular, I posited that contingency tables constructed from pairs of states in the inferred grammar models would provide support for independence between the choices made at those pairs of states as the birds transitioned out of those respective states during song generation. Intuitively, one would not expect such independence to obtain in grammar models inferred from songs that had instead been generated via simple string recombination, because there is no obvious reason to expect random crossover between songs to respect phrase boundaries in such a way as to maintain the symmetry necessary for independence at pairs of states in the corresponding inferred grammar.

Although such independence may be expected to occur in extreme cases where extensive recombination has produced nearly all possible syllable combinations, such an extensive shuffling of syllables would seem to obliterate certain forms of order observed in actual House Finch song. For example, the models inferred in Smith [16], as well as unpublished analyses of data from Pytte [14] performed by myself reveal persistent multi-syllable themes that one would not expect to survive in the presence of high levels of random crossover.

## **4. Methods**

### **4.1 Sound Recording**

Recording took place between February 18<sup>th</sup> and July 12<sup>th</sup>, 1998, and was conducted in eighteen distinct suburban neighborhoods in Pennsylvania and Maryland, USA.

Recording equipment consisted of a 31-inch Telinga Pro Universal parabolic reflector fitted with a Sennheiser ME62 microphone, and a Marantz PMD222 portable cassette recorder. Songs were recorded onto Maxell high-bias ( $\text{CrO}_2$ ) cassettes in the field and then digitized using 18-bit A/D conversion hardware and written to compact disk (CD-R).

Male House Finches were located by walking along a road and listening for song. When a singing finch was encountered, recording of that bird began immediately and continued until the bird ceased singing, the end of the current tape was reached, or another singing House Finch intruded upon the session. Thus, a single recording on the tape consisted of song from only one bird. Although the same bird may be represented in multiple recordings taken at a single site, this will not affect the grammar induction process, which was performed only within individual recordings.

### **4.2 Selection and Transcription of Recordings**

Each session was indexed according to media location and recording quality. Quality values consisted of a subjective assessment of recording clarity, with an emphasis on the expected sonogram quality. The 263 recordings in the highest quality category were then selected for analysis.

Individual songs in the selected recordings were separated based on a subjective assessment of inter-song delay. House Finch songs are generally separated in time by a delay much longer than that between the individual syllables of a song (pers. obs.), so this process is largely unambiguous and replicable. Each song was stored in a separate file in ".wav" format. Recordings consisting of fewer than fifteen songs were omitted from the analysis.

Sonograms were produced using the Spectrogram 2.3 software [9] and printed out on a 1440 dpi photo-quality printer for manual analysis. Individual syllables were subjectively identified by spacing. Although a syllable often consisted of a single "trace" in the sonogram, some syllables consisted of multiple traces which were positioned very close together (in time) and which consistently occurred together in all sonograms. Thus, the process of separating syllables is rather more subjective than that of identifying individual songs, but a relatively high degree of reproducibility should be expected nonetheless.

Syllables were compiled into site-specific "alphabets," with a single alphabet being associated with all of the birds recorded at a single site. Each syllable was assigned a unique symbol, and songs were then translated into sequences of these symbols. These steps were all performed by a single person, in order to eliminate any inter-observer differences.

### 4.3 Grammar Induction

All of the songs produced by a single bird during a single, continuous recording session were used to infer a grammar model for that bird. The class of grammar model chosen for this study was the Moore machine, a type of finite automaton [8]. A Moore machine consists of a set of labeled states and transitions between those states, with the state labels prescribing the output of the machine upon entering that state. Moore machines are similar to hidden Markov models [3] except that there are no explicit transition probabilities, and each state has a label with an associated emission probability of 1 [note 5].

The first step in inferring a model was to identify common subsequences with which to label the states. This was accomplished by computing a multiple alignment between songs [note 8], using techniques developed for molecular biology. Beginning with the first pair of songs in the session, pairwise alignments were computed using the Needleman-Wunsch global alignment algorithm [3] between each successive song and the overall consensus sequence of the foregoing alignments, with individual syllables being treated as *residues* for the purpose of alignment. The alignment algorithm was parameterized as follows: gap open penalty = 1.5, gap extension penalty = 0.66, match score = 100, and mismatch penalty = 1000. This parameterization effectively precluded the alignment of nonidentical syllables [note 9].

Gaps in the alignment were used to infer breaks in the consensus sequence, with a break being established immediately left of the beginning of each gap and immediately right of the end of a gap. The consensus sequence was then segmented into the substrings delimited by these breaks, with an implicit break occurring at the beginning and end of the sequence. Each such substring was then assigned as the label of a state, with identically labeled states being merged into a single state.

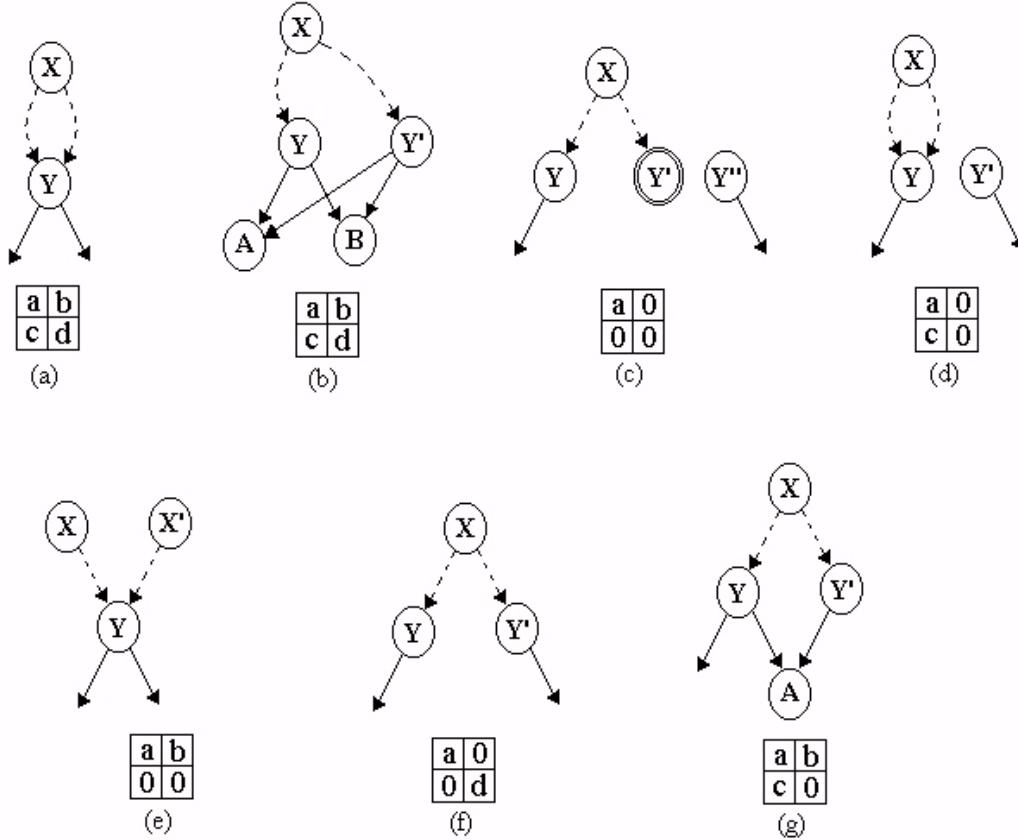
Because of the manner in which breaks were inferred, songs could be unambiguously recoded as sequences of states, based on the positions of their constituent symbols in the multiple alignment relative to the breaks. These state sequences were then used to establish the state transitions by observing which states immediately followed which others in the recoded songs, much the way one goes about training a Markov chain [3]. Start states and final states were identified by their occurrence at the beginnings and ends of songs, respectively.

Grammar models were visualized using the daVinci program [19].

### 4.4 Contingency Table Analysis

A basic assumption of the class of grammar models employed is that the choice of transition made while in one state is independent of the choice of transition made in any subsequent state during the generation of a song. Complete failure of syntactic structure in House Finch song to exhibit any such independence would preclude generation of song according to this class of grammar models, but occasional violations of independence may indicate only that distinct states had been erroneously merged during the grammar induction process. Thus, the results of any independence tests that we apply must be interpreted within the wider context involving these imperfect model induction techniques.

In order to test for independence between the choices of transition made at pairs of states within individual songs, a contingency table was constructed for each such pair of states. In particular, for each pair of states  $\langle X, Y \rangle$  such that  $Y$  follows  $X$  (potentially after some number of intervening states) in at least one song and such that both  $X$  and  $Y$  offer transitions to at least two successor states (fig. 1a), a  $2 \times 2$  contingency table was constructed by tabulating the occurrences of the four combinations of transitions leaving  $X$  and  $Y$  ( $\langle \text{left}, \text{left} \rangle$ ,  $\langle \text{left}, \text{right} \rangle$ ,  $\langle \text{right}, \text{left} \rangle$ ,  $\langle \text{right}, \text{right} \rangle$ ) in all of the songs from the current recording session (i.e., for a single bird). In cases where more than two transitions left a state, all pairs of outgoing transitions from that state were considered separately, so that only  $2 \times 2$  contingency tables were constructed. As an additional constraint on which of these state pairs would be considered, it was required that state  $Y$  be reachable from state  $X$  by at least two different paths.



**Figure 1:** (a) Contingency tables are naively assumed to arise from the depicted topology. Solid lines are transitions and broken lines indicate arbitrarily long paths. (b) A four-entry table could result from the merging of identically labeled states  $Y$  and  $Y'$ , but only if both have transitions to  $A$  and  $B$ . (c) A single-entry table could result from the merging of states  $Y$ ,  $Y'$ , and  $Y''$ , where  $Y''$  is a final state. (d) One form of two-entry table can be formed by merging  $Y$  and  $Y'$ . (e) Another form of two-entry table can be formed by merging  $X$  and  $X'$ . (f) Yet another form of two-entry table can be formed by merging  $Y$  and  $Y'$  in the topology shown. (g) A three-entry table can be formed by merging  $Y$  and  $Y'$ , but only if both have transitions to  $A$ .

Because it was supposed that a bird might perform a song several times in succession after generating it, and because my interest was in the generation of songs rather than their performance, any song which was identical to the song immediately preceding it in time was omitted from this part of the analysis.

Probabilities were assigned to contingency tables using the Fisher Exact Test [21] under the assumption of independence. Because a potentially large number  $N$  of such tests were to be conducted, an adjusted critical value was computed as follows:

$$1 - (1 - \alpha)^N = 0.05,$$

or, equivalently,

$$\alpha = 1 - 0.95^{1/N}.$$

Using this computed  $\alpha$  as the critical value sets the probability of committing at least one Type I error during the  $N$  tests to 5%.

Any particular statistically significant table is evidence only of a single instance of erroneous state-merging during model induction. A large number of significant tables would suggest that no state merging is appropriate, thereby seriously calling into question the applicability of this class of models. As previously mentioned, a small number of significant tables suggests only that the grammar induction algorithm commits occasional state-merging errors.

However, tables with four nonzero entries would seem to constitute positive evidence for the validity of this class of model, because a spurious four-entry table (regardless of P-value) would require not only erroneous state merging, but also the coincidental matching of transitions leaving such erroneously merged states (fig. 1b). Such a high level of coincidence is not required for the spurious generation of tables having one or more zero entries (fig. 1c-1g), with the possible exception of single-entry tables (fig. 1c). However, single-entry tables were deemed less informative than four-entry tables for the current application, because statistically nonsignificant single-entry tables tend to have very small entry sums, and they are therefore indicators primarily of inadequate sample size. Obviously, statistically significant tables with any number of nonzero entries were considered informative.

A large number of four-entry tables would therefore seem to be highly unlikely without the existence of an underlying grammar model of the type proposed, unless those tables were judged statistically significant by the Fisher Exact Test.

Unfortunately, the complexities of these grammar models make reliable estimation of probabilities for various types of grammar induction errors exceedingly difficult without adequate sample sizes. Because the number of possible paths through a model grows exponentially with the number of transitions, sample sizes obtained in practice can be expected to be woefully inadequate; therefore, a rigorous analytical treatment was not deemed practical.

My approach to compensating for this lack of rigor was to apply the same contingency table analysis to sets of simulated songs generated without the use of a grammar model (see below); any support for the existence of a grammar model obtained from this simulated data would nullify the value of any corresponding evidence found in the House Finch data.

#### 4.5 Simulation of Memetic Evolution of Songs

In order to investigate whether memetic evolution applied directly to songs (rather than to grammar models) could generate syntactic structure similar to that observed in House Finch song, a steady-state genetic algorithm was applied to strings of artificial syllables. Genetic algorithms are described in detail in [6].

Briefly, an initial population of 14 to 25 strings of 13 to 24 unique symbols drawn from nonoverlapping alphabets was subjected to between 14 and 40 generations of panmictic and selectionless evolution. During each generation, 15 new strings were generated by applying two-point crossover to pairs of strings selected uniformly at random from the current population. Each new string then replaced an existing string at random. A two-point crossover operation combines the beginning and ending portion of one parent string with the middle portion of another parent string, to produce a new child string. The crossover points are selected at random. Note that differently-sized parents were allowed to recombine, and crossover points were allowed to vary between the two parents, thereby facilitating the creation of repetitive elements and further variation in string length. The final population at the end of each run was then subjected to the grammatical inference algorithm described above to infer a grammar model which was in turn subjected to contingency table analysis. Several thousand runs were performed under varying parameterizations until several grammar models were obtained which adequately matched those inferred from House Finch song according to various criteria.

In particular, the following quantities were required to differ by no more than one (sample) standard deviation from the expected values established by the House Finch data: number of states, number of transitions, transition density, mean state label length, mean in-degree, mean out-degree, number of start states, number of final states, number of cycles, mean cycle length, number of unique syllables, and mean song length. In-degree and out-degree are the numbers of transitions entering and leaving a state, respectively. Transition density is the proportion of combinatorically possible transitions realized. Cycles were identified by observing back-edges during a depth-first-search originating at each start state [2]; no attempt was made to identify composite cycles, so cycle sizes and counts are somewhat underestimated.

To summarize the intent of these experiments, my prediction was that four-entry contingency tables would be common in the House Finch data, with no more than a few of these being statistically significant, thereby demonstrating independence between transition choices made at different states during the generation of a song, and thus supporting the possibility that House Finches actually use some form of grammar model to generate their songs anew at each performance. In contrast, I predicted that the simulated data would produce either very few four-entry tables, or that most of these would be statistically significant, thereby supporting the intuition that syntactic structure of the type observed in House Finch songs would be highly

unlikely to occur by chance if House Finches were simply reciting fixed songs that changed only through transmission "errors" during cultural evolution.

#### 4.6 Regression Analysis

Least-squares regression was used to fit a model of the form  $y = mx+b$  to the number of transitions in a grammar as a function of the number of states. This analysis was applied to the grammar models induced from the field data and separately to those derived from the simulated data. Coefficients of determination ( $r^2$ ) were computed to assess the quality of the regressed line, and t-tests were used to determine statistical significance. Regression analysis was performed using SPSS for Windows Release 8.0.0 [17]. Except where noted, all analyses were performed using unpublished software written by the author. Means are given  $\pm$  SD.

### 5. Results

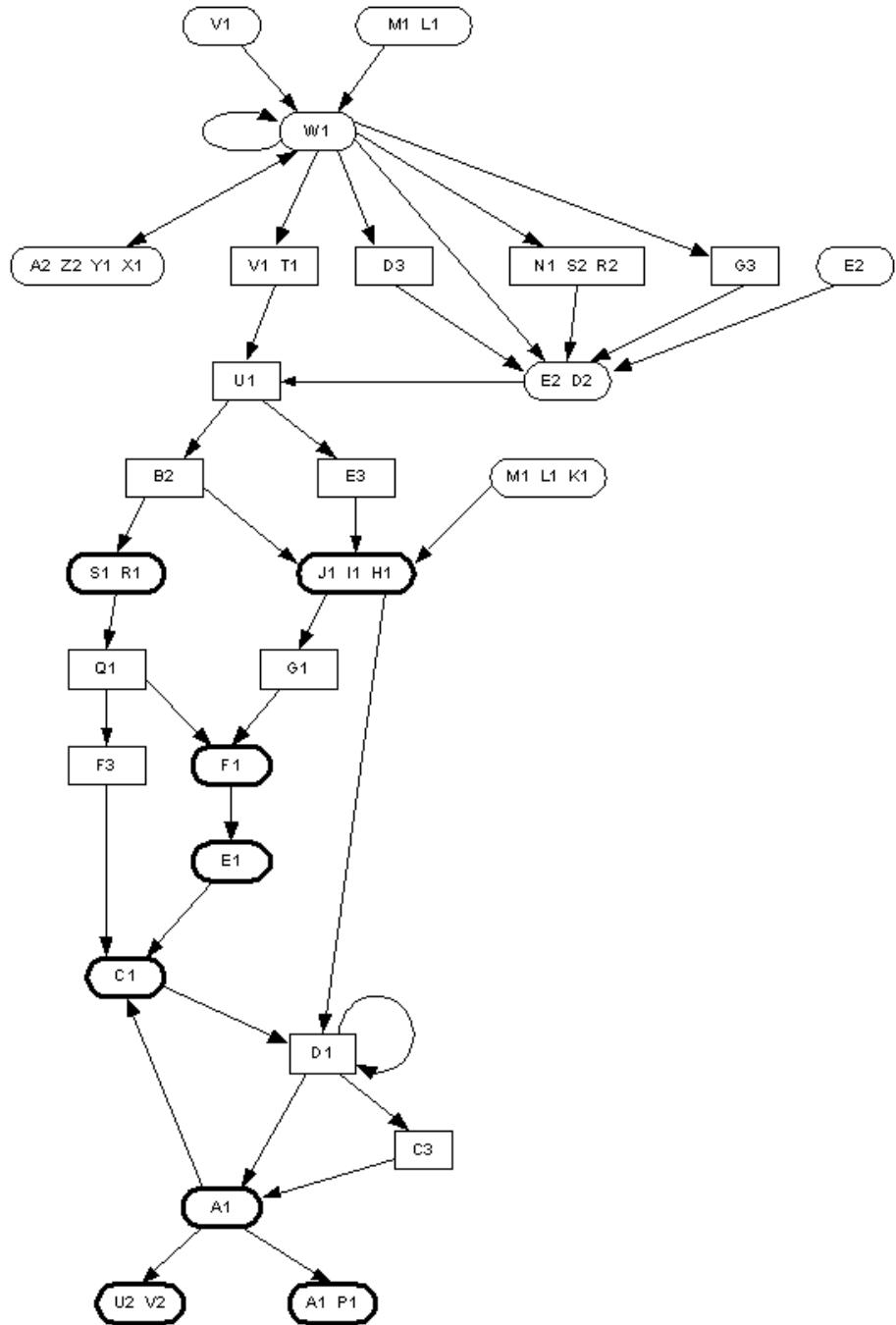
#### 5.1 Properties of House Finch Data

The 44 birds used for the analysis each sang 15 to 104 songs (mean  $27.3 \pm 19.4$ ), for a total of 1,172 songs. Table 1 provides summary statistics for various other quantities. *State sequence length* refers to the length of a song when recoded as a sequence of states. Sample standard deviation is denoted by s.d. *Transition density* is the number of transitions divided by the total number possible in a complete graph. *In-degree* and *out-degree* are the numbers of transitions entering and leaving a state, respectively.

Variable	mean	s.d.	min	max
number of states	28.21	14.38	9	75
number of transitions	32.47	17.7	8	88
transition density	0.10	0.04	0.03	0.22
song length	14.64	4.91	4	47
state sequence length	9.23	4.03	1	34
state label length	1.89	1.62	1	13
number of songs per bird	27.3	19.4	15	104
unique syllables per bird	39.28	15.29	21	97
number of cycles	4.16	4.41	0	17
cycle length	5.22	4.82	1	21
number of start states	4.79	2.60	2	12
number of final states	10.30	4.03	4	21
in-degree	1.47	5.86	0	69
out-degree	1.47	1.02	0	7

**Table 1:** Summary statistics for House Finch data. Terms are defined in the text.

An example model inferred from the field data is shown in figure 2. There was considerable diversity in the overall structure of the topologies observed. In particular, not all of the models were long and narrow like that shown in figure 2, some had states with longer labels, and some had larger cycles than shown in the example.



**Figure 2:** An example of a grammar model inferred from songs produced by a House Finch. Rectangles denote states, thin-lined ellipses denote start states, and thick-lined ellipses denote final states. Transitions between states are shown as directed arrows.

A total of 3,282 contingency tables were constructed from the inferred models. However, 3,176 of these (97%) contained two or more zero entries, and only four tables (0.12%) had four nonzero entries. These were:

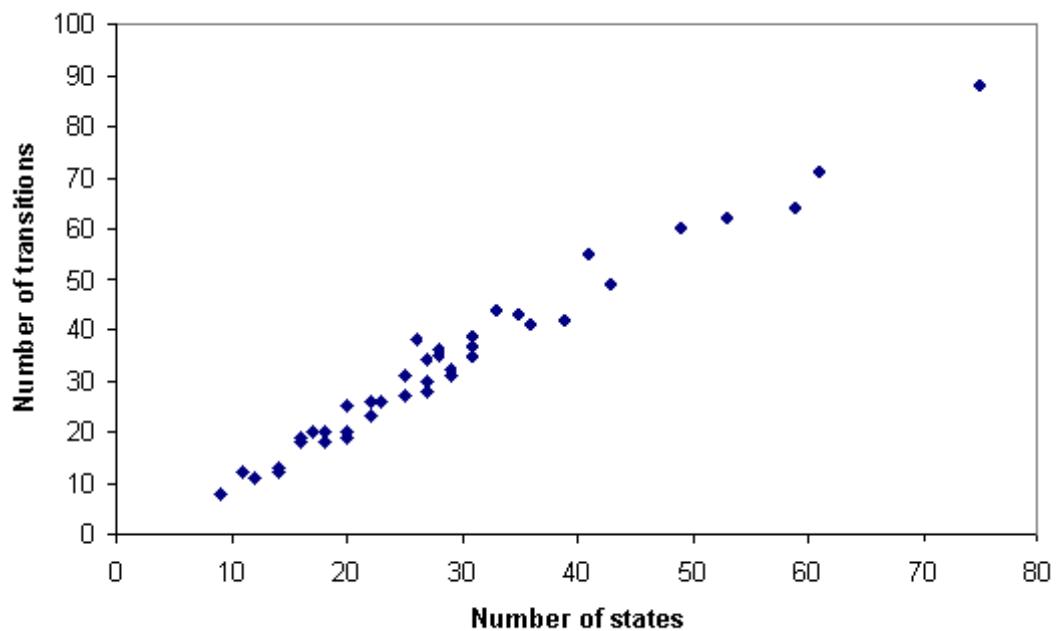
<table border="1"><tr><td>1</td><td>1</td></tr><tr><td>1</td><td>3</td></tr></table>	1	1	1	3	<table border="1"><tr><td>1</td><td>1</td></tr><tr><td>2</td><td>1</td></tr></table>	1	1	2	1	<table border="1"><tr><td>31</td><td>13</td></tr><tr><td>1</td><td>2</td></tr></table>	31	13	1	2	<table border="1"><tr><td>25</td><td>12</td></tr><tr><td>3</td><td>2</td></tr></table>	25	12	3	2
1	1																		
1	3																		
1	1																		
2	1																		
31	13																		
1	2																		
25	12																		
3	2																		
P=0.53	P=0.60	P=0.207	P=0.35																

Of the 3,282 tables, only one (0.03%) had a P-value less than 0.05. This was:

<table border="1"><tr><td>0</td><td>12</td></tr><tr><td>5</td><td>3</td></tr></table>	0	12	5	3
0	12			
5	3			
P=0.0036				

However, due to the large number of tests conducted, an event with this probability could be expected to occur several times; the adjusted critical value given N=3,282 is 0.000016. Even if tables that extreme could not possibly be obtained with the given sample size, the fact that far fewer than 5% of the tables had a P-value of less than 0.05 supports the decision not to reject the assumption of independence.

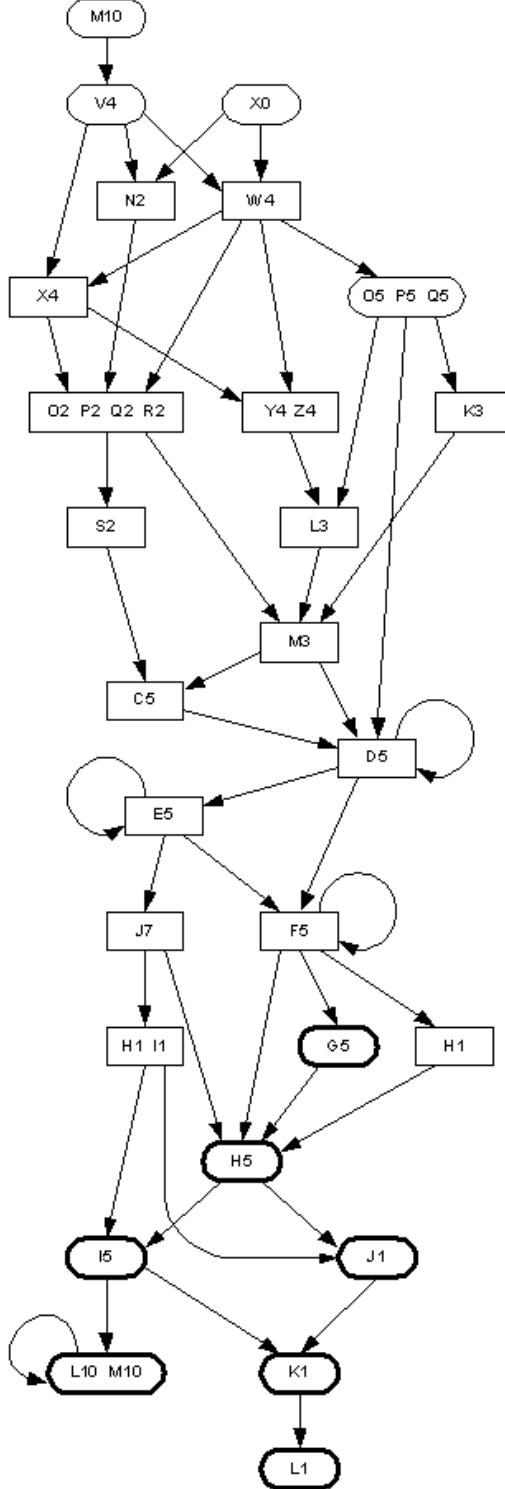
The number of transitions in the inferred grammar models was found to vary linearly with the number of states ( $m=1.22$ ,  $b=-1.81$ ,  $r^2=0.97$ ;  $t=38.816$ ,  $P<0.001$ ), as shown in figure 3. Because the number of transitions combinatorically possible between  $n$  states varies with  $n^2$ , this linear relationship places appreciable constraints on the topological structure of the models.



**Figure 3:** Number of transitions (vertical axis) as a function of the number of states (horizontal axis) in grammar models inferred from House Finch songs.

## 5.2 Properties of Simulated Data

The models inferred from the simulated data appeared remarkably similar to those inferred from the House Finch songs. An example is shown in figure 4, which can be seen to be qualitatively like figure 2 in many respects. All of the relevant quantities for this run fell within one standard deviation of the mean values for the corresponding House Finch data, as shown in table 2.



**Figure 4:** An example of a grammar model inferred from simulated songs produced by a genetic algorithm.

<b>Variable</b>	<b>field data</b>	<b>simulation</b>
number of states	28.21	27
number of transitions	32.47	48
transition density	0.1	0.07
song length	14.64	13.79
state sequence length	9.23	11.50
state label length	1.89	1.35
number of songs per bird	27.3	14
unique syllables per bird	39.28	33
number of cycles	4.16	4
cycle length	5.22	1
number of start states	4.79	3
number of final states	10.3	7
in-degree	1.47	1.78
out-degree	1.47	1.78

**Table 2:** Results of an example run of the genetic algorithm.

This particular run produced a single four-entry contingency table:

A total of approximately 300,000 contingency tables were obtained during approximately 5,200 runs. Of these tables, 3,128 (1.04%) had four nonzero entries, and only 47 (0.02%) were significant at the 0.05 level. Thus, the simulations produced proportionally more four-entry tables and proportionally fewer statistically significant tables than analyses of the House Finch data.

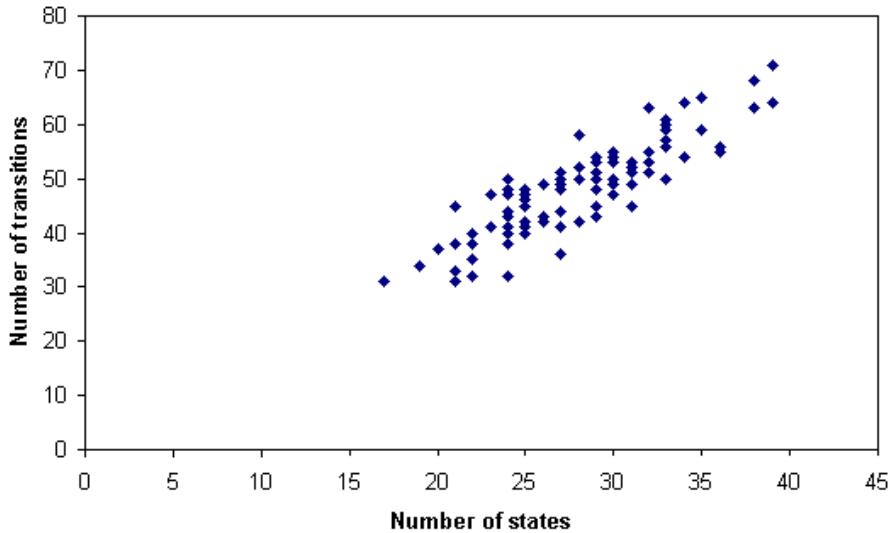
As with the field data, the models inferred from simulated data exhibited a wider variety of topologies than can be communicated by a single example such as that given in figure 4. Like the House Finch songs, the simulated data gave rise to topologies with varying numbers and sizes of cycles, varying numbers of states and transitions, and clearly discernible variation in overall structure.

Table 3 gives the mean values of the previously-identified key quantities from a sample of 1000 runs. With the exception of the number of final states, all of the values are within one standard deviation of the mean values observed in the field data. Although this numeric agreement resulted from a conscious effort to match these quantities to the field data via refinement of the parameters to the genetic algorithm, the fact that such agreement could be achieved between these two systems is notable.

variable	field data	simulation
number of states	28.21	28.31
number of transitions	32.47	48.49
transition density	0.10	0.06
song length	14.64	15.04
state sequence length	9.23	12.92
state label length	1.89	1.32
number of songs per bird	27.30	19.63
unique syllables per bird	39.28	32.11
number of cycles	4.16	3.79
cycle length	5.22	1.58
number of start states	4.79	2.41
number of final states	10.30	4.85
in-degree	1.47	1.72
out-degree	1.47	1.72

**Table 3:** Mean results for 1000 runs of the genetic algorithm.

As with the field data, the number of transitions was found to vary linearly with the number of states ( $m=1.57$ ,  $b=4.19$ ,  $r^2=0.75$ ;  $t=17.240$ ,  $P<0.001$ ), as shown in figure 5.



**Figure 5:** Number of transitions (vertical axis) as a function of the number of states (horizontal axis) in grammar models inferred from simulated data.

## 6. Discussion

Contrary to my predictions, simple string recombination was found to be capable of generating structures highly similar to those observed in House Finch song. Although considerable diversity was observed in both the models inferred from the field data and those inferred from the simulated data, suitable parameterizations were readily found for the genetic algorithm which would constrain various key measurements of structural complexity to within one standard deviation of the mean values obtained from the field data. Furthermore, the contingency table analysis failed to uncover more convincing evidence for transition independence in the House Finch data than in the simulated songs; indeed, the trend was in the opposite direction, with the genetic algorithm producing proportionally more four-entry tables (1.04% vs 0.12%) and proportionally fewer tables significant at the 0.05 level (0.02% vs. 0.03%) than the field data.

Lacking clear and compelling evidence for the existence of grammar rules in House Finch singing behavior, one might be tempted to invoke the principle of parsimony to establish rote memorization as the preferred hypothesis. However, it is not clear that either of the two hypotheses considered here (rote memorization vs. grammatically constrained improvisation) is significantly more parsimonious than the other. Although rote memorization would appear to be a simpler task than grammar induction, the potential for data compression afforded by a grammar model can lead to significantly greater space efficiency. Thus, while grammar models may be more complex in acquisition and operation, they are generally less complex in memory requirements and might therefore demand preference under a Minimum Description Length philosophy [15].

Thus, the question of whether House Finches generate their song using some form of grammar model remains unanswered, and it is questionable whether data collected in the field can definitively resolve the issue. What is likely needed is a set of controlled laboratory experiments, where large samples can be obtained from each bird and where training material can be specifically designed to create various informative testing scenarios.

An important point which has heretofore not been emphasized is that the hypothesis being considered is not simply that House Finches utilize some grammar model in generating their song, but that they specifically adopt a model which accounts for patterns that they identify in training songs. Thus, it is conceivable that controlled experiments might be designed which were capable of detecting a bird's ability to identify and recreate syntactic structures in the songs of an artificial tutor. The task is essentially one of distinguishing between concept learning versus rote memorization in a complex adaptive system.

Failure to foresee the results of the memetic simulation is perhaps attributable to the interaction between the shuffling effects of random crossover and the propensity of sampling error to create spurious order. The same susceptibility can be expected whenever a grammar induction algorithm is applied to moderately sized samples. Even for a collection of strings known to have been generated using a grammar model, one should not expect a black-box algorithm to reconstruct that model with complete accuracy in even the majority of cases. Not only are sample sizes generally limited in practice by the difficulties of working in the field, but grammar

induction algorithms are truly only heuristics, and are guaranteed to correctly reconstruct the original grammar only under special circumstances ([4, 5, 13]), which may not obtain in practice. If it can be demonstrated that House Finches do indeed employ syntax rules in the generation of their song, then a next step would be to characterize the range of complexity possible in those rules. This would require analysis of bird song from a diverse range of geographical locations, because the structural properties of song recorded in one locale can be highly sensitive to the evolutionary history of the local song tradition. That significantly different conclusions could be drawn from different populations of House Finches can be clearly seen by comparing the results of research conducted in different regions of the U.S. (Mundinger [12] in New York and Connecticut, Bitterbaum and Baptista [1] in California, Pytte [14] in Wisconsin, and Tracy and Baker [18] in Colorado).

Alternatively, if direct song recombination is ultimately established as the mechanism behind the evolution of, and apparent order observed, in House Finch song, then it is conceivable that syntax analyses may still be useful in elucidating some of the parameters of that evolutionary process. Furthermore, the recombination mechanism need not be as simple as the two-point crossover approach used here. The details of this recombination remain to be determined. All of this would require large-scale analyses.

There does not appear to be any logical way to distinguish between the use of grammar models by juveniles which later crystallize their song into fixed syllable strings versus the direct use of syntax rules by singing adults without resorting to invasive techniques such as neuronal monitoring of singing birds, and it is doubtful that current technology and neuroanatomical knowledge are adequate for such an undertaking [11].

Finally, Lynch and Baker [10] predicted that a "general theory of cultural evolution in bird song" would eventually emerge. Before this can happen, the actual nature of memes in bird song must be determined. Lynch and Baker [10] defined memes as syllables and sequences of syllables, but they did not consider higher-level structural elements, such as state transitions or syntax rules. Only a greater understanding of bird song generation will determine whether such extensions to their quantitative theory of cultural evolution are necessary. As stated recently by Greene [7], "We currently lack a coherent conceptual model that explains the diversity of singing behavior and song learning programs in an evolutionary framework."

## 7. Conclusion

Statistical analyses were conducted to test the hypothesis that House Finches generate their song using a simple form of grammar model similar to a hidden Markov model. Results failed to convincingly show that such models are necessary to explain the syntactic structure observed in House Finch song. However, it was argued that this hypothesis cannot be dismissed simply on the grounds of parsimony.

Although song-centric models are perhaps the most natural and convenient way for human researchers to conceptualize vocal behavior in birds, until we have a more thorough understanding of song learning and performance in species having complex song, more sophisticated models of song structure such as those considered here should be viewed as valid

avenues of investigation. Given the potential complexity of these models, addressing such important issues as the selective advantage they confer, their impact on the form and rate of cultural evolution, and the bias they exert on observed syntactic structure in bird song should make for challenging and fascinating topics of future research.

## Acknowledgments

Carolyn Pytte kindly provided sequence data for preliminary analyses which inspired this study, and provided copies of several relevant papers. Mark Yandell and Daniel Edelstein provided valuable comments on a previous version of the manuscript, as did the anonymous reviewers; I am indebted to them all for their suggested improvements.

The views expressed in this paper are mine alone, and not necessarily those of my employer.

## Notes

Grammar models generally consist of an *alphabet* of symbols and a mathematical construct specifying the way in which those symbols can be combined to generate *strings*, or finite symbol sequences. For birdsong, we can take the individual syllables comprising a song to be the symbols and the entire song to be the generated string. Two prominent classes of grammars in the Chomsky hierarchy are the *regular grammars* and the *context-free grammars*. Context-free grammars are the more powerful of the two, and involve the use of *productions*, or *rewrite rules* (such as "X can be replaced by ABC") that may be recursively applied to transform an initial string of variables into the final observed form of the string. Context-free grammars are in general implicated in processes that generate matching pairs of symbols, and palindrome-like patterns. Regular grammars, which constitute a proper (and less powerful) subset of the context-free grammars, are equivalent in their computing power both to *regular expressions* and *finite automata* (also called *finite state machines*), which consist of *states* and *transitions*, and which are often depicted as small circles with arrows connecting them into allowable paths.

This is in no way meant to imply, however, that all, or even most, bird species utilize grammars for song generation. Many species have song which is too simple or too stereotyped to suggest utilization of a grammatically-constrained, stochastic song generation procedure. On the other hand, it is not unreasonable to consider grammar-like mechanisms that may be at play as songs change over evolutionary time, or to consider the use of grammatical theories to describe the syntactic structure of songs produced by birds which are not themselves using grammars to generate their song. However, this work is concerned only with the possible use by birds (House Finches in particular) of grammar models during singing behavior.

Mundinger [12] refers to an unpublished manuscript that deals with song learning in this species, but the manuscript in question actually describes experiments involving the training of House Finches in the laboratory by another species of bird [pers. communication], and I know of no other work to date that examines song learning in this species, either in a controlled setting or in the wild.

Markov chains and hidden Markov models (*HMM's*) are generalized forms of stochastic finite automata which provide explicit *transition probabilities* and (in the case of HMM's) *emission probabilities*. The model begins in the start state and moves at random from state to state, according to the transition probabilities, emitting a symbol or string of symbols upon entering each state, until a final state is reached. In this way, the model generates a string consisting of the concatenated symbols emitted by the states that were visited during operation of the model. In HMM's, the emission probabilities of a state determine what symbol is emitted when that state is entered, whereas in a Markov chain, there is only one choice of symbol for each state, and the symbols labeling different states are unique.

A Moore machine operates much like a Markov chain. The machine starts out in one designated *start state*, transitions stochastically from state to state according to the allowable transitions, and emits the symbol or string of symbols labeling each state when that state is entered. The computation can end only when the machine is in one of the designated *final states* (although transitions out of final states are permitted), and the result of the computation is taken to be the concatenated labels of the states visited during operation of the machine.

In general, the more compact models tend to generate larger, less constrained, sets of strings, so if natural selection were to favor either higher-entropy signals or more compact models, one might expect the ultimate outcome to be a species which generates its song by simply emitting syllables entirely at random, without following any grammatical constraints. However, other selective pressures operating simultaneously may prevent the reaching of this extreme condition (as may lack of evolutionary time or insufficient genetic variation), so it is not unreasonable to expect that we may observe a species producing grammatically-constrained song.

This assumes, of course, that the grammars are learned rather than innate. There are certainly species for which this is not a good assumption, as mentioned above in note 2.

An alignment between two strings of symbols, or *residues*, is a representation that explicitly shows the subsequences common to both strings. Short runs of spaces, or gaps, may be inserted into one or both sequences in order to shift portions of the sequence to the left or right in order to maximize the number of matching symbols which are aligned between the two strings and minimize the number of nonmatching symbols which are aligned. A *multiple alignment* is an alignment between three or more strings, whereas a *pairwise alignment* includes only two strings.

The Needleman-Wunsch algorithm utilizes a dynamic programming approach to pairwise alignment by filling in an alignment matrix indicating the cumulative scores of all the different possible partial alignments up to the current position in the two strings. By specifying different parameters, such as the *gap open penalty* or the *mismatch penalty*, one can adjust the behavior of the algorithm in various ways to suit the task at hand. Although the Needleman-Wunsch algorithm computes only pairwise alignments, it can be used iteratively to incorporate another sequence into an existing alignment, thereby producing a multiple alignment. This is similar to the popular CLUSTALW program, which first computes a *guide tree* in order to determine the order in which the pairwise alignments should be performed. Although a guide tree was not used

in this study, future work may wish to do so, in order to account for evolutionary relationships between the strings to be aligned.

## References

- [1] Bitterbaum, E., and L. F. Baptista, 1979. Geographical variation in songs of California House Finches (*Carpodacus mexicanus*). *Auk* **96**:462-474.
- [2] Cormen, T. H., C. E. Leiserson, and R. L. Rivest, 1990. *Introduction to algorithms*. MIT Press, Cambridge, Massachusetts.
- [3] Durbin, R., S. Eddy, A. Krogh, and G. Mitchison, 1999. *Biological sequence analysis: probabilistic models of proteins and nucleic acids*. Cambridge University Press, Cambridge, United Kingdom.
- [4] Gold, E. M., 1967. Language identification in the limit. *Information and Control* **10**:447-474.
- [5] Gold, E. M., 1978. Complexity of automaton identification from given data. *Information and Control* **37**:302-320.
- [6] Goldberg, D. E., 1989. *Genetic algorithms in search, optimization and machine learning*. Addison-Wesley, Reading, Massachusetts.
- [7] Greene, E., 1999. Toward an evolutionary understanding of song diversity in Oscines. *Auk* **116**:299-301.
- [8] Hopcroft, J. E., and J. D. Ullman, 1979. *Introduction to automata theory, languages, and computation*. Addison-Wesley, Reading, Massachusetts.
- [9] Horne, R. S., 1995. Spectrogram version 2.3.  
<<http://www.visualizationsoftware.com/gram.html>>.
- [10] Lynch, A., and A. J. Baker, 1993. A population memetics approach to cultural evolution in Chaffinch song: meme diversity within populations. *American Naturalist* **141**:597-620.
- [11] Margoliash, D., and A. C. Yu, 1996. Temporal hierarchical control of singing in birds. *Science* **273**:1871-1875.
- [12] Mundinger, P., 1975. Song dialects and colonization in the House Finch, *Carpodacus mexicanus*, on the East Coast. *Condor* **77**:407-422.
- [13] Pitt, L., and M. Warmuth, 1989. The minimum DFA consistency problem cannot be approximated within any polynomial. STOC '89: Proceedings of the Twenty First Annual ACM Symposium on Theory of Computing.

- [14] Pytte, C. L., 1997. Song organization of House Finches at the edge of an expanding range. *Condor* **99**:942-954.
- [15] Rissanen, J., 1978. Modeling by shortest data description. *Automatica* **14**:465-471.
- [16] Smith, A. R., 1996. Cultural evolution in House Finch song. Unpublished honors thesis, Williams College, Williamstown, MA.
- [17] SPSS, Inc. SPSS for Windows Release 8.0.0. SPSS, Inc., Chicago, Illinois.
- [18] Tracy, T. T., and M. C. Baker, 1999. Geographic variation in syllables of House Finch songs. *Auk* **116**:666-676.
- [19] Werner, M., 1998. daVinci V2.1.x online documentation.  
<<http://www.tzi.de/davinci/docs/>>. Universitat Bremen, Bremen, Germany.
- [20] Wolff, J. G., 1982. Language acquisition, data compression and generalization. *Language and Communication* **2**:57-89.
- [21] Zar, J. H., 1996. *Biostatistical Analysis*. Prentice Hall, Supper Saddle River, New Jersey.

## **DAWKINS BAD IDEA: MEMES, GENES, AND THE METAPHORS OF PSYCHOLOGY**

By

**Arthur J. Marr**

Benecom Technologies, Inc.

P.O. Box 29445, New Orleans, La. 70160

Lanna@bellsouth.net; and drmezmer.com

In times past, if the devil didn't get you, the vapors would, and if they didn't, humours, poisons, bile or any number of fanciful entities would do you in. With Pasteur and 19<sup>th</sup> century biology, these agents of illness were replaced with microscopic organisms, and the invention of the disease model made it easy to attribute your aches and pains to malicious bacteria, viruses, or other little microscopic buggers. Of course, then as now, few people understand the actual biological processes that are responsible for disease, but the *metaphors* for disease do just fine, and have been duly incorporated into the common vernacular. Some may say that these metaphors have worked a bit too well, since they have made a Procrustean stretch to cover all sorts of behaviors, from alcoholism to gambling. Nonetheless, the incorporation of metaphors from the syntax of science does provide a correcting influence to common sense, which earlier had only recourse to metaphors that engaged evil spirits and deadly vapors to help explain the world.

Nowadays, man has access to a wealth of metaphors from modern science. Cancers, black holes, laser beams, and computer viruses have replaced the vitalistic metaphors that assigned causes to ethereal spirits, evil demons, or invisible ethers. Of course, the incorporation of the metaphors of the biological and physical sciences into common language does not entail the ability to map actual processes, but only *suggest* those processes. The metaphorical description of a cold and its viral causes does not equate with a biochemical or biological description that requires a strict syntax and data language all its own. Thus when we mix our metaphors by talking about rampaging viruses spreading like wildfire from person to person, we know that our description only suggests what viruses are doing in the large, not the actual processes that cause them to propagate and harm.

The mixing of metaphors from different data languages can be poetic insofar as it suggests the juxtaposition or correlation of causes and events, or it can be interpreted as *literal* insofar as it presumed to denote actual processes. But how do we know if 'rampaging viruses' are a literal or a figurative representation of the truth? The ability in science to distinguish the literal from the alliterative is the mark of good science, and good science writing. A physicist may talk about matter using the metaphors of billiard balls, time warps, and cosmic string, but the literalness of those concepts is intentionally undermined by a continuous restatement of the mathematical metaphors that belie their literal reality. So why does a physicist engage two sets of metaphors when he can participate in his science quite well without the need to postulate billiard ball atoms and the like? It is simply because 'understanding' requires it. On the one hand, common sense metaphors are easily understood through their appeal to conceptual domains that we readily perceive (e.g. up, down, fast, slow, hot, cold), but mathematical metaphors (e.g. E=mc<sup>2</sup>) correct

for the tendency to make their existence literal. Well-written books that popularize ‘hard’ physical science all recognize the necessity to utilize two different sets of metaphors that correct the deficiencies of the other. Thus an understanding of the physical world can engage metaphors that are derived from our native experience and those that are derived from the abstract language of mathematics. Understanding consists in our ability to move from one set of conceptual metaphors (e.g. green grass, hot suns, and expanding universes) to another (e.g. the calculus, non-Euclidean geometry). It is no less than our ability to shift between different languages that enables us to envision the world.

Sometimes however, two entirely different sets of conceptual metaphors may be quite similar in terms of the processes they describe, and proceed to confirm rather than contrast with one another. Newtonian ideas of acceleration, mass, gravity, force, etc. have long been assimilated into our popular lexicon because a Newtonian view of the universe coheres with our own naïve experience. The mechanical universe of Newton corresponds with common sense theories of the physics of cars, boats, apples, and other physical objects, and is much easier to understand and accept than other more accurate physical theories that are reflected in the conceptual metaphors of Einsteinian relativity and Quantum theory. However, the conceptual metaphors associated with relativity and the invisible quantum correct for the literal interpretation of Newtonian mechanics as a representation of reality, and have entailments (e.g. time travel, multiple universes, quantum indeterminacy) that are dramatically at odds with the Newtonian conception of a clockwork universe.

In the biological sciences, the Darwinian principles of natural selection have an import comparable to Newton, and the conceptual metaphors of evolution find an equal correspondence with common sense theories of human psychology. Common sense or ‘folk’ theories of psychology tell us that we are motivated by ordinary objects (e.g. cars, jewelry, money, sex) whose value we determine consciously, and either impel (as in eliciting drives and reflexes) or compel (as in rewards, reinforcers, or punishers) behavior. In a Newtonian sense, our lives revolve around the collection of ordinary objects that push and pull us to them from a distance, and populate but scarcely our psychological universe.

This common sense explanation of how behavior is selected bridges quite easily to conceptual metaphors that describe how biological entities are selected, and by implication to the behavior that is instantiated by those entities. Thus patterns of behavior that are elicited by instinctive events can be ultimately attributed to individual genetic influences that are objectified in the metaphors of the activities of individual genes. Similarly, the common sense notion that ideas are selected by some obscure competition between objective alternatives also finds an equal bridge to selectionist principles that are derived from biology. Thus, just as Newtonian physics and common sense physics seem to confirm each other, common sense psychology and Darwinian biology share similar metaphorical principles that explain respectively how behavioral and biological selections are made.

As with the blending of metaphors that saw the adoption of Newtonian terms into common sense physics, metaphors from biology and ‘folk’ psychology have also become commingled, and thus form a new explanatory framework for behavior that explains behavior as a Darwinian process. Thus genes become ‘selfish’, and ideas or memes become ‘contagious’. But these are only two

different levels of thinking, and do not implicate the metaphorical schemes that explain the actual neurological processes that underlie behavior.

Darwinian or sociobiological explanations strongly imply that the molar processes of cumulative selection that led to bumblebees and human beings are isomorphic with neural or ‘molecular’ processes of the mind that lead to the selection of behavior. But an implication is *not* a demonstration, as a sociobiological explanation merely establishes a similarity between the *metaphors* of common sense and natural selection. The overriding question is not whether genes and memes represent a good metaphorical bridge between common sense and biological explanations for behavior, but whether they provide an equally good metaphorical representation of the *biobehavioral* processes that instigate behavior. The question is analogous to that posed by Quantum physicists to the common sense and academic views that extended the metaphors of Newtonian physics to the molecular world of the atom. The answer to that question was not a reaffirmation of the perspective of atoms as mere baby solar systems, but of the creation of an entirely new science that was equally rooted in the empirical tradition of science. That science was quantum physics.

A biobehavioral explanation of behavior represents the mapping of the actual neurological processes that comprise behavior to the patterns of information or environmental contingencies that parallel and elicit them (Donahoe and Palmer, 1993). This information is in turn mediated by somatic events that are perceived as emotion (Damasio, 1996), activating neurological events that comprise attentional processes (Donahoe and Palmer, 1993), and nativistic (i.e. inborn or genetic) sensitivities to certain abstract patterns of information (Bolles, 1976). Whereas a biobehavioral explanation is neurally realistic because it ties behavior to actual neural and informative events, a sociobiological explanation is neurally unrealistic, and merely substitutes neurological processes with Darwinian metaphors.

Biobehavioral science, which is also known as 2<sup>nd</sup> generation cognitive science (Lakoff, 1999), or theoretical behaviorism (Staddon, 1990), is, like evolutionary psychology, entirely informed by evolutionary principles. However, it is more rigorously empirical because of its insistence on ultimately observing or reliably inferring the neural processes that intervene between information and behavior. Biobehavioral and evolutionary psychology are represented by entirely different sets of conceptual metaphors that are respectively entailed by molecular (small scale processes and time frames) and molar (large scale processes and time frames) processes. The polarity of these metaphors is remarkable, and can be reduced to the following contrasting principles. Evolutionary psychology and folk psychology share the implicit presumption that decision making is generally based on the conscious and disembodied appraisal of ordinary objects that lead to the maximization of our self interest. In contrast, biobehavioral psychology has demonstrated that most reasoning is not conscious but nonconscious (Lewicki, 1992; Greenwald, 1992), and is guided by embodied non-verbalized somatic (Damasio, 1994) and neural activation events (Donahoe and Palmer, 1997) that are ‘just as cognitive as any other perceptual image’ (Damasio, 1994). Because nonconscious embodied reasoning is computable but not directly accessible by conscious reasoning, we often find that our conscious reasoning about what is a ‘best outcome’ conflicts with our nonconscious determination of ‘best outcome’. Thus there is no univocal or self-consistent locus of value (Lakoff and Johnson, 1999). (Remember this next time you reluctantly try to get up in the morning.) Values are due to the binding of many information

streams that are mediated by disparate neural and somatic processes, and motivate behavior in real time as they are perceived, and not when the physical or objective entity that denotes such information is attained. (In other words, it's the thought that counts) Thus value is not found in some conceptual object like a meme, but in separate threads of information that are mediated by the mind and body that individually have salience to an individual and are perceived presently, independently, virtually, and for the most part nonconsciously. In other words, the concept of ideas as compartmentalized memes leads us to find value in the obvious topographical or 'surface' aspects of an idea, whereas it is the non-obvious abstract properties that are actually selected. Because value resides in information that is often incognizant to us, it cannot be subject to the economic models that are based on the rationing of value according to some single utilitarian measure, or the mathematical decision models such as game theory that conceptualize value simplistically as being no more than material wealth.

The definition of a meme as an independent conceptual object is ultimately not simple, but simplistic, since it does not denote the web of informative relationships between behavior and the environment that is denoted by consciously and nonconsciously by the mind and body proper. For example, the concept of the sport of football is a well traveled meme to be sure.

Football represents a rather involved information pattern that has infected the minds of young men nationwide, and football games, commentary, and assorted chatter has parasitized not only the minds of people, but the network airwaves, the written media, and many unwilling housewives. But is a football game an indivisible meme like entity, or is it somewhat different than the sum of its parts? Actually, the 'meme' of football is not a singular information pattern that replicates like a strand of DNA, but rather emerges from a web of separate patterns of information that are mediated not only by consciously perceived information but by neural and somatic activating processes that we otherwise call emotion. The meme of football is not just a compendium of rules, but comprises the memory of the somatic responses that occur while watching (excitement, depression), the natural feeling of elation that occurs with a high state of alertness, the virtual extension of control over all those partisans of the losing team, the constantly changing and stimulating prediction error that occurs as one play after another unfolds, the smell and taste of hot dogs and beer, the camaraderie of friends, and so on. The meme of football is in other words a web of perceptual relationships that is volatile and constantly changing. Moreover, different aspects of the meme football may be present in one circumstance, and not present in another. Watching your team lose at a hostile stadium on a rainy day is a whole lot less rewarding than if you were watching your team at home while among friends. And what is a meme when it becomes frayed around the edges?

The most important distinction contrast between evolutionary psychology and bio-behavioral psychology is that bio-behavioral psychology denotes value not in the assimilation of ideas or memes, but in changes in the *relationships* between memes, or behavioral discrepancies (Donahoe and Palmer, 1993). To explain this, we must understand first how a meme does *not* reproduce. Although a meme represents a self replicating packet of information, unlike a virus it possesses no internal instructions that secure its influence on behavior, let alone its retention in memory. Memes or ideas take root in memory because they are rehearsed, and they are rehearsed because of their contingent relationship to a myriad other ideas that comprise the stimulus context of a behavior. This idea of contingency is critical to the methodology of modern

behaviorism, and underlines the fact that it is not ideas alone that motivate, but the dependencies between ideas. The meme of a fishhook for example hardly comes to mind until it is perceived as part of a means-end (memes-end?) expectancy. We think of fish hooks because of the fish it can catch, but to even think about fishing one must also think about the time, place, and equipment that allows one to fish. If any of these events fail to take place, there is hardly a need to think about fishing, or for that matter the meme of fishhooks. An atomized universe of memes does not implicate the contingent relationship between ideas that secures the rehearsal and retention of a ‘good idea’. Behavior is elicited not by individual memes but by global maps of means-end expectancies that are constantly changing, and are in general non-consciously perceived.

However, what causes us to think about fishhooks, fishing trips or other ideas is the fact the relationships they denote are selected and are mentally rehearsed. We constantly think about a fishhook as it winds its way from our tackle box to the end of our fishing line because in every moment the relationship between the fishhook and the line changes, and it is the change that gains our attention. Thus, we select not only memes, but also the abstract *relationships* between memes as they are moderated by our thoughts and overt behavior. Ultimately, as Alexander the Great found out when he wept upon having no more new worlds to conquer, what motivates is not the end, but in the traveling.

Unfortunately, Darwinian and common sense models can no more describe the molecular ‘environment-behavior’ relationships that comprise human motivation than a weatherman’s description of an impending cold front describes how a snow storm forms over your head. We would err in using a molar analysis (cold fronts) to describe molecular process (the formation of clouds) because the inherent processes implied by storm systems and storm clouds are different. Likewise, the human brain is a massively parallel biological computer, and metaphors from information processing are far more apt than biological metaphors that liken ideas to viruses and their spread to contagion, let alone the metaphors from common sense that posit a disembodied objectified reasoning. The lack of ‘fit’ of Darwinian and common sense metaphors to bio-behavioral science does not invalidate the selectionist principles that inform all of the sciences. But it does point out the level confusion that occurs when a set of principles from one level of understanding (biology) are invoked not just to explain another (an in the juxtaposition of the metaphors of cosmic string and mathematics) but to *embody* another. In other words, because the uses of memes and genes is not corrected by an understanding of how the brain as a neural system actually works, the metaphorical conception of memes and genes can easily be seen as not just figuratively real, but literally real.

The ultimate danger in assigning a literal reality to the means-end rationality imposed by utilitarian memes and genes is that it implies that we implicitly know what is in our best interest. Furthermore the convergence of the metaphors of common sense and Darwinism reinforce the idea that value is objectified and is a limited commodity, and must be allocated to those who are most fit to achieve it. In this way, a meme world becomes ‘mean’ world, wherein our memetic impulses robotically drive us forward to achieve our goals, with the long term survival of our genes and memes being the only necessary outcome.

In contrast to this cold and sterile vision, bio-behavioral psychology defines value not as a scarce commodity, but in the creation of information that is consumed virtually. That is, if value is

denoted in abstract informative relationships between ideas, then it is prospectively unlimited, and is constrained not by our inability to manufacture physical things, but by our ability to create and perceive information. But perception requires the skills that enable us to mentally model the world, from the implications of the cheers of a crowd in a football game to the thoughts of a proud parent. To experience the world is to model it, and that is nothing more than empathy. Universal empathy allows us to expand and enhance the rewards we perceive, but it also constrains our behavior due to the virtual penalties (e.g. shame, embarrassment) we perceive.

Cultures that understand that value derives from the development of empathy will take an entirely different course than the materialistic societies that posit value as the accumulation of objects. Indeed, information is more economically produced by a societal exaltation of sports, art, literature and music than by the manufacture of a new prestige automobile. Ironically, the lasting legacy of a psychology that is informed by evolutionary principles is not the amoral world driven by the erroneous metaphors of selfish genes and infectious memes, but by the evolutionary mandate of an expanding empathy, and our innate interest in the cultivation of beauty.

Bolles, Robert C. (1976) *Theory of Motivation*. 2nd ed. New York: Harper and Row

Damasio, Antonio R. (1994) *Descartes Error: Emotion, Reason, and the Human Brain*. Avon

Greenwald, A.G. (1992) Unconscious Cognition Reclaimed, *American Psychologist*, 766-775

Lewicki, P., Hill, T., & Czyewska, M. (1992) Nonconscious acquisition of information, *American Psychologist*, 47, 796-801

Donahoe, J. W. and D. C. Palmer (1993) *Learning and Complex Behavior*, Allyn and Bacon

Donahoe, J. W., D. C. Palmer, and Jose E. Burgos (1997) The Unit of Selection: What do reinforcers reinforce?, *Journal of the Experimental Analysis of Behavior*, 67, 259-273

Donahoe, J. W. (1997) *Neural-Networks Models of Cognition*, J. W. Donahoe and V. Packard Dorsel (Eds.)

Lakoff, George, and Johnson, Mark (1999) *Philosophy in the Flesh: The Embodied Mind and Its Challenge to Western Thought*, Basic Books

Staddon, John (1993) *Behaviorism*, Duckworth

(The best introduction to bio-behaviorism comes from Shull's article on the website of the Journal for the Experimental Analysis of Behavior, and the scholarly commentary also at the JEAB website that discussed this new school of behaviorism. An understanding of metaphor and how it heavily influences ideas in evolutionary psychology can be found on the many web sites that discuss the work of the cognitive linguist George Lakoff, and in particular his new book: *Philosophy in the Flesh: The Embodied Mind and its challenge to Western Thought*... and of course there is my own site, drmezmer.com, which is pure irony. )

# CRASH CONTAGION

By  
Paul Marsden  
1997

Graduate research Centre in the Social Sciences  
University of Sussex  
[PaulMarsden@msn.com](mailto:PaulMarsden@msn.com)

## Abstract

*Using a memetic interpretation of the events following the death of Diana, this paper introduces memetics, by way of example, to social scientists and social researchers. It is argued that standard social scientific interpretations of human behaviour are overly rational and may ultimately be quite unhelpful in understanding human behaviour. The paper explores the possibility that memetics may provide researchers with a viable non-Cartesian conceptualisation of the human individual and behaviour. Substantively, the paper argues that much human behaviour may be a product of non-rational imitation and suggests that such an interpretation may help explain the mass hysteria following the death of Diana, Princess of Wales.*

"It gives you a scare. Initially it just looks like a white mass."

"A shiver went down my spine. It was Di... Seriously top right hand corner. It's just there"

"Princess Diana's face is looking out of it. Everybody's seeing her face looking out"

"We saw it, as clear as day. You know the pose, the one with her head cupped in her hands. She's got the tiara on"

These people are referring, of course, to the visions of Diana that appeared to mourners waiting patiently in line to sign one of the Books of Remembrance at St James Palace in September 1997. A vision, we were told by these tearful mourners, that was a copy-conform of Diana's Vogue cover pose. The visions started the same day, in the middle of the planetary Diana'thon, that an Iraqi newspaper, the Babel, reported that Princess Diana had been assassinated by the British secret service.

Di' mentia had seized the world's press. The morning following Diana's death, the Observer newspaper didn't have time to cancel the article on Diana which concluded that "if her IQ were five points lower (she) would have to be watered daily". The Sunday Mirror which following her death was creased with grief had just a few weeks earlier described Diana as "trivial and brain-dead". But the accident changed all that.

From Sunday morning, news-readers, sounding increasingly like narrators of a badly written romantic tragedy, delivered hour upon hour of repetitive commentary, only punctuated from time to time by solemn, ashen-faced "experts" who would advise us "grieve in our own way" the death of this, the People's Princess.

The media may have been suffering from Di'mentia, but the world was suffering from Di'mania. The deification and the arguments for canonising the icon Diana began well before the visions. With Diana's death, a cult has been born; a secular saint had been created by popular acclamation. Britain's own Mother Theresa Lite, with the option on good legs included, was perfectly packaged for the age of the video-clip and the sound-bite. "The Queen of Hearts" became "the light of our life", "a candle in the wind" who possessed, as a member of the British House of Lords confirmed, a "genuine gift of healing".

Saturday 6th September 1997 was a world first; half of the human inhabitants of Planet Earth were wired up and tuned in to witness the most widely experienced event in our history. The funeral of Diana, Princess of Wales. The funeral was broadcast live to over 60 countries in 44 languages with an estimated audience of 2.5 billion people.

Television stations dropped their scheduled programmes and replaced them with tributes to Diana. Pop radio stations decided that pop music was inappropriate, some stopped broadcasting altogether whilst others threw themselves into weeklong session of mournful "sombre music only..." One London radio station even offered free counselling to those in distress. The event temporally but fundamentally modified peoples' behaviour. Elton John even took out his earrings, and the Union Jack was flown at half mast over Buckingham Palace for the first time. Millions of people converged on London for Diana's funeral, more people were present than had been at the celebrations in London at the end of the Second World War.

It is our interpretation, as social scientists, to the reaction to the death of Diana that I would like to pause for a few minutes and reflect upon. Now, one obvious line of social scientific enquiry would be to try to provide an account of why people responded in such a way to the death of Diana. We could explain their behaviour by referring to notions of self and identity, to internal states, to attitudes and dispositions. We could interpret the mass mourning as an example of how our inner human selves can sometimes break out of the iron cage of rationality in an alienating capitalist (post) modernity. We could look for the real, true and underlying reasons for the Diana phenomena interpreting it from within one of the plethora of theoretical paradigms that we as social scientists have at our disposition.

To take a brief example of how we can use inner states to explain our behaviour to the Diana phenomenon, I want to concentrate just on our immediate reaction to the news of her death. Clairvoyance and other X-Files-type phenomena aside, the death of Diana must have been temporally antecedent to our reaction to it. For each one of us, there must have been one fascinating moment between the time that the patterns of information reporting the death of Diana entered our heads and the time that we reacted or responded to the news. That fascinating moment was the moment when the electrochemical signal travelling up the axons of neurons in our heads became conscious thought. Cartesian psychology would have that in some dark corner of our brains where consciousness presumably 'happens', the signals were translated into the language of thought, a sort of '*mentalese*', and were then projected as subjectively meaningful internal representations in our heads, to a special audience - the conscious self. (That same conscious self that "sees" the image of Diana when you shut your eyes and imagine her.) The incoming signal became conscious experience and we finally realised that Diana is dead. In this way, we, ourselves, the real, conscious selves, could then evaluate this shocking new fact from

the control centre of the mind and send out instructions to our bodies and brains about how to react to it. And just how did we react to the news when we became conscious of it?

Some of you may have responded with emotional upset, some of you even may have taken your earrings out as well, whilst others may have just carried on eating breakfast. Your reactions will have been rich and varied, fascinating stuff for social scientists, but I am willing to bet that all of you, every single one here today, reacted identically in one very special way. Despite our preciously individual inner selves, I suspect you all reacted in a slavish, machine-like fashion to the news of Diana's death with a response that doesn't need the Cartesian crutch of that inner self to explain it.

You just had to talk about it. You had to communicate the news in some way or another to someone else. You may have phoned somebody in order to talk to them about it, or the subject may have come up in conversation at a later time, but the information that you received got itself passed on, copying itself, infecting another brain. It was almost as if the words in your head wanted to get themselves said, rather as if your mind had been infected with a virus, much like a flu virus, which in spite of yourself, spread themselves in a fit of verbal sneezing.

I think there is an alternative to positing homunculi, mysterious internal agents or selves endowed with all sorts of intentional, evaluative and emotional capacities in our brains to explain the Diana phenomenon. In any case, these sort of 'explanations' don't really explain anything at all, they just take us down the tedious road of infinite regress: If you argue that we can evaluate information, and respond intentionally because we have selves, then how is that selves come to be intentional and have rational capacities in the first place? If our selves are looking on in the inner theatre of consciousness, who is looking in on their inner theatre consciousness? Do our selves have their own selves?

No, our reaction to the death of Diana does not need the theoretical and often circular crutch of the self. I think we need to understand our reaction as a symptom of having been infected with a virus; a mind virus. The death of Diana can be explained by the epidemiology of a contagion, a crash contagion that leapt from mind to mind infecting new brains. A virtual virus that hijacked the copying machinery in our brains to copy itself onto other brains through a process of communication. In this way the crash contagion spread through the population through a process of replication in a global game of Chinese Whispers, as if it was a virus.

Our behaviour following the death of Diana owes much more to epidemiology than it does to the reaction of some inner self. Just as virtual pets, Gameboys and Rubik's Cubes sweep through, and leap between, schools in a manner virtually identical to measles or chicken pox, so did the Diana Crash Contagion. Whether the symptoms are spots, sneezes or the fixed motor behaviours associated with feeding your virtual pet or reacting to the news of Diana's death, we need epidemiology and not internal agents to explain our behaviour.

A key issue in the spread of the Diana crash contagion was the role of the mass media as a vector. More generally, the mass media has been responsible for an explosive proliferation of all sorts of mind viruses that exist in today's infosphere spread via radio, television, the press, intranets, internets, fax machines, e-mail and voice mail. Every year, your brain is assaulted by

thousands of designer mind viruses trying to infect you in the form of advertisements. Our minds today are suffering from a chronic case of information overload, our brains are clogged up by virtual parasites, mind viruses that spread around the world at the speed of light, replicating at rates that make AIDS and meningitis appear quite unthreatening in comparison. And mind viruses can be just as deadly as their biological cousins, leaping promiscuously from vehicle to vehicle, and from medium to medium, coding for all sorts of behaviour, whilst proving to be virtually unquarantinable (Dennett 1990).

The study of mind viruses is known as memetics, and the unit of replication, the mind virus is called a meme. The term was coined by Richard Dawkins, author of *The Selfish Gene*, a successful popularisation of a gene-centred Darwinian understanding of evolution. In addition to his now familiar metaphor of the human body as a gene-machine, Dawkins argued that we were hosts not to just to one replicating entity (the gene) but also another, virtual mind viruses that he called memes.

"A meme is a unit of cultural transmission, or a unit of imitation. 'Mimeme' comes from a suitable Greek root, but I want a monosyllable that sounds a bit like 'gene' . . . it could alternatively be thought of as being related to 'memory' or to the French word *même*. . . Examples of memes are tunes, ideas, catch-phrases, clothes fashions, ways of making pots or of building arches. Just as genes propagate themselves in the gene pool by leaping from body to body via sperm or eggs, so memes propagate themselves in the meme pool by leaping from brain to brain via a process which, in the broad sense, can be called imitation." (Dawkins 1989)

A meme is a unit of cultural replication, a cultural instruction if you will, that codes for the cultural equivalent of sneezing. Berman's famous quote, "An idea is something you have; an ideology is something that has you" illustrates the logic in a memetic understanding of human behaviour. For memetics, the human agent can be understood as a system of co-apted meme-complexes that have infected or been inherited in the brain. Our behaviour is determined by an innate disposition for mechanically and automatically executing and thereby replicating these instructions. From this perspective, a prerequisite for understanding this sort of imitative behaviour is being able to audit and understand the infection strategies of the particular mind viruses resident in our brains. According to memetics, the legitimate focus in the study of human behaviour is the message not the source. We, ourselves are held to be emergent properties of an infectious and interacting complex of replicating mind viruses, and our bodies are but vehicles that have evolved to help these replicators, replicate.

Daniel Dennett has summed up the memetic paradigm succinctly

"A scholar is just a library's way of making another library". (Dennett 1990)

From this perspective, memes have us, we don't have them. We are simply user illusions, a neat shorthand for the emergent property of memes competing for space and salience so that they can get replicated. Returning to our reaction to the death of Diana, we have to understand that first rules of memes, or mind viruses - as it is for genes, is that replication is not necessarily *for* anything: Good replicators or good ideas are not necessarily true, good or beautiful ideas, they are simply good at replicating. There need be no rational reason *why* there was such a massive

collective reaction to the death of Diana, it was simply because the Diana crash contagion was particularly infectious, we were predisposed to imitating it. The challenge for memetics is to understand what it is that makes certain mind viruses more contagious than others, by understanding their structure and dynamics.

I know that it is more comforting to think that there are reasons behind our behaviour, that we consciously decide on how to respond to a stimulus, but often it just isn't the case. Memetics argues that we don't have to look for internal explanations for why we behave in certain ways; sometimes we can provide explanations by shifting our focus to the nature of the message itself. How is it that certain mind viruses are more infectious than others, for example, the meme or mind virus whose primary unfortunate symptom is an instruction for the host to go and jump off a cliff? What is the mechanism of infection? How do we develop immunity to some mind viruses and not to others? This is the focus of memetics: Understanding the epidemiology of mind viruses and their associated symptomalogies. Rather than ask why so many of us believe x, or why so many of us behaved in a certain way following say, the death of Diana, memetics asks what is it about our behaviour that made it so infectious, and how did it spread?

Mind viruses spread through a (Darwinian) process of imitation. We know for example that if a suicide is reported in the mass media, the suicide rate will rise in the following month. (Phillips 1986) A few people with poor immunity to this mind virus will become infected and display similar symptoms to the original victim - they will commit suicide. This is called the Werther effect, after Goethe's novel *The Sorrows of the Young Werther*, which was banned in several countries after readers started imitating the hero's suicide (Phillips 1974). We have all heard of copy cat riots, copy cat murders and various other "non-rational" behaviour such as mass suicides of the members of Heaven's Gate and the People's Temple. Conscious choice or contagion? The amount of violence that seen on US television screens correlates positively with US homicides (Phillips 1983). We know as well that suicide victims following a publicised suicide story will more likely than not resemble the reported victim (Phillips 1974). They are victims in very real ways, victims of virulent mind viruses, with symptoms much more deadly than a sneeze. Following the death of Diana, an event witnessed by nearly half of this planet, we can reasonably predict that some immuno-deficient individuals will imitate her behaviour and push up the car fatality figures for September 1997. The relevant statistics are not yet available but I would predict an increase in imitative behaviour among non-married couples. So watch out if you're driving home with your partner tonight - you may suffer from one of the rarer but deadly symptoms of the Diana crash contagion.

And just as we are beginning to see the fruits of genetic engineering, the possibility of memetic engineering is emerging. Because of their substrate neutrality, memes may evolve in computers just as well as human minds, and the Prediction Company in Santa Fe, New Mexico, has been sponsored by US affiliate of the Swiss Bank Corporation to evolve memes that make predictions useful for currency trading (J. Doyne Farmer in Brockman 1996). Memes based on Darwinian algorithms have also helped design fibre-optic telecommunication networks, detect enemy targets in infrared images, improve mining operations, and facilitate geophysical surveys for oil exploration (Cziko 1995). And myself, I make my living out of memes. I am at the grubby end of social science, I get my hands dirty with industry - I develop marketing campaigns and advertising campaigns for multinationals. I build and engineer designer mind viruses, and my

hope today is that I have infected you with a new meme, a meme that could be a useful tool in your new social scientific careers, the meme meme.

## Bibliography

Brockman, J. (1996) *The Third Culture* USA; Touchstone

Cziko, G. (1995). *Without Miracles: Universal Selection Theory and the Second Darwinian Revolution*. USA; MIT Press (A Bradford Book).

Dawkins, R. (1989) *The Selfish Gene* (2nd Edition) UK; OUP

Dennett, D.C. (1990) "Memes and the Exploitation of Imagination." in *Journal of Aesthetics and Art Criticism* 48: 127-135.

Phillips, D.P. (1974) "The Influence of Suggestion on Suicide: Substantive and Theoretical Implications of the Werther Effect" in *American Sociological Review* Vol. 39 (June) 340-354.

Phillips, D.P. (1983) "The Impact of Mass Media Violence on US Homicides" in *American Sociological Review* 1983, Vol. 48 (August) 560-568.

Phillips, D.P. (1986) "Clustering of teenage Suicides after Television News Stories about Suicide" in *The New England Journal of Medicine* Vol. 315 (Sep) 685-694.

# **OPERATIONALIZING MEMETICS: SUICIDE, THE WERTHER EFFECT, AND THE WORK OF DAVID PHILLIPS**

**By**

**Paul Marsden**

Graduate Research Centre in the Social Sciences  
University of Sussex  
e-mail PaulMarsden@msn.com  
tel/fax (44) (0) 117 974 1279

## **Abstract**

One of the major challenges currently facing memetics is the issue of how to successfully *operationalise* the emerging paradigm. In other words, how can we exploit the innovative analytical framework of memetics in order to generate a body of theoretically informed empirical research? Whilst there are some important *theoretical* issues that have yet to be resolved, the future success of memetics, qua academic discipline, may depend not so much on elaborate theoretical developments, but on the results of *empirical* research findings. Operationalising memetics will not only involve subjecting memetic theory *itself* to empirical testing, but it will also mean assessing the usefulness of the paradigm in describing, understanding and explaining the sociocultural patterns and phenomena that are the *traditional foci of the social sciences*. To this end, the following paper will explore the operationalisation of memetics by reviewing the work of the David P. Phillips, a sociologist who has been publishing empirical research on social contagions since 1974. Although Phillips has not explicitly referred to memetics, it will be suggested that a number of practical lessons and guidelines may be drawn from his research, and a possible outline for operationalising memetics will be proposed based on his approach.

Summary of empirical research on social contagions conducted by Phillips

<b>The Werther Effect: Fact or Fiction - Summary of Research by D.P. Phillips</b>	
<b>Key findings and Conclusion</b>	<b>Source</b>
Suicide rates increased significantly after suicide stories were reported newspaper stories. The increase was proportional to the amount of newspaper coverage devoted to the suicide stories	American Sociological Review 1974 Vol. 39:340-54
Car accident fatalities increased following media representations of suicide implying that some car accident fatalities were in fact imitative suicides	Science 1977 196: 1464-65
Car accident fatalities, particularly those resulting from single-car accidents increased significantly three days after a suicide story was publicised in the newspaper press. The increase was proportional to the intensity of the publicity, and the age of the accident victim was positively correlated with the age of the suicide story victim. There was also a corresponding correlation between murder-suicide stories and	American Journal of Sociology 1979 Vol. 84 No.5: 1150 -1174

multiple car crashes involving passenger deaths.	
Publicised murder-suicides were followed by an increase in aeroplane crashes (airline and non-commercial). The increase in aeroplane crashes was proportional to the degree of coverage that these stories received	Social Forces 1980 Vol. 58 (Jun): 1001-1024
Daily US suicide rates increased significantly (for a period of less than ten days) following the appearance of highly publicised suicide stories on television evening news programmes.	American Sociological Review 1982 Vol. 47: 802-809
Suicide rates, motor vehicle fatality statistics and non-fatal accidents all rose immediately following the transmission of fictional televised suicide stories in 1977	American Journal of Sociology 1982 Vol. 87 No. 6: 1340-1359
Homicides in the US increased following heavyweight championship prize-fights in a relationship that persisted after correction for secular trends, seasonal and other extraneous variables. The increase was found to be largest following heavily publicised fights.	American Sociological Review 1983 Vol. 48 (Aug): 560-568
Between 1973 and 1979 teenage suicides increased significantly following 38 nationally televised stories of suicide. The intensity of publicity devoted to the suicide stories was significantly correlated to its effect on teenage suicide rates	New England Journal of Medicine 1986 Vol. 315 No.11: 685-9

**"No fact is more readily transmissible by contagion than suicide." -- Emile Durkheim (Le Suicide [1897] 1951:141)**

In the mid-1770s a peculiar clothing fashion swept across Europe. For no immediately apparent reason, young men started dressing in yellow trousers, blue jackets and open-necked shirts. This mildly eccentric fashion spread from region to region in a manner strangely similar to the epidemics that were continuing to plague the Old Continent. It turned out that these 18th century fashion victims all had one thing in common; they had all been exposed to first novel of Johann Wolfgang von Goethe, *The Sorrows of Young Werther*. Goethe's novel recounted the desperate plight of Werther, a young man hopelessly in love with a happily married woman called Charlotte. In this intense and romantic tale, Goethe describes Werther's rather peculiar penchant for wearing a colourful mélange of blue jackets, yellow trousers and open-necked shirts.

Shortly after being published in 1774, Goethe's novel was banned in several areas across Europe. This was not because certain authorities held it responsible for the spread of a fashion of rather doubtful taste, but because there were signs that the book was also the vector for an altogether more serious contagion. The tale recounted how Werther's clumsy, but painfully sincere, attempts at winning Charlotte's heart ultimately failed. Destroyed by rejection, Werther saw no way out of his desperate plight other than suicide, and using a pistol, he dramatically put an end to his sorrows.

For a memeticist, the social consequences of the publication of Werther's tragic story are entirely predictable: Not only was Werther's dress code the object of imitation but so was his somewhat extreme code of behaviour. Anxious authorities around Europe received reports of

increasing numbers of young men imitating Werther's desperate act. Goethe himself became convinced that his tale was responsible for a continental wave of suicides.

"My...friends thought that they must transform poetry into reality, imitate a novel like this in real life and, in any case, shoot themselves; and what occurred at first among a few took place later among the general public..." (Goethe quoted in Phillips 1974:340)

In an attempt to prevent the suicides from reaching epidemic proportions, a number of authorities banned *The Sorrows of Young Werther* in the hope that the imitative behaviour would cease.

Two hundred years later, in 1974, the sociologist David Phillips coined the term "the Werther effect" to describe imitative suicidal behaviour transmitted via the mass media. Phillips devised an empirical research programme to establish whether media reporting of suicide stories really did affect suicide rates. In over a decade of research, Phillips produced important evidence that supported the hypothesis that behavioural patterns in society can in fact operate as contagions. Methodologically speaking, Phillips' research was interesting because it described one possible solution to the general problem of operationalising the memetic paradigm which, to date, has been dominated by anecdotal evidence.

Phillips' methodological approach could be described as *quasi-experimental* in that his analysis was based on an experimental protocol, but he worked with exclusively historical data sourced in the real, uncontrolled, world. Using newspaper records and official suicide statistics, he identified a number of 'control periods' defined by the absence of front-page newspaper suicide stories. Using the suicide statistics from these control periods, he generated a number of expected suicide rates for a pre-defined selection of 'experimental periods' during which front-page suicide stories were published. Working with a null hypothesis that the front-page newspaper reporting of suicide stories had no effect on aggregate suicide rates, he compared the expected rates with the actual rates. After controlling for seasonal and other spurious effects, Phillips tested for significance between the two values, and by comparing a number of different experimental periods he was able to test for a correlation between suicide rates and the intensity of media representations.

Phillips' technique was unusual in that other experiments on imitation and the mass media have tended to be conducted under artificial laboratory conditions. Whilst some of these experiments have yielded results that point to a correlation between media representations and imitative behaviour (e.g. Bandura, Ross and Ross 1963), they have been largely discounted for failing to accurately replicate the conditions and environment under which memes are transmitted in the heterogeneous and multifaceted social world.

The results of Phillips' quasi-experimental research could not be subjected to this "non-relevance" argument precisely because he was testing for a correlation between media representations and actual social behaviour *in* the social world. Importantly, Phillips was not testing for the relationship between individual behaviour and media representations; rather he was testing for a relationship between suicide rates and media representations. Put differently, Phillips was attempting to provide an explanation of social facts (suicide rates and media representation levels) through social forces (imitation) mediated via the communication

infrastructure of society. Such a macro-level of analysis is of important theoretical significance to memetics since it adds a structural element to a theory otherwise open to the charge of methodological individualism. Severe limitations on space preclude the development of this more theoretical tangent, but it is important to recognise that Phillips was testing for the replication of structural patterns in society rather than investigating the individual process of replication/transmission per se. In this way, his approach is an example of what might usefully be called macro-memetics, in contradistinction to the equally valid micro-memetic approach that currently dominates our paradigm.

Following Phillips' research protocol, such a macro-memetic analysis can be broken down into a certain number of stages that together might outline a possible method for investigating the structural epidemiology of memes:

1. Define the phenotypical expression/symptomatology that will be used to measure levels of meme infection (suicide)
2. Measure the prevalence of meme infection over time within a given population (suicide rates)
3. Measure the exposure rate within a population to this meme through a particular (mass) medium over time (circulation/viewing figures)
4. Calculate an index of exposure intensity (exposure level multiplied by the share of total medium content (column length/no. of days on front page))
5. Define a series of control periods where media transmission intensity = 0
6. Define an experimental period where media transmission intensity > 0
7. Regress meme infection levels during control period(s) to generate an expectation for the experimental period based on the null hypothesis that media representations of the meme have no effect on the incidence or prevalence of that meme
8. Test for significance between the expected and actual results
9. If expected and actual results are significantly different, test for further relationships (host similarity/correlation of intensity and suicide rates)

#### **Figure 1: A New Approach for Memetics?**

Such a research programme might yield results providing evidence to support the hypothesis that meme exposure partly determines the incidence and prevalence of meme infection. Put differently, Phillips' macro-memetic approach would help determine whether patterns of behaviour in society do in fact operate in a manner similar to that of contagions. This would provide the empirical foundations for the development of a fuller memetic account of behaviour, providing not only evidence for replication/transmission but also incorporating the processes of variation and selection. If it can be established that social facts may sometimes operate as

contagions, meta-studies of the epidemiologies of various different traits and practices (symptoms of meme infection) might yield clues as to what it is that makes a social contagion infectious, and which factors influence immunity/susceptibility.[1] Further, it might then be possible to measure both the fidelity of sociocultural replication as well as the average time-lapse between exposure and replication (the incubation period). Thus we could begin to generate results around a complete evolutionary loop of variation, selection and replication operating within the (infra) structure of various lines of communication. Mapped over time, structural lineages of mediated traits and practices could provide the basis for a memetic phylogeny of society. This would raise the interesting possibility of an innovative memetic typology of societies derived from the various modes, and relationship to the means, of cultural reproduction. Again, this is a theoretical theme that falls outside the scope of the paper, but such a typology might usefully draw on the social evolutionary theory developed by Jürgen Habermas (Habermas 1979).

Turning now to the empirical results of this 'macro-memetic' research Phillips has provided evidence to support the claim that suicide does indeed behave as a social contagion. Specifically, he demonstrated that exposure to suicide in media stories was a significant variable in accounting for UK and US suicide rates (Phillips 1974). He was also able to show that the intensity of meme transmission correlated positively with suicide rates. In a similar way, Phillips also correlated aeroplane and car accidents, murder and violent crime rates to mass media reporting. (Phillips 1977, 1979, 1980, 1982a, 1982b, 1983, 1986).

The implications of Phillips' research findings were radical; suicide appeared to behave as a contagion mediated via, and dependent upon, lines of mass communication.[2] However, despite consistent findings in favour of this replication/imitation thesis, the Werther effect has not been widely integrated into a comprehensive social scientific understanding of suicide. The reason for this is particularly relevant to memetics and thus merits a brief review.

There are certainly some important methodological problems in Phillips' research that could partly explain why replication/imitation is not a now central concept in contemporary suicide research. In interpreting the results, Phillips appears to make individual-level inferences from aggregate findings (ecological fallacy), he is uncritical of the reliability of official suicide statistics, and he completely ignores the key issues of definition, intention and performance (suicide as opposed to para-suicide (Platt 1984), suicidality (Thorlindsson and Bjarnason 1998), attempted suicide and more generally, risk taking behaviour (Taylor 1982)). These problems have led to an inevitable discounting of his findings in social scientific circles (e.g. Baron and Reiss 1985). The key lesson for memetics here is that it will be important for empirical research to take cognisance of, and address these issues of validity, reliability and meaning. The heterophenomenological approach adopted by Dennett (Dennett 1991) may go some way to address the problems of definition, intention and performance, and the problems of official statistics may be obviated by directly recording prevalence rates. However, the greatest problem for the replication/imitation thesis is theoretical; the Werther effect fundamentally undermines the still dominant Cartesian understanding of the human subject that underpins much social scientific explanation. Social contagions are fundamentally at odds with such an understanding that generally presupposes an irreducible source of intentionality and rationality (economic and cognitive) behind human behaviour. This homuncular understanding of the human subject is part

of what Barkow et al. (1992) have called the Standard Social Scientific Model (SSSM), and its domain assumptions virtually preclude taking evidence of the Werther effect seriously. Whilst the SSSM might accommodate evidence to the effect that the media acts as gate-keeper and agenda setter in the process of communication, any evidence suggesting that we, ourselves, are emergent properties of this communication process, rather than vice versa, is incompatible with the 'dogma of the ghost in the machine' that still dominates social science. An alternative understanding of human consciousness, perhaps similar to Dennett's 'Multiple Drafts' functionalist model, will probably have to become firmly established in social science before memetics starts to be taken seriously. Put differently, until Cartesian materialism/dualism ceases to underpin social scientific explanation, the obvious parallel between Darwin and the Creationists will continue to run deep.

However, I do think that Phillips' research has highlighted a weakness in the SSSM, which might provide a more direct opportunity for the establishment of a memetic understanding of suicide and other social behaviour within "respectable social science". This opportunity lies in the very inability of the SSSM to account for the phenomenon of replication/imitation from within its own homuncular paradigm. The more concrete evidence that memetics can provide to support the hypothesis that traits and practices in society do evolve according to principles of variation, selection and replication/transmission, then the more, I believe, the SSSM will be subjected to an increasingly acute crisis of legitimisation. This is why I believe the operationalisation of memetics is so crucial to the development of the paradigm.

Interestingly enough, the memetic challenge to the sociology of suicide does have an historical precedent, dating back a hundred years to the competing analytical frameworks of Gabriel Tarde (Tarde [1903] 1962) and Emile Durkheim (Durkheim [1897] 1952). However, Tarde's proto-memetic framework was largely abandoned by sociologists in favour of the approach proposed by Durkheim. In his seminal text, *Le Suicide*, Durkheim explicitly discounted the effects of imitation on suicide rates, arguing that any effect of imitation would be precipitative, minor, and local, and thus insignificant at aggregate level. Whilst the social scientific understanding of suicide has certainly evolved since Durkheim's time, theoretically informed research has tended to either develop out of his original regulation/integration thesis (e.g. Thorlindsson and Bjarnason, 1998), or limit itself to a largely interpretivist critique of his method (e.g. Douglas 1967). However more recently, *empirical* research has begun to include evidence on replication/imitation. This third approach, sometimes known as risk factor analysis, is characterised by the absence of any organising or explanatory principle (Hood-Williams 1996). More of a collection of statistical correlations between antecedent conditions and suicidality than a theory per se, risk factor analysis provides explanations for variations within and between populations that are couched in terms of differential exposure to risk factors (Charlton et al. 1993), *including social contagions*.

This theory-neutral territory is perhaps the ideal ground for developing a comprehensive memetic understanding of suicide. Recent research has tended to confirm the replication/imitation thesis, and has concerned itself with measuring the effects of this phenomenon (Gould et al. 1989 1990). Stack (1990), in a review of research results, concluded that the degree of suicide content in the mass media is indeed an important variable in accounting for varying suicide rates. Further research has been undertaken in order to demonstrate the

circumstances under which the suicide contagion is particularly infectious (Stack, 1987, Bjarnason and Thorlindsson 1994). It is now generally accepted that a spatio-temporal clustering of imitative suicidal behaviour does occur, and that the Werther effect may account for a significant proportion of youth suicides in the US (Gould et al 1989). In consequence, a US Government endorsed programme on suicide contagion has now been set up to provide guidelines and recommendations for minimising the Werther effect,[3] and a similar Government project also exists in Australia.[4]

However, as critics of Phillips' research have pointed out (e.g. Baron and Reiss 1985:361), without a credible theory of contagion or imitation upon which hypotheses may be tested, quasi-experimental research findings, such as those of Phillips, may invite ex-post reinterpretations as to the nature and existence of these putative social contagions. This means that until a comprehensive theory of replication/imitation is developed, evidence will probably be simply ignored because it doesn't 'fit' the SSSM. This, of course, is where I believe the opportunity for memetics lies, through the *theoretically informed provision* of empirical evidence of contagion and imitation from an alternative non-homuncular paradigm.

The purpose of this paper has not been to develop a memetic theory of suicide, (a central task of my DPhil thesis) but to demonstrate the pertinence of Phillips' research for the memetic paradigm methodologically, empirically and politically. Health authorities, if not academia, already accept the central tenet of our paradigm, and this important fact may provide memetics with a key opportunity for developing academic respectability and funding. A memetic approach, operationalised as *the epidemiology of social contagions in society* could build on research such as that conducted by Phillips, and could capitalise on the inability of current paradigms to adequately deal with the phenomenon of replication/imitation. Operationalising memetics as epidemiology suggests that one of the central tasks of our discipline may be to map the structural evolution and spread of social patterns of behaviour within society in terms of the familiar evolutionary loop of variation, selection, and replication. This macro-memetic approach could be complemented by, and conceptually integrated with, the micro-memetic analysis of differential infection as proposed by Lynch (1997) and Brodie (1996). These complementary approaches might provide the basis for a comprehensive analytical framework for generating a body of theoretically informed evidence relating to social contagions, which in turn might lay the foundations for the long overdue Kuhnian paradigm shift that will finally see the integration of social science within a broader evolutionary paradigm.

## Notes and References

[1] This research might perhaps draw on the research of McGuire (1964) who has developed an inoculation model in the social psychology of persuasion.

[2] Other researchers have subsequently adopted Phillips' quasi-experimental approach to demonstrate that other traits and practices may operate as social contagions transmitted via the mass media (e.g. Mazur 1982 on bomb threats)

[3] The suicide contagion thesis has impacted on public policy in the U.S. following a recent CDC (Center for Disease Control and Prevention) endorsed national workshop (CDC 1994).

[4] National Youth Suicide Prevention Strategy National Projects  
<http://www.health.gov.au/hsdd/mentalhe/nypsps/projects/13.htm>

Bandura, A., Ross, D., and Ross, S. (1963) "Imitations of Aggressive Film-Mediated Models," in *Journal of Abnormal and Social Psychology*, Vol. 66: 3-11

Barkow et al. (1992) *The Adapted Mind* New York OUP

Baron, J.N., Reiss, P.C. (1985) "Same Time Next Year: Aggregate Analyses of the Mass Media and Violent Behaviour" in *American Sociological Review* 50 (Jun): 347-363.

Brodie, R. (1996) *Virus of the Mind: The New Science of the Meme*. USA; Integral Press

Centers for Disease Control and Prevention (1988) "Recommendations for a Community Plan for the Prevention and Containment of Suicide Clusters" in *Morbidity and Mortality Weekly Review* 37(S-6): 1-12.

Centers for Disease Control and Prevention (1994) "Suicide Contagion and the Reporting of Suicide: Recommendations from a National Workshop" in *Morbidity and Mortality Weekly Review* 43 (RR-6).

Charlton, J. et al (1993) "Suicide Deaths in England and Wales 1982-1992: Trend Factors associated with Suicide Deaths" in *Population Trends* 71: 10-16.

Dennett, D.C. (1991) *Consciousness Explained*. UK. Penguin.

Douglas, J. (1967) *The Social Meaning of Suicide*

Durkheim, E. [1897] (1952) *Suicide: A Study in Sociology* London. Routledge.

Gould, M.S., Wallenstein, S., and Davidson, L. (1989) "Suicide Clusters: A Critical Review" in *Suicide and Life Threatening Behavior* 19: 17-29.

Gould, M.S. (1990) "Suicide Clusters and Media Exposure" in Blumenthal, S.J., Kupfer D.I.(Eds.), *Suicide Over the Life Cycle*. Washington, DC. American Psychiatric Press.

Habermas, J. (1979) *Communication and the Evolution of Society*. London. Heinemann

Lynch, A. (1997) *Thought Contagion: How Belief Spreads Through Society. The New Science of Memes*. Basic Books.

McGuire, W. (1964) "Inducing Resistance to Persuasion: Some Contemporary Approaches" in Berkowitz, L. (Ed.), *Advances in Experimental Social Psychology* Vol. 1: 191-229

Mazur, A. (1982) "Bomb Threats and the Mass Media: Evidence for a theory of Suggestion" in *American Sociological Review* 47 (Jun): 407-411.

Phillips, D.P. (1974) "The Influence of Suggestion on Suicide: Substantive and Theoretical Implications of the Werther Effect" in *American Sociological Review* 39 (Jun): 340-354

Phillips, D.P. (1977) "Motor Vehicle Fatalities Increase Just after a Publicised Suicide Story" in *Science* 196: 1464-65.

Phillips, D.P. (1978) "Airplane Accident Fatalities Increase Just after Newspaper Stories about Murder and Suicide" in *Science* 201: 748-750.

Phillips, D.P. (1979) "Suicide, Motor Vehicle Fatalities, and the Mass Media: Evidence towards a Theory of Suggestion" in *American Journal of Sociology* 84: 1150-1174.

Phillips, D.P. (1980) Airplane Accidents, Murder, and the Mass Media: Towards a Theory of Imitation and Suggestion in *Social Forces* 58 No. 4 (Jun): 1000-1024.

Phillips, D.P. (1982) "The Impact of Fictional Television Stories on U.S. Adult Fatalities: New Evidence on the Effect of the Mass Media on Violence" in *American Journal of Sociology* 87 No. 6: 1340-1359.

Phillips, D.P., Bollen, K.A. (1982a) "Imitative Suicides: A National Study of the Effects of Television News Stories" *American Sociological Review* Vol. 47 (Dec): 802-809

Phillips, D.P. (1983) "The Impact of Mass Media Violence on US Homicides" in *American Sociological Review* 48 (Aug): 560-568.

Phillips, D.P., Bollen, K.A., (1985) "Same Time Last Year: Selective Data Dredging for Negative Data Findings" in *American Sociological Review* Vol. 50 (Jun) 364-371

Phillips, D.P., Carstensen, L.L. (1986) "Clustering of teenage Suicides after Television News Stories about Suicide" in *New England Journal of Medicine* 315 (Sep): 685-694.

Platt, S. (1984) "Unemployment and Suicidal Behaviour" in *Social Science and Medicine* 19: 93-115

Stack, S (1990) "Media Impacts on Suicide" in Lester, D. *Current Concepts of Suicide* 107-120. Philadelphia. Charles Press

Tarde, G [1903] (1962) *Laws of Imitation*. Mass. Peter Smith

Thorlindsson, T., Bjarnason, T. (1994) "Manifest Predicators of Past Suicide Attempts in a Population of Icelandic Adolescents" *Suicide and Life Threatening Behavior* 24: 350-58

Thorlindsson, T., Bjarnason, T. (1998) "Modeling Durkheim on the Micro Level: A Study of Youth Suicidality" in *American Sociological Review* 63 (Feb) 94-110

Wasserman, I.A. (1984) "Imitation and Suicide: A Re-examination of the Werther Effect" in *American Sociological Review* 49 (Jun): 427-436

## **MEMETICS ON THE EDGE OF CHAOS**

A Review of:

**Shifting the Patterns: Breaching the Memetic Codes of Corporate Performance**  
by If Price and Ray Shaw, (1998). Management Books 2000 Ltd, UK. ISBN 1-85252-253-4

By

**Paul Marsden**

Graduate Research Centre in the Social Sciences

University of Sussex

[PaulMarsden@msn.com](mailto:PaulMarsden@msn.com)

Shifting the Patterns (STP) is the first of a new generation of books that attempts to apply a memetic stance to describe, explain, and suggest solutions to real-world issues and problems. The memetic stance involves taking a 'meme's-eyes view' of the social world, and analysing social phenomena as if human agents were vehicles for these replicating cultural traits (Marsden). The authors apply such a memetic stance to corporate culture and management theory, and in doing so develop a broadly memetic understanding of organisations.

Essentially, STP is an excursus into the idea that companies are both constructed and constrained by the patterns of beliefs and rules (memes) that produce behaviour, and ultimately determine commercial performance. To change the patterns of performance, it is necessary to change the patterns of beliefs and rules (both explicit and implicit) that determine that performance. To this end, the authors argue against the application of quick fix management fads, and for a deeper understanding of the patterns that code for performance.

According to Price and Shaw, company performance is largely the product of the unchallenged and blind application of past patterns of behaviour. The authors urge us to realise that we are not simply narrators or passive victims in this process; rather we are creative actors in the ongoing production of these patterns. We have the power to shift the patterns. The authors promote a velvet revolution in corporate culture where managers rebel against the proverbial "*tyranny of the replicators*" (Dawkins 1976/1982), see the strings by which they are moved, and take the first step to freedom (and, of course, improved performance).

STP is both engaging and stimulating, drawing together many themes of contemporary management theory and presenting them in one broadly memetic framework. In doing this, the authors demonstrate the flexibility, power and potential of the memetic stance. The book provides useful chapter by chapter summaries as well as comprehensive listings of sources for further reading. Whilst the specific recommendations in STP for corporate strategy and management theory are interesting in their own right, the focus of this review will restrict itself to an assessment to the authors' particular application of meme theory.

Importantly, STP is not a book about memetics; rather it uses the memetic stance as a vehicle for providing a critique of companies that can become mindless structures replicating themselves through time and space, and where employees can become largely passive cogs in the corporate machine. In Chapter 4: The Pattern's Eye View, the authors borrow from meme theory by

*"set[ting] out an unconventional perspective on companies, arguing that they are best interpreted as vehicles, or hosts, for a meme's - or pattern's - replication."*

Rather than adopting the conventional view that individuals are the hosts for memes, this *"unconventional perspective"* holds that companies should be understood as the proper hosts for memes. Embracing 19th Century organicism (Spencer 1969), the authors suggest that organisations can be understood *"as 'living', self-organising, self-maintaining entities"*, and these entities are the proper hosts for memes. Specifically, the meme concept is introduced as a heuristic device for understanding the components (sets of beliefs and rules) of the patterns which define the organic company.

Memes (beliefs and rules), both explicit and implicit, are held to coalesce into paradigms that can be understood as *"memomes"*, that produce the phenotypic expression of the company; its pattern and performance. This unorthodox understanding of the relationship between meme and organisation such that *"as the organism is to the gene, so the organisation is to the pattern or meme"* is not without problems, particularly when compounded with the organism/company (dis)analogy. (Organisms are not physically bonded, rather relations between their constituent parts are purely informational: Organisations themselves do not reproduce whereas organisms do: Unlike organisms, organisations have phylogenies: Organisations do not have needs themselves, unlike organisms.)

Just as the company/organisation analogy is problematic, the authors seem to exacerbate problems in the already partial gene/meme analogy by conflating 'pattern' and 'meme': It is not clear whether they understand memes as the nodes of patterns or relations between them, or indeed, both. However, the problem becomes yet more confused when the authors employ a more conventional understanding of memes, as socially inherited cultural instructions. Finally, the authors' blanket statement that memes are also the stuff of *"mental models, paradigms, languages, traditions, habits and rules; inter-related units of a cultural or economic code; widely shared patterns of perception, communication, understanding, appreciation and tradition"* stretches the meme concept to such an extent that it loses not only explanatory power and heuristic provocativeness, but also meaning. Importantly, the authors seem to confuse the issue of whether individual brains or companies are hosts to memes. In arguing that companies are both sets of rules, and hosts of those rules, the authors appear to be arguing that a set of memes somehow constitutes a host. This is deeply confusing and results in the loss of any potential clarity and coherence that the memetic stance might have.

The ambiguity over what the authors mean when they are referring to memes is further exacerbated by a tendency to switch freely between loose gene and virus analogies. Memes are held to code for organisations, whilst at other times they are presented as viral *"antibodies to change"* that hijack the organisation for their own reproduction. This confuses both the unit and the level of evolution. It is not at all clear whether rules evolve, organisations evolve, or more amorphously, patterns (of rules, behaviours and organisations) evolve. The definition of the meme as proposed by the authors does little to clear this ambiguity.

*"An organism is coded via chemical strands of DNA, sections of which form genes, the smallest units capable of being copied. An organisation is coded via 'ideas and images of the mind',*

*abstract strands of thinking, perception and language, the smallest units of which can be thought of as memes which may be interpreted as: the smallest element capable of being exchanged, with an associated sense of meaning and interpretation, to another brain."*

'Abstract strands of thinking' aside, the cogency of STP only holds together if memes are conceptualised as rules or instructions that code for an organisation. This would render the framework presented by the authors broadly compatible with Cloak's cultural ethology of i-culture and m-culture (Cloak 1975). Indeed, the authors might have usefully exploited the clarity of Cloak's framework in the development of their own position.

The authors' central tenet is that that these rules/patterns/organisations/memes/abstract-strands-of-thinking evolve selfishly to their own, rather than their human hosts', agenda, a 'fact' that is posited more by partial analogy to genes/viruses as opposed to supported by evidence. However, it is not at all clear why or how these cultural critters should evolve selfishly, since the environmental conditions (resource insufficiency) necessary for evolution to occur are not made explicit. In fact, whilst this is a central theme of the book, virtually no consideration is given to how patterns evolve within the framework developed. A simpler and more credible explanation would be to account for alienation in terms of conflicts of interests; anonymising exploitation using convenient selfish memes is a cheap trick. If the authors wished to develop a truly memetic approach to organisations, they might have usefully drawn on Allison's (1992, 1993) work on the spread of non utility-maximising rules through populations.

Whilst it is unfair to criticise STP for ideas that are not developed, it is nonetheless surprising that the authors chose not to develop, or even refer to, the leading proponents of their own understanding of organisations, that is, the neo-Darwinian theory of institutions developed by Burns and Dietz (1992, 1996, 1997a, 1997b, Dietz and Burns 1992, Dietz, Burns and Buttel 1990). Similarly, the absence of any reference to the established field of ethnomethodology (Garfinkel 1967, Zimmerman 1971) whose central concern is the understanding social behaviour in terms of the maintenance and evolution of the rules (written and unwritten) that code for that behaviour, is somewhat surprising. Had the authors chosen to develop insights from these perspectives, STP might have gained in coherence and clarity.

As it stands, the coherence of STP is apparent only when it is understood that the authors are simply borrowing elements from evolutionary thought and meme theory, and employing these elements to make a number of specific points not necessarily related to either evolution or memetics. Evolutionary and memetic concepts are simply used as metaphors full of pregnant analogies for the corporate world. It would be unfair to criticise the authors for being selective, uncritical and unrepresentative in this use of these bodies of thought, because they did not set out to be comprehensive, critical and representative in the first place. However, readers unfamiliar with the central ideas of evolutionary thought and meme theory are likely to come away from STP with a somewhat misleading understanding of the two bodies of thought. For example, the authors authoritatively state that evolution is characterised by long periods of stasis punctuated by crisis and rapid change, and that "*evolution is either revolution or it is nothing*". Perhaps, but reports of the death of gradualism have been greatly exaggerated, and Gould is not the only evolutionist around.

The most controversial aspect of STP for memeticists is probably the authors' insistence that we, ourselves, can step miraculously outside the memetic evolutionary process and, by virtue of our non-memetic independent freewill, rebel against the tyranny of the cultural critters. In other words, the dualist homunculus is left untouched by the authors' conceptualisation of memetics. Memes are considered alien parasites that infect companies (but oddly, also code for them at the same time) but which are nevertheless dependent on our willingness to replicate them. Thus, according to the authors, we can direct the evolution of companies through the selective transmission and the creation of new memes derived from our new understanding of the patterns that code for performance.

*"The key difference is that this does not have to be left to accident or random variation. We can, by human intention and creative act, wilfully change the basis upon which any pattern of company is built."*

*"People can choose to break the grip of a prevailing pattern. Punctuating organisational equilibria intentionally, shifting patterns rather than waiting for history to shift them for you, is the key to the self-generative system."*

This homuncular-speak sits uneasily with meme theory, which has largely been concerned with deconstructing pervasive homuncularism in understanding human behaviour (Blackmore forthcoming, Dawkins 1976/1982, Dennett 1990, 1991, 1995, Rose 1998, Marsden). If the authors wish to argue for empowerment and change, and against the blind following of socially inherited rules that perpetuate past patterns of individual and corporate behaviour, then they do not need memetics; indeed the use of the memetic stance is positively confusing in this respect. Neither do they need memetics to argue that companies, or indeed the rules that code for companies, evolve in ways that are alienated from the concerns of those that perpetuate them. Standard conflict theory and the introduction of the notions of power, coercion and exploitation to the analysis would make for adequate explanation of this phenomenon within the safe homuncular framework that the authors seem to wish to retain. By shying away from the deconstructive implications of the memetic stance, and by refraining from developing their theory in terms of the evolutionary loop of replication, variation and selection, STP gets a raw deal from memetics and evolutionary thought, inheriting many of the weaknesses and ambiguities, without capitalising on the strengths.

The final chapter of STP concludes with a heady cocktail of chaos theory and memetics couched in hyperbole under section headings such as "*When butterflies flap their wings: The shifting patterns of our time*", "*Slaves to, or shifters of, patterns?*", "*Where butterflies flap their wings, people infect minds!*" "*Breaking out of the cocoon: Existential fears to existential breakthroughs*", and "*Bringing forth a world: letting the genie out of the bottle*". The message here is one of empowerment, that whilst we are the products of our memetic inheritance, we also have the power to shift the patterns, to rebel against the tyranny of the memetic replicators, in a world where small changes can have major effects. This truly is memetics on the edge of chaos.

## References

- Allison, P. (1992). Cultural relatedness under oblique and horizontal transmission rules. *Ethology and Sociobiology* 13:153-169.
- Allison, P. (1993). The cultural evolution of beneficent norms. *Social Forces* 71:279-301.
- Blackmore, S. (forthcoming). *The Meme Machine*. Oxford University Press, Oxford.
- Burns, T.R. and Dietz, T. (1992). Cultural Evolution: Social Rule Systems, Selection, and Human Agency. *International Sociology*, 7: 259-283.
- Burns, T.R. and Dietz, T. (1996) Emerging Order and Disorder: Modernization and Globalization in an Evolutionary Perspective. Report, Uppsala Theory Circle, Uppsala University, Uppsala, Sweden.
- Burns, T.R. and Dietz, T. (1997a). Evolutionary Sociology: Selective Environments, Human Agency and Institutional Dynamics. Paper presented at the Conference on "Sociological Theory and the Environment", Zeist, The Netherlands, March 20-23, 1997.
- Burns, T.R. and Dietz, T. (1997b). Human Agency and Evolutionary Processes: Institutional Dynamics and Social Revolution. In: B. Wittrock (ed.), *Agency in Social Theory*. London: Sage Publications.
- Cloak, F.T. (1975). Is a cultural ethology possible? *Human Ecology* 3:161-182.
- Dawkins, R. (1976, 1982). *The Selfish Gene*. Oxford University Press, Oxford.
- Dennett, D.C. (1990). Memes and the Exploitation of Imagination. *Journal of Aesthetics and Art Criticism*, 48:??-???
- Dennett, D.C. (1991). *Consciousness Explained*. UK; Penguin.
- Dennett, D. C. (1995). *Darwin's Dangerous Idea: Evolution and the Meanings of Life*, New York: Simon and Shuster
- Dietz, T, Burns T, and Buttel F.H. (1990). Evolutionary Theory In Sociology: An Examination Of Current Thinking. *Sociological Forum* 5, 2, 155-171.
- Dietz, T. and Burns, T.R. (1992). Human Agency and the Evolutionary Dynamics of Culture. *Acta Sociologica*, 35: 187-200.
- Garfinkel, H. (1967). *Studies in Ethnomethodology*. Prentice-Hall, Englewood Cliffs.

Marsden, P. (1998). Memetics and social contagion: Two sides of the same coin? *Journal of Memetics - Evolutionary Models of Information Transmission* 2. [http://jom-emit.cfpm.org/1998/vol2/marsden\\_p.html](http://jom-emit.cfpm.org/1998/vol2/marsden_p.html)

Rose, N. (1998) Controversies in meme theory. *Journal of Memetics - Evolutionary Models of Information Transmission*, 2. [http://jom-emit.cfpm.org/1998/vol2/rose\\_n.html](http://jom-emit.cfpm.org/1998/vol2/rose_n.html).

Spencer, H (1969). The Social Organism. In D. MacRae (Ed.), *The Man Versus the State: With Four Essays on Politics and Society*. Hammondsorth. Penguin

Zimmerman, D. (1971). The practicalities of rule use. In J. Douglas (ed.), *Understanding Everyday Life*, RKP. London.

# **MEMETICS AND SOCIAL CONTAGION: TWO SIDES OF THE SAME COIN?**

**By**

**Paul Marsden**

Graduate Research Centre in the Social Sciences

University of Sussex

[PaulMarsden@msn.com](mailto:PaulMarsden@msn.com)

## **Abstract**

Following a thematic overview of social contagion research, this paper examines the question of whether this established field of social science and the nascent discipline of memetics can be usefully understood as two sides of the same coin. It is suggested that social contagion research, currently lacking a conceptual framework or organising principle, may be characterised as a body of evidence without theory. Conversely, it is suggested that memetics, now over two decades old but yet to be operationalised, may be characterised as a body of theory without evidence. The article concludes by proposing a memetic theory of social contagion, arguing that social contagion research and memetics are indeed two sides of the same social epidemiological coin, and ends with a call for their synthesis into a comprehensive body of theoretically informed research.

## **1 The Contagion Phenomenon**

Two centuries ago, a wave of suicides swept across Europe as if the very act of suicide was somehow infectious. Shortly before their untimely deaths, many of the suicide victims had come into contact with Johann von Goethe's tragic tale "The Sorrows of Young Werther," in which the hero, Werther, himself commits suicide. In an attempt to stem what was seen as a rising tide of imitative suicides, anxious authorities banned the book in several regions in Europe (Phillips 1974, Marsden 1998).

During the two hundred years that have followed the publication and subsequent censorship of Goethe's novel, social scientific research has largely confirmed the thesis that affect, attitudes, beliefs and behaviour can indeed spread through populations as if they were somehow infectious. Simple exposure sometimes appears to be a sufficient condition for social transmission to occur. This is the social contagion thesis; that sociocultural phenomena can spread through, and leap between, populations more like outbreaks of measles or chicken pox than through a process of rational choice.

The term contagion (kan-tā'jan) itself has its roots in the Latin word contagio, and quite literally means "and from touch". Contagion therefore refers to a process of transmission by touch or contact. The Microsoft Dictionary (Microsoft 1997) defines contagion as the

*"transmission of a disease by direct contact with an infected person or object; a disease or poison transmitted in this way; the means of transmission; the transmission of an emotional state, e.g. excitement; a harmful influence."*

From this definition, contagion refers to 1) the social transmission, by contact, of biological disease, and 2) the social transmission, by contact, of sociocultural artefacts or states.

The contagion concept first became popular as both a descriptive and explanatory device for social, as opposed to biological, phenomena in the late 19th century France, notably through the work of James Mark Baldwin (1894), Gabriel Tarde (1903) and Gustave Le Bon (1895).

Empirical research into the phenomenon did not, however, begin until the 1950s. This more recent research has unequivocally established the fact of the social contagion phenomenon, and has identified its operation in a number of areas of social life. The implications of this social contagion research are radical: The evidence suggests that under certain circumstances, mere 'touch' or 'contact' with culture appears to be a sufficient condition for social transmission to occur.

Despite this promising start, social contagion research has evolved into a field that is now unorganised, disparate and incoherent, lacking both an organising principle and a conceptual framework (Levy and Nail 1993). There is, in fact, a complete absence of agreement among researchers as to the particular mechanism that underlies social contagion. This lack of consensus has led to a proliferation of definitions of the phenomenon which range from the vague to the plain contradictory. For example, the Penguin Dictionary of Psychology (Reber 1995) defines contagion simply as the "*spread of an activity or a mood through a group*". The Concise Oxford Dictionary of Sociology (Marshall 1994) adopts a similarly vague definition of "*ideas moving rapidly through a group*." Other definitions, whilst more extensive, provide little in the way of increased clarity or utility. The Macmillan Dictionary of Psychology (Sutherland 1995) defines contagion as "*the spread of ideas, feelings and, some think, neuroses through a community or group by suggestion, gossip, imitation etc.*" Some definitions attempt to clarify the concept in terms of a putative uncritical and non-rational mode of inheritance/infection. Thus, The Encyclopaedic Dictionary of Psychology (Furnham 1983) defines contagion as a process and form of collective excitement "*in which emotions and behavioural patterns spread rapidly and are accepted uncritically by the members of a collective*." In contrast, other definitions make no mention of this uncritical nature of inheritance but specify instead the perception of non-intentionality in transmission (e.g. Levy and Nail 1993) such that contagion becomes the "*spread of affect, attitude or behaviour from Person A ("the initiator") to person B ("the "recipient") where the recipient does not perceive an intentional influence attempt on the part of the initiator.*" A very different definition of contagion has also been proposed, referring to neither the non-intentional nature of transmission, nor the uncritical nature of inheritance, but rather to a putative phenomenon of disinhibition. Thus, Wheeler (1966) states that:

*"If the set of test conditions T1 exists, then contagion has occurred if and only if Person X (the observer) performs behaviour N (BN) where T1 is specified as follows: a) A set of operations has been performed on Person X which is known to produce instigation toward BN in members of the class to which X belongs; b) BN exists in the response repertoire of X, and there are no physical restraints or barriers to prevent the performance of BN; c) X is not performing BN; d) X observes the performance of BN by Person Y (the model)." (p. 180)*

Together, these very different definitions of contagion have been operationalised to produce studies that have little in common except the observable phenomenon of spread by contact. Most of these insist on the presence of a number of internal states and mechanisms (intentionality, approach-avoidance, conflict etc.) for the process to count as 'true', as opposed to merely 'apparent' contagion. However, these various qualifications have not only contributed to the confused nature of social contagion research, but they have also undermined the central rationale of the metaphor; that observable culture spreads as if it has contagious properties.

One of the clearest and most inclusive definitions of social contagion is that proposed by The Handbook of Social Psychology (Lindzey and Aronsson 1985). This definition refrains from positing as necessary internal states. Instead social contagion is held to be "*the spread of affect or behaviour from one crowd participant to another; one person serves as the stimulus for the imitative actions of another.*" Such a definition has the advantage of focusing and clarifying the observable contagion phenomenon, whatever internal states may or may not be present. It should be noted however that there is no reason for the contagion phenomenon to be restricted to the crowd scenario, to be sure, the mass media allows for the possibility of contagion through dispersed collectivities.

## 2 Social Scientific Research on Social Contagion

Despite the varied definitions of contagion, the empirical research has tended to confirm that the hypothesis that human behaviour clusters in both space and time even in the absence of coercion and rationale. This tendency towards homogeneity has been identified in a number of types of behaviour using one or more of three basic approaches. In the case of dispersed collectivities or masses, evidence for and against the social contagion phenomenon has been typically drawn from correlational studies where aggregate statistics on exposure and infection are correlated, such as media reporting on suicide stories and suicide rates (e.g. Phillips 1974, Marsden 1998). In the case of local collectivities such as crowds, research methods have included field studies using participant or non-participant observation (e.g. Reicher 1984 on Bristol riots), or formal experimental studies under laboratory conditions (e.g. Freedman, Birsky and Cavoukian 1980). Bringing the disparate data from the various methods together, meta-analyses of the contagion phenomenon have also been conducted (e.g. Levy and Nail 1993).

Substantively, social contagion research can be broken down into two major areas, studies investigating emotional contagion (the spread of mood and affect through populations by simple exposure) and studies investigating behavioural contagion (the spread of behaviours through populations by simple exposure). Behavioural contagion research can itself be broken down into six broad areas, based on the nature of the behaviour that is spread; hysterical contagions, deliberate self-harm contagions, contagions of aggression, rule violation contagions, consumer behaviour contagions, and financial contagions.

A hysterical contagion is "*the dissemination of a set of symptoms among a population in which no manifest basis for the symptoms may be established.*" (Kerckhoff and Back 1968). Also known as contagious psychogenic illness (Cohen, Colligan, Wester II and Smith 1978), hysterical contagions involve the spread by contact of reported symptoms and experiences usually associated with clinical hysteria (hallucinations, nausea, vomiting, fainting etc.) in the

absence of a biological contagion. The paradigmatic example of hysterical contagion is the "June Bug" incident that occurred in a US textile factory in 1962, where 62 factory workers reported having been bitten by a mythical bug that 'caused' symptoms such as numbness and nausea (Kerckhoff and Back 1968). More recently, Colligan and Murphy (1982) have analysed a further 23 examples of hysterical contagion - "*the collective occurrence of a set of physical symptoms and related beliefs among two or more individuals in the absence of an identifiable pathogen*", and found that it was the verbal reporting of the symptoms that spread in a contagious-like manner rather than the symptoms themselves. Their research also largely confirmed Kerckhoff and Back's theory that those susceptible to hysterical contagion were suffering from intra-psychic stress. More recently still, Showalter (1997) has suggested that chronic fatigue, Gulf war and multiple personality syndromes might spread by contagion, and Pfefferbaum and Pfefferbaum (1998) have argued post-traumatic stress disorder also spreads by contagion. The hallucinatory component of hysterical contagion may also account for the spread of supernatural phenomena such as the sightings of Diana ghosts following the death of the princess in 1997 (Marsden 1997), as well as reports of UFO sightings and alien abductions (Houran and Lange 1996, Showalter 1997).

A second class of behaviour that appears to spread through populations by contagion is rule breaking or rule violation behaviour. Evidence has tended to support the thesis that an individual's exposure to rule violations increases their likelihood of engaging in similar or identical behaviour. Such rule violation contagions have been identified in teenage smoking (Ritter and Holmes 1969, Rowe, Chassin, Presson, Edwards and Sherman 1992), speeding (Connolly and Aberg 1993), substance abuse (Ennett, Flewelling, Lindrooth and Norton 1997), delinquency (Jones 1998), youth sex (Rodgers and Rowe 1993) and criminality (Jones and Jones 1995).

A third class of behaviour, which has been the focus of empirical social contagion research, is deliberate self-harm (DSH), of which suicide is the paradigmatic example. Specifically, research has shown that suicide rates and other examples of DSH vary proportionally to the extensity, intensity and content of exposure, both in local and dispersed collectivities (Phillips 1974, 1980, 1982, Stack 1987, 1990, Higgins and Range 1996, Gould 1990, 1996, Gould, Wallenstein and Kleinman 1987, Gould, Wallenstein and Davidson 1989, see Marsden 1998 for a 'memetic' overview). Contagion is now an accepted risk factor in suicide research, and the overwhelming evidence has prompted the establishment of several government programmes to minimise the effects of suicide contagion.

Another, very different focus of social contagion research has been the financial contagion phenomenon, manifested in the behaviour of stock markets which lurch from state to state as a result of selling panics and buying frenzies that sweep across the globe. Financial contagion research has tended to investigate the various factors that may exacerbate and contribute to the phenomenon such as analysis techniques, the level and nature of information available to dealers, and social communication networks (e.g. Orlean 1992, Temzelides 1997, Lux 1998).

A fifth area of contagion research has investigated the contagious properties of consumer behaviour which sometimes results in the spread of consumer fashions and fads through populations in a manner more indicative of an influenza epidemic than rational behaviour

(Marsden, forthcoming b). This phenomenon has prompted the development of deterministic and stochastic models with good predictive power that forecast both sales realisation and new product adoption patterns based on the 'infectiousness' of consumer goods (Bass, Mahajan and Muller 1990, Rashevsky 1939, 1951, Rapoport 1983, Rogers 1995).

A sixth focus of social contagion research has been the contagion of aggressive behaviour, a phenomenon that has been shown to operate in both local and dispersed collectivities. Whilst much of this research has been of a descriptive nature within transitory and unpredictable angry crowds (mobs) (Bandura 1973, Reicher 1984, Lachman 1996), results have been supported with experimental evidence (Bandura, Ross and Ross 1963, Wheeler and Caggiula 1966, Wheeler and Levine 1967, Wheeler and Smith 1967, Goethals and Perlstein 1978). In dispersed collectivities, where the contagion of aggression is mediated by the mass media, research has focused on measuring exposure and infection rates and then testing for correlations. (Atkin, Greenberg, Korzenny, and McDermott 1979, Sheehan 1983, Phillips 1983).

Finally, social contagion research has not only restricted itself to the spread of behaviours, a significant number of studies have identified a variety of emotional contagions. The emotional contagion phenomenon was originally defined by McDougall (1920) as "*the principle of direct induction of emotion by way of the primitive sympathetic response*" and more recently by Sullins (1991) as "*the process by which individuals seem to catch the "mood" of those around them*". The proposed mechanism for this spread of mood is an automatic and continuous human tendency to synchronise facial expressions, voices, and postures with others in the immediate environment (Hatfield, Cacioppo and Rapson 1993). These behavioural cues then appear to trigger the appropriate emotions in a system of feedback. The Emotional Contagion Scale (Doherty 1997) has been recently developed and validated to assist further research in this area, which has already identified various examples of emotional contagion including mood (Hsee, Hatfield and Chemtob 1992), anxiety (Behnke, Sawyer and King 1994), fear (Gump and Kulik 1997), appreciation (Freedman and Perlick 1979) and enjoyment (Freedman et al. 1980).

### **3 The Social Contagion Phenomenon Explained (Away)?**

Whilst the vast majority of social contagion research has demonstrated the existence and voracity of the empirical phenomenon, the theoretical implications of the results have not been addressed. The results of contagion research suggest that just as we do not choose to be infected with, and pass on, biological contagions, we often behave as if we have little control over the culture we become infected with and consequently spread. Such an observation undermines the traditional understanding of the human subject as an autonomous agent whose action is defined by individual intentionality and rational evaluation. Whilst we may like to believe that we consciously and rationally decide on how to respond to situations, social contagion evidence suggests that some of the time this is simply not the case. Rather than generating and 'having' beliefs, emotions and behaviours, social contagion research suggests that, in some very real sense, those beliefs, emotions and behaviours 'have' us.

The failure of mainstream social science to take this implication of social contagion evidence seriously is certainly in part due to the above-mentioned disorganised and incoherent state of the field. However, the failure is also probably due to a fundamental incompatibility between the

concept of social contagion and the Cartesian voluntarism implicit in much social science. In fact, standard explanations of social contagion can be characterised by an almost desperate attempt to restore irreducible individual agency and rational action to the phenomenon. In trying to explain away the social contagion phenomenon, two types of theory have been developed. Firstly, a number of theories suggest that the spread of homogeneity is a consequence of conscious and deliberate imitation in situations usually defined by uncertainty or ambiguity. Secondly, contagion has been accounted for by putative latent homogeneities in terms of prior motivations that antecede the observable phenomenon.

An example of the first type of explanation is Emergent Norm Theory (e.g. Turner 1964) which states that the spread of behaviour through a population is not by contagion (contact) but is the result of conscious and deliberate attempts to adhere to norms and rules emerging out of complex and subtle interaction within collectivities. Similarly Social Learning Theory (e.g. Bandura 1971, 1986) holds that homogeneity is the result of the conscious and deliberate imitation that takes place when individuals are presented with uncertain and ambiguous situations. When we are unsure of how to react to a stimulus or a situation, these theories suggest that we actively look to others for guidance and consciously imitate them.

An example of the second type of explanation is Convergence Theory (e.g. Turner and Killian 1987) which suggests that homogeneity and clustering is not a result of contagion but the result of prior shared motivations that cause collectivities to converge in the first place. From this perspective, similarities cause collectivities and not vice versa. A similar explanation is provided by Disinhibition Theory (e.g. Freud 1922, Redl 1949, Wheeler 1966, Ritter and Holmes 1969, Levy and Nail 1993) which states that contagion is "*essentially imitation mediated by restraint release due to observing another perform an action that the individual is in conflict about performing himself*" (Freedman 1982). In other words, from this perspective, behaviours are not transmitted by contact; rather inhibited behaviours (sometimes unconscious and "primitive") that are already held in an individual's behavioural repertoire are simply released. Thus, homogeneity spreads as a result of the intra-psychic conflict resolution that occurs through social evidence. Another variation on the 'prior motivations' theme is Deindividuation Theory (e.g. Diener 1976, 1979, Festinger, Pepitone and Newcombe 1952, Zimbardo 1969). This theory holds that the anonymous nature of collectivities can engender a restraint reduction in individuals. This sense of anonymity is held to cause a reduction in the individual's sense of personal accountability and responsibility, allowing them to engage in behaviour from which they might otherwise abstain. When anonymity leads to restraint reduction of similar behaviours within individuals within a collectivity, this produces the appearance of contagion.

Both the 'conscious choice' and 'prior motivations' theories may explain the social contagion phenomenon in some circumstances, but none of them can comprehensively explain the phenomenon. Indeed, it is difficult to see how any of the theories could provide a credible explanation of either emotional or hysterical contagion, except by maintaining that we either choose illnesses or emotional states based on those that are around us, or worse we have hidden desires to be ill, angry or anxious! Social contagion stretches Cartesian rational action theory to such a degree that the latter becomes an untenable explanation of the former. Valiant attempts at squeezing irreducible individual agency and rational evaluation into the phenomenon are simply at odds with data. The evidence shows that we inherit and transmit behaviours, emotions, beliefs

and religions not through rational choice but contagion. Does this rejection of rational choice/action theories mean that social contagion is a homeless body of research, a body of evidence with no theoretical home to go to? No, I think there is an alternative paradigm that has the potential to explain more of the data more of the time. That paradigm is memetics.

## 4 The Memetic Stance

Social theorists often use the language of architecture, they speak of theory building, laying theoretical foundations, or constructing theoretical edifices. This is useful language; it indicates the step by step, laborious nature of their enterprise. Meme theory is no different in this respect; many problems still have to be resolved within the new paradigm (Rose 1998). However, meme theory is developed enough to be operationalised conservatively by adopting what could reasonably be called a memetic stance. Not a fully blown theory, the memetic stance is more of a way of looking at the world, a set of guiding principles, a useful heuristic, based on some hopefully important insight into the nature of the social world. Whether the memetic stance turns out to be an explanatory device in an evolutionary extension of folk psychology, or a proper theory of mind where memes are internally instantiated in the neural networks of our brains, is an issue that will one day have to be resolved empirically. For now, by adopting the memetic stance, these issues may be bracketed, and research can proceed based on the utility of this ontologically minimalist heuristic.

So what exactly is the memetic stance? The memetic stance states that human condition is minimally defined by two selective processes operating in two different substrates, the biological and social (Marsden forthcoming a). This is because the necessary conditions for the evolutionary loop of replication, variation and selection are present in the two substrates. This is not contentious in itself, what is more contentious is that the memetic stance sees these processes operating at the level of what is being replicated, that is, the gene and the meme. Thus, the memetic stance involves taking a meme's-eye perspective and understanding of the social world, thinking not in terms of selfish genes, but selfish memes. Taking this memetic stance has allowed researchers to explain the spread of non-rational behaviour in terms of the fitness of that behaviour itself. Examples include altruism (Allison 1992, 1993, Blackmore forthcoming) chain letters (Goodenough and Dawkins 1994, Hofstadter 1995, Allison 1993), chain e-mail (Jones 1995) religions and cults (Dawkins 1993, Lynch 1996 and Cowley 1997), political revolutions and war (Vajk 1989), religious scriptures (Piper 1997) Usenet content (Best 1997), management practices (Price and Shaw 1996, 1998), media representations, (Rushkoff 1994), urban legends (Gross 1996) and consumer behaviour (Marsden forthcoming b, Brodie 1996).

The memetic stance suggests that design in the social world is at least partly a product of the evolutionary loop of replication, variation and selection operating on culture, or more specifically cultural instructions coding for behaviour (Cloak 1975, Marsden forthcoming a). It is not necessary to invoke conscious choice or rational evaluation by an entity - homuncular, divine or otherwise - standing miraculously outside evolution to explain design; given enough iterations, natural selection will inexorably and inevitably give rise to design.

Once we take the memetic stance, features of the world that are difficult to explain from the orthodoxy of traditional social science become non-miraculous and eminently explicable. The

memetic stance can explain not only apparent design in the social world, but importantly it can also explain phenomena that seem to negate the omnipresence of individual agency in human affairs. Put simply, the memetic stance states that the reason why some social behaviour doesn't seem to make sense from the perspective of the individual is because we are looking at that behaviour at the wrong level. We are taking an anthropocentric or homuncular view of a social world that was created at least in part at a memetic level. Trying to explain the social world from the perspective of the individual is like trying to explain the movements of a car without reference to the driver. The movements of a car can be rationally described, explained and understood in terms of the car's own needs as (somewhat circuitous) trips from petrol pump to petrol pump. However, by ignoring the driver much of what is observed makes no sense at all. The same argument holds for the social world, just as we can explain much of our (somewhat circuitous) social behaviour in terms of the needs of the meme-vehicle (individual), much of what is interesting about that behaviour is overlooked. By taking the memetic stance we can account for what happens when the needs of an individual cannot explain behaviour, the equivalent of all the non-petrol seeking activity of a car, and this stance provides an evolutionary rationale for explaining why the social contagion phenomenon occurs.

## 5 Social Contagion from the Memetic Stance

Taking the memetic stance involves, to use an overused concept, a true Kuhnian paradigm shift; just as evolution in the biological world evolves according to what is better (not best) for the gene in its environment, so too does the social world evolve according to what is better for the meme. The memetic stance involves describing, explaining and understanding social behaviour from this meme's-eye perspective. From the memetic stance "What makes this person want to do x?" becomes "What is it about x that makes people want to do it?" Social contagion can be explained by the memetic stance because culture has an independent evolutionary dynamic that is derived from the genetically evolved human capacity and predisposition to replicate culture (see Flinn 1997 for a review). Because social learning is an evolved psychological trait, it follows that we have an evolved predisposition to replicating the behaviour of those around us. Successful social contagions are those elements of culture that operate as both stimulus and response, and that are adapted to the evolved architecture of the human brain. No homunculus need be invoked, only evoked imitation.

In this way, the memetic stance deconstructs the homunculus into what can be understood as replicating cultural instructions (memes). This opens up an exciting research programme for memetics, as contagion is no longer understood as a metaphor but an evolutionary process. Social contagion research, from the memetic stance could focus on the particular characteristics that render behaviours and emotions infectious. It could also investigate why certain people are immune to certain contagions, or how they develop resistance to contagion, or conversely what makes certain people particularly susceptible to contagion, and others not. Similarly, memetic research could look for the limiting factors of the contagion phenomenon in both time and space. How, for example, is social contagion bounded? Can social contagion epidemics burn themselves out and if so, how? How does a contagious epidemic become an endemic trait in the social world? Is it possible to quarantine areas exposed to contagion, or quarantine those who have been infected? Can individuals be vaccinated against contagion? How long is the incubation period, that is, the time from exposure to infection? What are the primary vectors of

contagions, that is, what are the primary channels of infection? Are contagions specific or diffuse? These are all questions that are more or less precluded in a traditional paradigm dominated by a Cartesian homuncularism and rational action theory which essentially deny the existence of the social contagion phenomenon. By deconstructing the homunculus into a web of replicating instructions, the memetic stance allows the social contagion phenomenon to become a theoretically informed research enterprise.

## **6 Memetics and Social Contagion: Two Sides of the Same Coin?**

Taking the memetic stance allows research to proceed with the objective of explaining the spread of non-rational behaviour in terms of the fitness of that behaviour itself. Until recently, such memetic research has been of a largely non-rigorous and anecdotal nature. Despite offering the exciting prospect of being an autonomous social theory that is compatible and coherent with, but not reducible to, our knowledge of the biological world, the emerging discipline of memetics has yet to produce any concrete results. This is essentially because the issue of how to successfully operationalise the emerging paradigm has yet to be addressed; memeticists have yet to exploit their innovative analytical framework to build a body of theoretically informed empirical research. It is here that the body of social contagion research may be of particular use to memetics, offering itself up as a rich source of empirical evidence, whilst offering important methodological lessons and inspiration for future research. For example, the emotional contagion scale developed by Doherty (1997) could be used by memeticists, as could the field studies, correlational and experimental methods that have been exploited by social contagion researchers.

More generally, memeticists could develop the social contagion research tradition of using the substrate neutral tools of epidemiology to assist their research programme. These tools could be adapted to provide useful information about differential incidence and prevalence of evolving cultural traits, as well as structure of endemic and epidemic features of society.

The use of epidemiological tools would have the advantage of allowing memetic research to proceed without making any ontological claims as to the nature or status of what exactly is being spread. Epidemiology is not the study of the inheritance of particular diseases or pathologies per se; rather it is the study of the distribution and pattern of the measurable effects of infection. Memetics qua social epidemiology might aspire to a similar goal. In the same way that causal mechanisms in the epidemiology of disease depend on, and vary with, the particular pathology that is being studied, taking the memetic stance does not require that social patterns be reduced to any one particular selective mechanism. Epidemiology may proceed independently of the aetiology of social products being researched; no assumptions about the heterogeneity or homogeneity of causal mechanisms are necessary.

## **7 Conclusion**

The emerging paradigm of memetics and the established tradition of social contagion research do not simply have much to learn from each other, they are in fact two sides of the same social epidemiological coin, the former a theory-rich version of the latter, and the latter an evidence-rich version of the former. Taking the memetic stance is a radical move, but there are some wheels, particularly of the methodological variety, that just don't need to be reinvented by

memeticists, because they can already be found in social contagion research. Equally, memetics brings to the social contagion table an innovative conceptual framework with an important evolutionary component that the latter currently lacks. By integrating social contagion research and the memetic paradigm we would allow for the development a robust body of theoretically informed empirical research. In doing this we will be laying one more foundation for the long overdue Kuhnian paradigm shift that will finally see the integration of social science within a broader evolutionary paradigm.

## Acknowledgements

The author would like to thank members of 'Meme Lab', Sue Blackmore, Nick Rose and Derek Gatherer, for discussing and reviewing earlier drafts of this article and for their useful constructive criticism. The usual qualifier of course applies.

## References

- Allison, P. (1992). Cultural relatedness under oblique and horizontal transmission rules. *Ethology and Sociobiology*, 13:153-169.
- Allison, P. (1993). The cultural evolution of beneficent norms. *Social Forces*, 71:279-301.
- Atkin, C.K., Greenberg, B.S., Korzenny, F., and McDermott, S. (1979). Selective exposure to televised violence. *Journal of Broadcasting*, 23:5-13.
- Baldwin, J. M. (1894). Imitation: A chapter in the natural history of consciousness. *Mind*, 3:25-55.
- Bandura, A. (1971). *Social Learning Theory*. New York: General Learning Press.
- Bandura, A. (1973). *Aggression: A Social Learning Analysis*. NJ: Prentice-Hall.
- Bandura, A. (1986). *Social Foundations of Thought and Action*. Engelwood Cliffs, NJ: Prentice-Hall.
- Bandura, A., Ross, D., and Ross, S. (1963). Imitations of aggressive film-mediated models. *Journal of Abnormal and Social Psychology*, 66:3-11.
- Bass, F.M., Mahajan, V., and Muller, E. (1990). New product diffusion models in marketing: A review and directions for research. *Journal of Marketing*, 54:1-26.
- Behnke, P.R., Sawyer, C.R., and King, P.E. (1994). Contagion theory and the communication of public speaking state anxiety. *Communication Education*, 43:246-251.
- Best, M.L. (1997). Models for interacting populations of memes: Competition and niche behaviour. *Journal of Memetics - Evolutionary Models of Information Transmission*, 1. [http://jom-emit.cfpm.org/1997/vol1/best\\_ml.html](http://jom-emit.cfpm.org/1997/vol1/best_ml.html).

- Blackmore, S. (forthcoming). *The Meme Machine*. Oxford: OUP.
- Brodie, R. (1996). *Virus of the Mind: The New Science of the Meme*. Seattle, WA: Integral Press.
- Cloak F.T. (1975). Is a cultural ethology possible? *Human Ecology*, 3:161-182.
- Cohen, B.G.F., Colligan, M.J., Wester II, W., and Smith, M.J. (1978). An investigation of job satisfaction factors in an incident of mass psychogenic illness at the workplace. *Occupational Health Nursing*, January 1978:10-16.
- Colligan, M.J., and Murphy, L.R. (1982). A review of mass psychogenic illness in work settings. In M.J. Colligan, J.W. Pennebaker, and L.R. Murphy (eds.) *Mass Psychogenic Illness*, NJ: Erlbaum, 171-182.
- Connolly, T. and Aberg, L. (1993). Some contagion models of speeding. *Accident Analysis and Prevention*, 25:57-66.
- Cowley, G (1997). Viruses of the mind: How odd ideas survive. *Newsweek* April 14:14.
- Dawkins, R. (1993). Viruses of the mind. In B. Dahlbohm (ed.) *Dennett and his Critics*, Oxford: Blackwell Publishers, 13-27.
- Diener, E. (1976). Effects of prior destructive behavior, anonymity, and group presence on deindividuation and aggression. *Journal of Personality and Social Psychology*, 33:462-470.
- Diener, E. (1979). Deindividuation, self-awareness and disinhibition. *Journal of Personality and Social Psychology*, 37:1160-1171.
- Doherty, R.W. (1997). The Emotional Contagion Scale: A measure of individual differences. *Journal of Nonverbal Behavior*, 21:131-154.
- Ennett, S.T., Flewelling, R.L., Lindrooth, R.C., and Norton, E.C. (1997). School and neighborhood characteristics associated with school rates of alcohol, cigarette, and marijuana use. *Journal of Health and Social Behavior*, 38:55-71.
- Festinger, L., Pepitone, A., and Newcombe, T. (1952). Some consequences of deindividuation in a group. *Journal of Abnormal and Social Psychology*, 47:382-389.
- Flinn, M.V. (1997). Culture and the evolution of social learning. *Evolution and Human Behavior*, 18:23-67.
- Freedman, J.L. (1982). Theories of Contagion as they relate to mass psychogenic illness. In M.J. Colligan, J.W. Pennebaker, and L.R. Murphy (eds.) *Mass Psychogenic Illness*, N.J: Erlbaum, 171-182.
- Freedman, J.L., and Perlick, D. (1979). Crowding, contagion and laughter. *Journal of Experimental Psychology*, 15:295-303.

Freedman, J.L., Birsky, J. and Cavoukian, A. (1980). Environmental determinants of behavioral contagion: Density and number. *Basic and Applied Social Psychology*, 1:155-161.

Freud, S. [1922] (1959). *Group Psychology and the Analysis of the Ego*. London: Norton.

Furnham, A. (1983). Contagion. In R. Harre and R. Lamb (eds.) *The Encyclopaedic Dictionary of Psychology*, 119.

Goethals, G.R. and Perlstein, A.L. (1978). Level of instigation and model similarity as determinants of aggressive behavior. *Aggressive Behavior*, 4:115-124.

Goodenough, O. R. and Dawkins, R. (1994). The "St. Jude" mind virus. *Nature*, 371:23-24.

Gould, M.S., Wallenstein, S., and Kleinman, M. (1987). *A study of time-space clustering of suicide. Final report*. Atlanta, Georgia: Centers for Disease Control.

Gould, M.S., Wallenstein, S., and Davidson, L. (1989). Suicide clusters: A critical review. *Suicide and Life Threatening Behavior*, 19:17-29.

Gould, M.S. (1990). Suicide clusters and media exposure. In S.J. Blumenthal and D.I. Kupfer (eds.), *Suicide Over the Life Cycle*. Washington, DC: American Psychiatric Press.

Gould M.S. (1996). Suicide Contagion. At American Foundation for Suicide Prevention Website. <http://www.afsp.org/research/gould.html>.

Gross, D. (1996.) The Blue Star Meme: Applying Natural Selection Thinking to Urban Legends. <http://www.lycaeum.org/drugs/other/tattoo/meme.html>.

Gump, B.B., and Kulik, J.A. (1997). Stress, affiliation and emotional contagion. *Journal of Personality and Social Psychology*, 722:305-319.

Hatfield, E., Cacioppo, J.T. and Rapson, R.L. (1993). *Emotional Contagion*. Cambridge: CUP.

Higgins, L and Range, L.M. (1996). Does information that a suicide victim was psychiatrically disturbed reduce the likelihood of contagion? *Journal of Applied Social Psychology*, 26:781-785.

Hofstadter D.R. (1985). *Metamagical Themes: Questions for the Essence of Mind and Pattern*. New York: Basic Books.

Houran, J., and Lange, R. (1996). Hauntings and poltergeist-like episodes as a confluence of conventional phenomena: A general hypothesis. *Perceptual and Motor Skill*, 83:1307-1316.

Hsee, C.K., Hatfield, E., and Chemtob, C. (1992). Assessments of the emotional states of others - conscious judgements versus emotional contagion. *Journal of Social and Clinical Psychology*, 11:119-128.

- Jones, L (1995). Good Times Virus Hoax. <ftp://usit.net/pub/lesjones/GoodTimes-HoaxFAQ.txt>.
- Jones, M.B. (1998). Behavioral contagion and official delinquency: Epidemic course in adolescence. *Social Biology*, 45:134-142.
- Jones, M.B., and Jones, D.R. (1995). Preferred pathways of behavioural contagion. *Journal of Psychiatric Research*, 29:193-209.
- Kerckhoff, A.C. and Back, K.W. (1968). *The June Bug: A Study in Hysterical Contagion*. New York: Appleton-Century-Crofts.
- Lachman, S.J. (1996). Psychological perspective for a theory of behavior during riots. *Psychological Reports*, 79:739-744.
- Le Bon, G. [1895] (1903). *The Crowd: A Study of the Popular Mind*. London: T. Fisher Unwin.
- Levy, D.A., and Nail, P.R. (1993). Contagion: A theoretical and empirical review and reconceptualization. *Genetic, Social and General Psychology Monographs*, 119:235-183.
- Lindzey, G. and Aronson, E. (eds.) (1985). *Handbook of Social Psychology: Group Psychology and the Phenomena of Interaction (3rd Ed.)*, New York: Random House.
- Lux, T (1998). The socio-dynamics of speculative markets: interacting agents, chaos, and the fat tails of return distribution. *Journal of Economic Behavior and Organization*, 33:143-165.
- Lynch A. (1996). *Thought Contagion. How Belief Spreads Through Society*. The New Science of Memes. New York: Basic Books.
- Marsden, P.S. (1997). Crash contagion. Unpublished Paper delivered at University of Sussex conference 'Death and Diana' [http://members.tripod.com/~Paul\\_Marsden/Crash-Contagion.html](http://members.tripod.com/~Paul_Marsden/Crash-Contagion.html)
- Marsden, P.S. (1998). Operationalising memetics - Suicide, the Werther effect, and the work of David P. Phillips. Proceedings of the 15th International Congress on Cybernetics, Namur, Belgium.
- Marsden, P.S. (forthcoming a) The selectionist paradigm: More implications for sociology. *Sociological Research Online*.
- Marsden, P.S. (forthcoming b). Memetics as a new paradigm for understanding customer behaviour. *Marketing Intelligence and Planning*.
- Marshall, G. (ed.) (1994). *Concise Oxford Dictionary of Sociology*. Oxford: OUP.
- McDougall, W. (1920). *The Group Mind*. New York: Knickerbocker Press.

Microsoft (1997). Microsoft Bookshelf Basics; *Chambers Dictionary in Microsoft Office 1997* (Professional Edition). <http://www.microsoft.com/office>

Orlean, A. (1992). Contagion of opinion in financial markets. *Revue Economique*, 43:685-698.

Pfefferbaum, B. and Pfefferbaum, R.L. (1998). Contagion in stress - An infectious disease model for posttraumatic stress in children. *Child and Adolescent Psychiatric Clinics of North America* 7:183.

Phillips, D.P. (1974). The influence of suggestion on suicide: Substantive and theoretical implications of the Werther effect. *American Sociological Review*, 39:340-354.

Phillips, D.P. (1980). Airplane accidents, murder, and the mass media: Towards a theory of imitation and suggestion. *Social Forces* 58:1000-1024.

Phillips, D.P. (1982). The impact of fictional television stories on U.S. adult fatalities: New evidence on the effect of the mass media on violence. *American Journal of Sociology*, 87:1340-1359.

Phillips, D.P. (1983). The impact of mass media violence on U.S. homicides. *American Sociological Review*, 48:560-568.

Price, I. and Shaw, R. (1996). The learning organisation meme: Emergence of a management replicator or Parrots, patterns and performance. <http://members.aol.com/ifprice/ppatperf.html>

Price, I. and Shaw, R. (1998). *Shifting the Patterns: Breaching the Memetic Codes of Corporate Performance*. UK: Management Books 2000.

Pyper, H.S. (1997). The selfish text: The Bible and memetics.  
<http://www.lycaeum.org/~sputnik/Memetics/bible.html>

Rapoport, A. (1983). *Mathematical Models in the Social and Behavioral Sciences*. New York: John Wiley and Sons.

Rashevsky, N. (1939). *Mathematical Biophysics*. Chicago: University of Chicago Press.

Rashevsky, N. (1951). *Mathematical Biology of Social Behavior*. Chicago: University of Chicago Press.

Reber, A.S. (ed.) (1995). *The Penguin Dictionary of Psychology* (2nd ed.). London: Penguin.

Redl, F (1949). The phenomenon of contagion and "shock effect" in group therapy. In K.R. Eissler (ed.), *Searchlights on Delinquency*, New York: International Universities Press, 315-128.

Reicher, S.D. (1984). The St. Paul's riot: An explanation of the limits of crowd action in terms of a social identity model. *European Journal of Social Psychology*, 14:1-21.

- Ritter, E.H. and Holmes, D.S. (1969). Behavioral contagion: Its occurrence as a function of differential restraint reduction. *Journal of Experimental Research in Personality*, 3:242-246.
- Rodgers, J.L. and Rowe, D.C. (1993). Social contagion and adolescent sexual-behavior - a developmental EMOSA model. *Psychological Review*, 100:479-510.
- Rogers, E. M. (1995). *The Diffusion of Innovations* (4th ed). New York: The Free Press.
- Rose, N. (1998). Controversies in meme theory. *Journal of Memetics - Evolutionary Models of Information Transmission* 2. [http://jom-emit.cfpm.org/1998/vol2/rose\\_n.html](http://jom-emit.cfpm.org/1998/vol2/rose_n.html).
- Rowe, D.C., Chassin, L., Presson, C.C., Edwards, D. and Sherman, S.J. (1992). An epidemic model of adolescent cigarette-smoking. *Journal of Applied Social Psychology*, 22:261-285.
- Rushkoff, D., (1994). *Media Virus*. New York: Ballantine.
- Sheehan, P.W. (1983). Age trends and the correlates of children's television viewing. *Australian Journal of Psychology*, 35:417-431.
- Showalter, E. (1997). *Hystories*. New York: Columbia University Press.
- Stack, S. (1987). Celebrities and suicide: A taxonomy and analysis, 1948-1983. *American Sociological Review*, 52:401-412.
- Stack, S. (1990). Media impacts on suicide. In D. Lester (ed.), *Current Concepts of Suicide*. Philadelphia: Charles Press, 107-120.
- Sullins, E.S. (1991). Emotional contagion revisited: Effects of social comparison and expressive style on mood convergence. *Personality and Social Psychology Bulletin*, 17:166-174.
- Sutherland, S. (ed.) (1995). *Macmillan Dictionary of Psychology*. London: Macmillan Press.
- Tarde, G. [1903] (1963). *The Laws of Imitation*. Mass: Peter Smith.
- Temzelides, T. (1997). Evolution, coordination, and banking panics. *Journal of Monetary Economics* 40:163-183.
- Turner, R.H. (1964). Collective behaviour. In R.E.L. Faris (ed.) *Handbook of Modern Sociology* Chicago: Rand McNally, 382-425.
- Turner, R.H. and Killian, L.M.(1987). *Collective Behavior* (3rd ed.) NJ: Prentice-Hall.
- Vajk, P. (1989). Memetics: Ideas and their transmission. Unpublished paper presented to Outlook at <http://www.geocities.com/Athens/2424/memetics.html>.

Wheeler, L. (1966). Towards a theory of behavioural contagion. *Psychological Review*, 73:179-192.

Wheeler, L. and Caggiula, A.R. (1966). The contagion of aggression. *Journal of Experimental Social Psychology*, 2:1-10.

Wheeler, L. and Levine, L. (1967). Observer-model similarity in the contagion of aggression. *Sociometry*, 30:41-49.

Wheeler, L. and Smith, S. (1967). Censure of the model in the contagion of aggression. *Journal of Personality and Social Psychology*, 6:93-98.

Zimbardo, P.G. (1969). The human choice: Individuation, reason and order versus deindividuation, impulse and chaos. In N.L. Arnold and D. Levine (eds.), *Nebraska Symposium on Motivation*, Lincoln: University of Nebraska Press, 237-308.

# **A STRATEGY FOR MEMETICS: MEMES AS STRATEGIES**

## **A Commentary on Gatherer's paper: Why the 'Thought Contagion' Metaphor is Retarding the Progress of Memetic**

**By**

**Paul Marsden**

Graduate Research Centre in the Social Sciences

University of Sussex

PaulMarsden@msn.com

Gatherer's (1998) demolition tour of mental memetics and the thought contagion metaphor is convincing and is to be applauded. Memetics organised around the interactions of thought contagions (e.g. Lynch 1996, 1998) is problematic because that interaction is neither observable nor measurable. This simple fact renders the thought contagion metaphor both conceptually vacuous and empirically redundant. It is certainly not by performing arbitrary mathematical manipulations and endowing the objects of introspection with contagious properties that will improve the status of a partial and misleading metaphor.

As Gatherer notes, the thought contagion metaphor persists largely because of what could charitably be called a 'disregard' for linguistics and cognitive science. To this I would add a dangerous 'disregard' for the corpus of social science in general, *and evolutionary social science in particular* (Dietz, Burns and Buttel 1990, Flinn 1997, Durham 1990, Sanderson 1990, and Runciman 1998 provide good primers on this research area).

At least two further elementary points can be added to demonstrate why the thought contagion metaphor should be definitively buried. First, a 'thought' (if memeticists could ever isolate a single thought) has meaning, like a letter in a word, only by virtue of its specific position within an environmental context. Therefore, there can be no general laws of combination and permutation. Secondly, the essentialism implicit in the thought contagion metaphor is incompatible with the very selectionist paradigm within which memetics is grounded (Palmer and Donahue 1992).

Dawkins has recently declared that he is "alarmed" that his readers have taken his memetic speculations at face value as a theory of culture (Dawkins forthcoming). Dennett (1995, 1998), Plotkin (1997) and a number of leading evolutionary psychologists (e.g. Pinker 1998, Flinn 1997) have also made known their scepticism as to the scientific viability of the whole memetic enterprise. Of course, the fact that the founding fathers of memetics have doubts about memetics, qua scientific enterprise, is no reason in itself to abandon research. But surely we should be addressing their specific concerns, and those of other evolutionary theorists, rather than burying our heads in outmoded pseudoscientific introspectionism (see Rose 1998 and Clewley 1998 for an exposition of some of problems underlying these concerns).

A further more practical point, not underlined by Gatherer, is that meme theory built on the thought contagion metaphor is simply bad strategy for the establishment of memetics as a respectable and respected research enterprise. If memetics wishes to be taken seriously, it must first explicitly address the established theories that already seek to explain the object of

memetics' own enquiry and demonstrate *theoretically and empirically* how the new memetic approach is superior. Growing memetics in an intellectual vacuum on a staple diet of partial metaphors is not only foolhardy, it is intellectual suicide. An alternative and more successful strategy might be to emulate the one taken in the establishment of evolutionary psychology (e.g. Barkow, Cosmides and Tooby 1992) as a credible alternative to the Standard Social Scientific Model. Current paradigms were attacked and undermined, their weaknesses exposed, and then an alternative proposed which was then tested to demonstrate its utility.

For Gatherer, a viable strategy open to memetics would be a return to the 'Dawkins A' conceptualisation of a meme, and to adopt a research programme broadly similar to that of the established field of social contagion research. (Note that Gatherer is *not* advocating the rejection of the contagion part of the thought contagion metaphor, indeed there is a wealth of evidence supporting the behavioural contagion hypothesis; Marsden 1998b, 1998c, Levy and Nail 1993). In other words, by focusing on the epidemiology of behaviours, which would of course include verbal behaviour, Gatherer is advocating a reconceptualisation of memetics as behavioural ecology.

Whilst I agree with Gatherer's diagnosis, and whilst I also agree that memetics has much to learn methodologically from social contagion research (Marsden 1998b), I find his prescription for memetics ultimately unsatisfactory. It is not that there would be anything intrinsically wrong with such an approach, it is merely that it is difficult to see what it would offer in terms of increased understanding of the social world. Gatherer's approach would reduce memetics to an, albeit innovative, counting exercise. To be sure, if trends and correlations were to be identified over time, such an approach would have explanatory and predictive power, as I have previously shown elsewhere with the case of suicidal behaviour and media representations of suicide (Marsden 1998a). However, non-memetic research can generate the same correlations and predictions quite adequately without the application of the selectionist paradigm (e.g. Phillips 1980), and it is unclear what Gatherer's behavioural stance would add to these approaches. Further, whilst counting behaviours may be more scientific than counting thoughts, neither is particularly useful in a social world which is defined by structure and context.

Gatherer asks at the end of the 'focus' whether there is a viable third approach for memetics, to which I would reply in the affirmative. As I assume Lynch and Gatherer would agree, memetics involves the application of the selectionist paradigm to the sociocultural world. However, unlike Lynch and Gatherer, I maintain that the appropriate unit of analysis in this application should be neither the 'thought' nor the 'behaviour', but rather the *strategy*. A meme conceptualised as a culturally transmitted behavioural strategy has a number of advantages.

First of all, an approach characterised by a focus on the differential replication of strategies would allow for a more comprehensive memetic theory of culture. Social learning and socialisation cannot be reduced to localised imitative behaviour; strategies are copied and executed differentially across space and time. Replicating (conditional) strategies, rather than behaviours, could account for such differential and contextual behaviour (see Cloak 1975 and Cohen and Machelek 1988 for early attempts at such a project).

Secondly, by conceptualising memes in terms of functional and conditional strategies, a possible path between thorny methodological and ontological problems can be navigated. A strategy is an abstract concept, a heuristic device that (unlike thoughts) may be taken to link observable stimuli and responses. By focusing on the populational dynamics of strategy selection, memetics could leave open the question of how strategies are actually instantiated in the brain, yet still retain explanatory power and falsifiability by employing a methodological (as opposed to ontological) social behaviourist stance. All such an ontologically minimalist approach would involve would be the positing of a heuristic device (strategies) to explain and predict behaviour; this is what I call taking the memetic stance (Marsden 1998b).

Thirdly, meme theory thus conceived would allow for a theory of institutionalisation and structure emerging non-miraculously out of the reciprocal execution of strategies (see Burns and Dietz 1992 and Runciman 1998 for a similar approach). The reciprocal execution of strategies, by producing reciprocal action, would result in the establishment of practices and the foundations of institutions including, importantly, institutional power.

Further, by grounding memetics in the evolution of culturally transmitted behavioural strategies our enterprise could be a) conceptually integrated with behavioural ecology (e.g. Krebs and Davies 1981), b) draw usefully from evolutionary game theory (e.g. Maynard Smith 1993), and c) carve itself a legitimate and coherent investigative niche within the field of evolutionary theory.

Finally, by understanding the social self as web of memes, qua culturally transmitted behavioural strategies, meme theory may provide a useful non-reductionist alternative to a naïve folk psychological conceptualisation of social agency. The self could be deconstructed into networks of contextually dependent strategies that evolve according to the familiar evolutionary loop of variation replication and selection (Dennett 1991, 1998). Herein lies the exciting potential of memetics, a potential that is alas being retarded by the partial and ultimately vacuous metaphor of thought contagion.

## References

- Barkow, J., Cosmides, L., and Tooby, J. (Eds.) (1992). *The Adapted Mind: Evolutionary Psychology and the Generation of Culture*. NY: Oxford University Press.
- Burns, T.R. and Dietz, T. (1992). Cultural evolution: Social rule systems, selection, and human agency. *International Sociology*, 7:259-283.
- Dietz, T, Burns, T.R., and Buttel, H. (1990) Evolutionary theory in sociology: An examination of current thinking. *Sociological Forum*, 5:155-190.
- Clewley R. (1998). Emergence without magic: the role of memetics in multi-scale models of evolution and behaviour. Paper given at the Memetics Symposium, 15<sup>th</sup> International Congress of Cybernetics, Namur, Belgium. <http://zeus.bris.ac.uk/~enrc/conf.html>
- Cloak F.T. (1975). Is a cultural ethology possible? *Human Ecology*, 3:161-182.

Cohen, L.E. and Machalek, R. (1988). A general theory of expropriative crime: An evolutionary ecological approach. *American Journal of Sociology*, **94**:465-501.

Dawkins, R. (forthcoming). Foreword to Blackmore, S. (forthcoming) *The Meme Machine*. Oxford: Oxford University Press.

Dennett, D.C. (1991). *Consciousness Explained*. UK: Penguin.

Dennett, D.C. (1995). *Darwin's Dangerous Idea: Evolution and the Meanings of Life*. New York: Simon and Shuster.

Dennett, D.C. (1998). Memes: Myths, misunderstandings and misgivings.  
<http://ase.tufts.edu/cogstud/papers/MEMEMYTH.FIN.htm>

Durham, W.H. (1990) Advances in evolutionary culture theory. *Annual Review of Anthropology*, **19**:187-210.

Flinn, M.V. (1997). Culture and the evolution of social learning. *Evolution and Human Behavior*, **18**:23-67.

Gatherer, D. (1998). Why the Thought Contagion Metaphor is Retarding the Progress of Memetics. *Journal of Memetics - Evolutionary Models of Information Transmission*, **2**.  
[http://jom-emit.cfpm.org/1998/vol2/gatherer\\_d.html](http://jom-emit.cfpm.org/1998/vol2/gatherer_d.html)

Krebs, J.R and Davies, N.B. (1981). *An Introduction to Behavioural Ecology*. Oxford, UK. Blackwell Science

Levy, D.A., and Nail, P.R. (1993). Contagion: A theoretical and empirical review and reconceptualization. *Genetic, Social and General Psychology Monographs*, **119**:235-183.

Lynch, A. (1996). *Thought Contagion: How Belief Spreads through Society: The New Science of Memes*. New York: Basic Books.

Lynch, A. (1998) Units, events and dynamics in memetic evolution. *Journal of Memetics - Evolutionary Models of Information Transmission*, **2**. [http://jom-emit.cfpm.org/1998/vol2/lynch\\_a.html](http://jom-emit.cfpm.org/1998/vol2/lynch_a.html)

Marsden, P.S. (1998a). Operationalising memetics - suicide, the Werther effect, and the work of David P. Phillips. *Proceedings of the 15th International Congress on Cybernetics*.  
<http://pespmc1.vub.ac.be/Conf/MemePap/Marsden.html>

Marsden, P.S. (1998b). Memetics and social contagion: two sides of the same coin. *Journal of Memetics - Evolutionary Models of Information Transmission*, **2**. [http://jom-emit.cfpm.org/1998/vol2/marsden\\_p.html](http://jom-emit.cfpm.org/1998/vol2/marsden_p.html)

Marsden, P.S. (1998c). The Selectionist Paradigm: More Implications for Sociology. *Sociology Research Online*, 3(4). <http://www.socresonline.org.uk>

Maynard Smith, J. (1993). *Did Darwin get it Right? Essays on Games, Sex and Evolution.* London: Penguin.

Palmer, D.C. and Donahoe, J.W. (1992). Essentialism and Selectionism in Cognitive Science and Behavior Analysis. *American Psychologist*, 47:1344-1359.

Phillips, D.P. (1980). Airplane accidents, murder, and the mass media: towards a theory of imitation and suggestion. *Social Forces*, 58:1000-1024.

Pinker, S. (1998). *How the Mind Works.* London: Allen Lane.

Plotkin, H.C. (1997). *Evolution in Mind.* London: Allen Lane.

Rose, N. (1998). Controversies in meme theory. *Journal of Memetics - Evolutionary Models of Information Transmission*, 2. [http://jom-emit.cfpm.org/1998/vol2/rose\\_n.html](http://jom-emit.cfpm.org/1998/vol2/rose_n.html)

Runciman, W.G. (1998) The selectionist paradigm and its implications for sociology. *Sociology*, 32:163-189.

Sanderson, S.K. (1990). *Social Evolutionism: A Critical History.* Cambridge: Blackwell.

## **THOUGHT CONTAGION: HOW BELIEF SPREADS THROUGH SOCIETY**

**Aaron Lynch, New York, NY: Basic Books (1996)**  
**Cloth: ISBN 0-465-08466-4; Paper: ISBN 0-465-08467-2**

**Reviewed by**  
**Paul Marsden**

Graduate Research Centre in the Social Sciences, University of Sussex, Falmer, Brighton BN1 9RH.

Social contagion theory is not new. In fact, it has been an established explanation for the clustering and spread of human behaviour for nearly two centuries. In his classic study of suicide at the turn of the century, Emile Durkheim (1897) critically reviewed a century of 19th century contagion theory, from Lucas' *Imitation Contagieuse* to Despine's *Contagion Morale* to the then influential and contemporary theory of Tarde (1903). In another landmark study, Le Bon (1895) used contagion as the central mechanism to explain crowd behaviour, and Baldwin (1894) used the concept, *qua* imitation, as "... the law and the only law of the progressive interaction of the organism and its environment." (Baldwin 1894, p. 54) Both Bagehot (1869) and James (1880) provided an evolutionary dimension this ecological framework by arguing that differential propagation could be accounted for in terms of selection processes. Similar evolutionary-ecological approaches to culture were subsequently (re)-proposed and developed by various authors including Zipf (1935), Lovejoy (1936), Medawar (1959), Siegfried (1960), Sperry (1965) and Monod (1971). More recently, Dawkins (1976) coined the term meme - in preference to Zipf's *actemes* or Lovejoy's *unit-ideas* - to denote the unit of focus for such an approach.

The differential replication and selection of non-genetic information (culture) over *time* has been the specific focus of evolutionary epistemology (Popper 1974, Campbell 1974, Hull 1988, Plotkin 1982), whilst the differential replication of culture through both *space* and time has been the focus for contemporary innovation diffusion (Rogers 1995) and social contagion models (Levy and Nail 1993, Marsden 1998) of culture. The evolutionary-ecological approach today competes with rational choice theory as the most useful approach for understanding and predicting the spread of human behaviour.

Recently, a number of models exploring different aspects of the evolutionary dynamics and biases of culture have been developed under the broad umbrella of evolutionary culture theory (ECT). These include the coevolution model (Durham 1991), the dual inheritance model (Boyd and Richerson 1985), the cultural transmission model (Cavalli-Sforza and Feldman 1981), the gene-culture transmission model (Lumsden and Wilson 1981), the i-culture and m-culture model (Cloak 1975) and an emerging memetic model (Dawkins 1993, Dennett 1990, 1991, 1995, Plotkin 1994, Marsden 1998, Blackmore 1999).

Although ECT has tended to focus on the ideational aspect of culture, including the epidemiology of representations (Rindos 1985, 1986, Sperber 1994, 1996), a number of alternative conceptualisations of culture have been proposed, including those emphasising

behaviour (Langton 1979, Gatherer 1998), norms (Allison 1992), practices (Runciman 1998), rules (Burns and Dietz 1992) and strategies (Marsden 1999).

Enter Aaron Lynch's *Thought Contagion: How Belief Spreads Through Society: The New Science of Memes*, a book which I hoped would provide an accessible introduction, overview and perhaps an assessment of a part of this exciting emergent field. Lynch's express aim in the book was one of "putting memetics on more serious footing in the sciences". (P. vii)

*Thought Contagion* is accessible and entertaining, and it has two particular strengths. Firstly, it demonstrates clearly the memetic shift in focus from the internal cognitive processing of social agents, to the identification of patterns of cultural information transmission. Secondly, Lynch shows how such a framework or stance can be used to generate new sets of hypotheses which do not intuitively follow from traditional perspectives. In this way, Lynch does a good job of demonstrating how memetics could be used to generate a novel research programme that is both informed and guided theoretically.

However, Lynch's aim is severely handicapped by his decision to ignore over a century of ECT in particular and social science in general. Instead, Lynch declares that he independently reinvented memetics in 1978 (p. vii) and embarks on a confused and largely uninformed account of how, according to him, ideas about masturbation, breast fetishism, circumcision, psychoanalysis, abortion, firearms, dieting and crucifixion are supposed to spread through society.

In this exposition, Lynch draws on little besides Dawkins' "meme" label. He redefines memes somewhat mysteriously, and without elaboration, as "actively contagious ideas". (p. 2) A modicum of clarification is gained 148 pages later when he states: "Though not exactly a contagion of thought, the addiction to a specific drug can become a replicating brain condition, and hence a meme." (p. 150) Confused? Well, things get even more confusing when Lynch talks of 'ownership' as "a kind of "transcendental meme"" (p. 19) and focuses on what he calls thought contagions, a sort of sub genus of memes, which (we are told) are made up of memes that do not spread through "passive adaptation". (p. 25) For the reader who requires further clarification, Lynch points to an "extremely technical" journal article "filled with propagation diagrams and mathematical equations" on his home page at <<http://www.mcs.net/~aaron>>. I have read this paper and, perhaps because of my own shortcomings, I am no more enlightened. A review of this paper can be found in Gatherer (1998).

Lynch goes on to suggest seven modes of transmission (see Appendix) which are largely at odds with (and uninformed by) anything in ECT, social psychology or social science in general. Of these seven modes, Lynch focuses overwhelmingly on what he calls "quantity parental" and "efficiency parental" modes for his analysis. The reasoning goes something like this: Belief B causes the holder of B to have more children than average in a particular population. Given effective parental socialisation, this means that the children of the holder of B are more likely to hold B too, and thus the prevalence of B will increase with each generation. For example, Lynch suggests that a masturbation taboo may result in an increase in sexual intercourse, presumably to reduce frustration, producing an increase in fertility, which (given effective belief transmission down genetic lines of descent) results in more offspring inheriting the masturbation taboo. (If

you find this sort of reasoning too preposterous to swallow, see p. 4 and p. 90 to check it, and note that this is one of Lynch's more sensible, if unsubstantiated claims.)

The problems with this sort of reasoning are manifold and are indicative of the standard of argument throughout the book. Firstly, Lynch provides not one jot of evidence to substantiate the proposed correlation between masturbation taboo holders and fertility rates. In fact, the entire book is devoid of evidence which substantiates any of his wild speculations. Even if there were a correlation between the taboo holders and fertility, an elementary knowledge of social psychology should have suggested to Lynch that publicly expressed beliefs and private behaviour are often completely at odds with each other (LaPiere 1934). Beliefs and attitudes are generally rather poor predictors of behaviour and have been largely abandoned by predictive models in favour of intentions, within an overall logic of utility maximisation (Petty and Cacioppo 1981). So, for Lynch's central argument to hold, he would first have to show that holding a masturbation taboo was predictive of masturbation behaviour and then show that masturbation and fertility were inversely correlated. He does neither. Furthermore, Lynch's argument would only hold if there was evidence of this correlation **and** if effective parent to child belief transmission could be demonstrated. In the case of sexual mores and norms (which change radically between generations), it is naïve in the extreme simply to assume that this is the case. In fact, Lynch ignores the hugely important effect of belief transmission **between** genetic lines of descent, through schooling, peer groups, the mass media and other institutions. The net result of this approach is a tower of unsubstantiated supposition often producing quite incredible conclusions.

**Why is the "man on top" sexual position so prevalent?** "Memes placing the man on top during intercourse can also boost fertility by giving men more control of the rhythm and pace" (p. 93). This belief can thereby be transmitted to more children, increasing its relative frequency.

**Why did the "domestic woman meme" spread?** Lynch states (without a reference) that recent research has found a positive correlation between women's fertility and their fat-to-lean ratio. "Vigorous physical work" reduces the ratio and hence fertility, therefore this explains, according to Lynch, why domestic wives and the "domestic woman meme" spread through the population. (p. 55, p. 60)

**Why does the "baby doll for girls meme" spread among girls?** "Early practice with a doll might foster the gentle nurturing that helps females perform well as real mothers, too. This would also help the meme replicate into more surviving offspring, accounting for still more of its prevalence." (p. 57)

**Why is circumcision so prevalent?** "Instead, circumcision may have spread by raising reproductive rates. Circumcision eliminates the self-gliding capacity of the penis, channelling more lust (sic) to vaginal intercourse by limiting nonlubricated arousal methods. Activities ranging from masturbation to nearly vaginal sex (sic) become more difficult without an applied lubricant. In ancient times, the need for extra lubricants could have posed some challenge, leading couples to just go ahead and do it (sic) vaginally. The circumcision meme would therefore have flourished parentally." (p. 136)

**Why do "breast privacy memes" spread?** Breast feeding in private "disrupts the hormonal mechanism behind breast-feeding's contraceptive effects, raising reproductive rates for cultures that hide their breast feeding." (p. 137)

To be fair, Lynch does recognise other cases in which beliefs are transmitted *between* genetic lineages, and he introduces what perhaps is the most interesting concept in the book, that notion that thoughts have a specific "propagative profile" made up of the modes by which they are transmitted. Unfortunately, Lynch does not develop or exploit this theme, choosing instead to continue with a series of examples, including "girlish helplessness" and "family-man finder" memes that challenge the credibility of the book. For example:

**What memes deter homosexual behaviour?** "For instance, the belief that male homosexuality requires anal sex deters homosexual mating in those males who dread anal sex." (p. 84)

**Why do socialistic ideas spread?** "... people with little property may expect personal gain by spreading socialist or communist beliefs." (p. 20)

**Why is the breast fetish meme so prevalent?** "Homosexuality and its taboo affect the "fashionability" of certain sexual fetishes in heterosexual males. Specifically, a "mandatory breast fetish" meme evolves and spreads ..." By verbalising attraction to breasts, men "... imply their exclusive heterosexuality." (p. 85)

**Why might progun memes spread better than antigun memes?** "Fiercely progun memes can also intimidate hosts of gun-control memes into silence ..." because "Most antigun folks would probably assume that fiercely progun people own firearms and might use them if angered. Yet few people worry about being shot by an angry gun-control advocate." (p. 163)

**Why does the family home meme spread?** The family home meme means that "... dependent homemakers, in particular, will worry less that a highly employable husband would depart for a new job and a new life far away." (p. 51)

**Why does the big house meme spread?** "The more money a couple spends on joint real estate, the more confident they can feel in each other's commitment to stay settled together. Extra children resulting from this mutual confidence can partly explain the proliferation of plans to buy as large a house as the income allows." (p. 51)

Now there is nothing wrong in using a memetic or any other framework to generate hypotheses, however wild, but they are of no value until evidence is provided to support them. But even in the absence of evidence, Lynch's speculations seem to be devoid of any rationale. For example:

**Why do men talk more about/prefer women's breasts than posteriors?** "Indeed, young males may feel inhibited from discussing female posteriors because it reminds them of anal sex and homosexuality ..." (p. 86)

**Why is masturbation "prime material for commercial use"?** Masturbation is a "prime material for commercial use" because it can make "... people pay attention long enough to hear a

commercial, and then improve recall of the commercial by "downloading" it to an aroused audience." (p. 91)

**Why do women want to have children?** "As the girl grows up, she learns that she is too old to play with dolls anymore. But the desire for comfort and attachment remains, and translates into a desire for a real baby". This is what lies behind the phenomenon of "baby lust". (p. 57)

The problem I have with all this uninformed speculation is not that there is the slightest risk that anybody could possibly take it seriously, but rather that it might be construed as representative of work that is being done in ECT and memetics, because it is not. Fortunately, suspicion is likely to be aroused in even the most uninitiated reader when Lynch includes a discussion of the science fictional "psychohistory" in his brief misrepresentation of the various social sciences, in the process of claiming that his brand of memetics is "the missing link".

I have argued elsewhere (Marsden 1999), as have others (Gatherer 1998, Flinn 1997) that, when taken literally as Lynch does, the thought contagion metaphor is quite disingenuous. I will not repeat the arguments here, suffice it to say that even if thoughts are anything other than a heuristic description in folk psychology, we know nothing about how they combine or recombine and it is certainly not ignoring cognitive processing and endowing thoughts with contagious properties that makes the memetic approach any more viable. When Lynch conflates the partial metaphor of thought contagion with population genetics by arguing that "Population memetics is the study of how proliferating memes combine and separate in a population", which "roughly parallels" population genetics (p. 12), his whole approach becomes even more confused.

What, for example, is it about the "mandatory breast fetish meme" that qualifies it as a meme? Just in what sense is it "actively contagious"? Or how is it a replicating brain condition? We know that folk psychological concepts do not match specific brain states; there is no breast fetish brain neuron or state that replicates and is identical across different brains. How does this meme recombine and perhaps then separate from a 'non mandatory nipple revulsion meme'? If they actually exist at all, thoughts and their 'meanings' are certainly contextually and subjectively dependent (Heyes *et al.* 1993, *passim*), and this fact alone strongly suggests that there can be no laws of combination and separation.

Does *Thought Contagion* demonstrate that an epidemiological or evolutionary-ecological approach to understanding non-genetically transmitted information (culture) is ultimately redundant, vacuous or false? The answer, I suggest, is no simply because exposure to culturally transmitted information clearly does influence the likelihood of an agent (re)-transmitting the information in that or another form. Call it social, vicarious or observational learning, replication or imitation, the fact is that our behavioural strategies are in part determined by the behavioural strategies of those around us. (See Marsden 1998 for a review.) From this simple observation, it is possible to adopt an alternative stance, a memetic stance, for interpreting, understanding, explaining and predicting human behaviour. Just as it is possible to interpret human behaviour *as if* it emanates from a homuncular centre of intentionality, our propensity to copy behavioural strategies, and the behavioural clustering that results, means that it is sometimes useful to understand non-genetic information transmission *as if* it were a contagion. By this, all I mean is

that our production of behaviour is sometimes dependent on the extensity and intensity of our exposure to that behaviour, or at least some representation of that behaviour.

By understanding the social agent as differentially selecting and replicating the information it is exposed to, models such as the Bass equation (Bass 1969) can be used to predict and model the spread of behaviours and products through a population without invoking or relying on homuncular folk psychology. Such an approach does not ignore cognitive processing, it simply shifts the focus to the description of how one subset of non-genetically transmitted input information is *selected* for output over another.

Of course, social agents do not blindly and directly imitate all the strategies they are exposed to, as in a simple game of Tit For Tat. Replication is altogether more sophisticated, temporally and contextually cued, and critically, involving the selection of one strategy instead of another. We are highly discriminating and discerning imitators. Indeed, the evolved structure of our brains has produced a remarkable intentional system out of our learning capacities such that we can usually view our behaviour as if it were produced by a principle of utility maximisation. Social agents are pretty good control systems, selecting various strategies to attain antecedently fixed ends.

However, the frame problem precludes optimality in this utility maximisation process, and this means that the intentional stance sometimes breaks down - we sometimes seem to behave irrationally. This, I suggest, is where memetics may be useful. Just as evolutionary psychologists may focus on how *biases* in strategy selection might result from the evolved cognitive architecture of our brains, memetics might inform us of how the differential exposure to culture influences strategy selection. Just as the evolved architecture of the brain may produce an intentional system with certain biases towards inclusive genetic fitness enhancement (within the environment of evolutionary adaptedness), a memetic stance might legitimately invoke an analogous heuristic of ***inclusive memetic fitness enhancement*** (within a particular context) to explain non-rational strategy selection. Because differential exposure to stimuli impacts upon how we respond to situations (by increasing the likelihood of replicating that stimuli), the trajectory and evolution of cultural information through a cultural environment will appear as if the information "wants to get itself" replicated.

The principle of inclusive memetic fitness enhancement is simply a heuristic (as opposed to an ontological) claim for interpreting the production of culture as if it were an attempt to enhance the survival chances of that culture, that is, as if the culture itself had a rudimentary intentionality. Crudely speaking, rather than always behave selfishly for what I personally may want (whatever that may mean), I will fight for my ideas and help others who share my ideas. Just as I may lay down my life for more than 2 brothers or more than 8 first cousins, I might also lay down my life for more than 2 brothers in arms, or more than 8 individuals with whom I share more than an eighth of my culture. Note that this memetic stance is neither incompatible with the intentional stance, nor with folk psychology; it simply provides a different stance for analysing human interaction, one that adumbrates certain features of culture and clarifies others. In sum, what I am proposing is that the memetic stance may provide a useful heuristic for interpreting non-rational behaviour, i.e. when the intentional stance breaks down. This memetic stance of course needs to be tested, but if it turns out to be useful in modelling and predicting behaviour,

integrating it with its sociobiological counterpart may one day provide a real alternative to homuncular folk psychology.

Theorists often speak in terms of architecture. They describe their work in terms of constructing theoretical edifices and laying theoretical foundations. This is useful language, it highlights the laborious step by step nature of theory building, where each move forward, whilst a tentative and speculative move into the dark, is both theoretically and empirically informed by what has gone before it, and is designed so as to be open to testing. If Aaron Lynch's intention in writing *Thought Contagion* was to put memetics on more serious footing in the sciences, in this he must be seen to have failed quite spectacularly. Fortunately, memetics has somewhat more secure foundations, as exemplified in the recent publication of Sue Blackmore's *The Meme Machine* (Blackmore 1999), which I commend to you as a useful contribution to ECT.

## Appendix: Lynch's Seven Modes of Memetic Transmission

**"Quantity Parental"** - Have lots of children and use your position of authority to indoctrinate them and increase the prevalence of your beliefs.

**"Efficiency parental"** - Propagate the belief that parents are good sources of indoctrination, then have lots of children and use your position of authority to further indoctrinate them and increase the prevalence of your beliefs.

**"Proselytic"** - Convince your non-kin that your belief is both right and urgent so that they will spread it for you. Lynch uses the example of appeals to fear. A belief that spreads will spread ...

**"Preservational"** - Ideas that protect your ideas from being challenged, such as the belief that one should never talk about religion or politics in polite company.

**"Adversative"** - A belief involving "selectively killing those who refuse to convert" to your beliefs, or at least sabotaging alternative beliefs. "The Muslim belief" is cited as an example of the adversative mode.

**"Cognitive Advantage"** - Make sure your belief appears "well founded" and it will tend to spread.

**"Motivational"** - Make your idea a good idea and it will tend to spread.

## References

ALLISON P. 1992. The Cultural Evolution of Beneficent Norms, *Social Forces*, 71:279-301.

BAGEHOT W. 1887 [1869]. *Physics and Politics or Thoughts on the Application of the Principles of 'Natural Selection' and 'Inheritance' to Political Society*, eighth edition, Kegan Paul, Trench and Company, London.

- BALDWIN J. M. 1894. Imitation: A Chapter in the Natural History of Consciousness, *Mind*, 3:25-55.
- BASS F. M. 1969. A New Product Growth Model for Consumer Durables, *Management Science*, 15:215-227.
- BLACKMORE S. 1999. *The Meme Machine*, Oxford University Press, Oxford. [JASSS review]
- BOYD S. and P. J. Richerson 1985. *Culture and the Evolutionary Process*, University of Chicago Press, Chicago, IL.
- BURNS T. R. and T. Dietz 1992. Cultural Evolution: Social Rule Systems, Selection, and Human Agency, *International Sociology*, 7:259-283.
- CAMPBELL D. T. 1974. Evolutionary Epistemology. In P. A. Schilpp, editor, *The Philosophy of Karl Popper*, The Library of Living Philosophers Volume 14, Book 1, Open Court, La Salle, IL, 413-463. [Reprinted in Plotkin 1982.]
- CAVALLI-SFORZA L. L. and M. W. Feldman 1981. *Sociocultural Transmission and Evolution: A Quantitative Approach*, Princeton University Press, Princeton, NJ.
- CLOAK F. T. 1975. Is a Cultural Ethology Possible? *Human Ecology*, 3:161-182.
- DAWKINS R. 1976. *The Selfish Gene*, Oxford University Press, Oxford.
- DAWKINS R. 1993. Viruses of the Mind. In B. Dahlbohm, editor, *Dennett and His Critics*, Basil Blackwell, Oxford, 13-27.
- DENNETT D. C. 1990. Memes and the Exploitation of Imagination, *Journal of Aesthetics and Art Criticism*, 48:127-35.
- DENNETT D. C. 1991. *Consciousness Explained*, Penguin, Harmondsworth.
- DENNETT D. C. 1995. *Darwin's Dangerous Idea: Evolution and the Meanings of Life*, Allen Lane, London.
- DURHAM W. H. 1991. *Coevolution: Genes, Culture and Human Diversity*, Stanford University Press, Stanford, CA.
- DURKHEIM E. 1952 [1897]. *Suicide: A Study in Sociology*, Routledge, London.
- FLINN M. V. 1997. Culture and the Evolution of Social Learning, *Evolution and Human Behavior*, 18:23-67.

- GATHERER D. 1998. Why the Thought Contagion Metaphor is Retarding the Progress of Memetics, *Journal of Memetics - Evolutionary Models of Information Transmission*, 2, <[http://www.cpm.mmu.ac.uk/jom-emit/1998/vol2/gatherer\\_d.html](http://www.cpm.mmu.ac.uk/jom-emit/1998/vol2/gatherer_d.html)>.
- HEYES S. C., L. J. Heyes, H. W. Reese and T. R. Sarbin, editors, 1993. *Varieties of Scientific Contextualism*, Context Press, Reno, NV.
- HULL D. L. 1988. *Science as a Process: An Evolutionary Account of the Social and Conceptual Development of Science*, University of Chicago Press, Chicago, IL.
- JAMES W. 1880. Great Men, Great Thoughts, and the Environment, *The Atlantic Monthly*, October, 46(267):441-459.
- LANGTON J. 1979. Darwinism and a Behavioural Theory of Sociocultural Evolution: An Analysis, *American Journal of Sociology*, 85:288-309.
- LA PIERE R. T. 1934. Attitudes Versus Actions, *Social Forces*, 13:230-237.
- LE BON G. 1903 [1895]. *The Crowd: A Study of the Popular Mind*, T. Fisher Unwin, London.
- LEVY D. A. and P. R. Nail 1993. Contagion: A Theoretical and Empirical Review and Reconceptualization, *Genetic, Social and General Psychology Monographs*, 119:235-183.
- LOVEJOY A. O. 1936. *The Great Chain of Being: A Study of the History of an Idea*, Harvard University Press, Harvard, MA.
- LUMSDEN C. J. and E. O. Wilson 1981. *Genes, Mind and Culture*, Harvard University Press, Harvard, MA.
- LYNCH A. 1996. *Thought Contagion: How Belief Spreads through Society: The New Science of Memes*, Basic Books, New York, NY.
- MARSDEN P. S. 1998. Memetics and Social Contagion: Two Sides of the Same Coin? *Journal of Memetics - Evolutionary Models of Information Transmission*, 2, <[http://www.cpm.mmu.ac.uk/jom-emit/1998/vol2/marsden\\_p.html](http://www.cpm.mmu.ac.uk/jom-emit/1998/vol2/marsden_p.html)>.
- MARSDEN P. S. 1999. A Strategy for Memetics: Memes as Strategies, *Journal of Memetics - Evolutionary Models of Information Transmission*, 3, <[http://www.cpm.mmu.ac.uk/jom-emit/1999/vol3/marsden\\_p.html](http://www.cpm.mmu.ac.uk/jom-emit/1999/vol3/marsden_p.html)>.
- MEDAWAR P. 1959. The Future of Man, The BBC Reith Lectures. [Transcript ]
- MONOD J. 1971. *Chance and Necessity*, Alfred E. Knopf, New York, NY.
- PETTY R. E. and J. T. Cacioppo 1981. *Attitudes and Persuasion*, William C. Brown Company, IA.

PLOTKIN H. C., editor, 1982. *Learning, Development, and Culture: Essays in Evolutionary Epistemology*, John Wiley and Sons, Chichester.

PLOTKIN H. C. 1994. *Darwin Machines and the Nature of Knowledge*, Penguin, Harmondsworth.

POPPER K. R. 1979 [1974]. *Objective Knowledge: An Evolutionary Approach*, revised edition, Clarendon Press, Oxford.

RINDOS D. 1985. Darwinian Selection, Symbolic Variation and the Evolution of Culture, *Current Anthropology*, 26:65-88.

RINDOS D. 1986. The Evolution of the Capacity for Culture: Sociobiology, Structuralism, and Sociocultural Selectionism, *Current Anthropology*, 27:315-332.

ROGERS E. M. 1995. *Diffusion of Innovations*, fourth edition, The Free Press, New York, NY.

RUNCIMAN W. G. 1998. The Selectionist Paradigm and Its Implications for Sociology, *Sociology*, 32:163-189.

SIEGFRIED A. 1960. Germs and Ideas: Routes of Epidemics and Ideologies. [Partial transcript]

SPERBER D. 1994. The Modularity of Thought and the Epidemiology of Representations. In L. Hirschfeld and S. Gelman, editors, *Mapping the Mind, Domain Specificity in Cognition and Culture*, Cambridge University Press, Cambridge, 39-68.

SPERBER D. 1996. *Explaining Culture: A Naturalistic Approach*, Basil Blackwell, Oxford.

SPERRY R. 1965. Mind, Brain and Humanist Values. In R. Platt, editor, *New Views on the Nature of Man*, University of Chicago Press, Chicago, IL, 71-92.

TARDE G. 1963 [1903]. *The Laws of Imitation*, Peter Smith, Clouchester, MA.

ZIPF G. K. 1965 [1935]. *Psychobiology of Languages*, The M.I.T. Press, Cambridge, MA.

# FOREFATHERS OF MEMETICS: GABRIEL TARDE AND THE LAWS OF IMITATION

By

**Paul Marsden**

Graduate Research Centre in the Social Sciences

University of Sussex

[PaulMarsden@msn.com](mailto:PaulMarsden@msn.com)

*"Meanwhile, let us not forget that every invention and every discovery consists of the interference in somebody's mind of certain old pieces of information that have generally been handed down by others. What did Darwin's thesis about natural selection amount to? To have proclaimed the fact of competition among living things? No, but in having for the first time combined this idea with the ideas of variability and heredity. The former idea, as it was proclaimed by Aristotle, remained sterile until it was associated with the two latter ideas. From that as a starting point, we may say that the generic term, of which inventions is but a species, is the fruitful interference of repetitions."*

Gabriel Tarde, *The Laws of Imitation* p.382

One way to conceptualise memetics is as a stance that seeks, using a selectionist rationale, to interpret the human social world in terms of the ongoing differential reproduction of traits describing that social world (Marsden 1999). Reproduction in the social world may occur at many levels; objectified individual acts are replicable and can become typified and habitual, thereby becoming practices, themselves reproducible, and through typified reciprocal interaction, institutions can emerge with roles that serve their own propagation (cf. Price and Shaw 1998). Through this multi-level sociological dance of reproduction, we both produce and are produced by the social world. The idea is that by employing a selectionist rationale, essentially the Law of Effect and perhaps a heuristic of inclusive memetic fitness, memetics may come to provide some purchase on how acts, practices and institutions provide for their own self-emplacement in a world where the products of variation outnumber the capacity for subsequent selection and reproduction.

Such a memetic research focus is not, however, without precedent; a century ago, one of the founding fathers of sociology, Gabriel Tarde, outlined a programme for sociology in his treatise, *Les Lois de l'Imitation* that has much in common with this memetic project (*The Laws of Imitation* 1890, translated 1903, reprinted in English 1962 to which page references here refer). The purpose of this short research note is to familiarise those participating in the construction of a memetic perspective with Tarde's model, a model that certainly has historical significance for contemporary memetics and possibly theoretical significance as well.

Gabriel Tarde was born in Sarlat, the Dordogne, France in 1843 where he grew up to become a lawyer and juge d'instruction. Early on in his career, he observed that particular crimes appeared to spread in waves through society as if they were fashions. Becoming increasingly interested with how this epidemiological aspect of criminal activity might be just one instance of a more general feature of the social world, Tarde published a number of papers in the *Revue*

Philosophique between 1880 and 1901, including a paper entitled, "Darwinisme naturel et Darwinisme social" (Tarde 1884) which developed this idea and outlined a general research programme for sociology.

Basically, what Tarde proposed was a different way of looking at the social world, not from the perspective of the individual or the group, but from the point of view of the products, acts and ideas that were used to classify those individuals or groups. By focusing on how these features were differentially reproduced, Tarde suggested it was possible to infer certain regularities or laws that appeared to pattern the social world. In 1890, he presented his arguments along with a preliminary set of "laws" in his major work, *Les Lois de l'imitation*, and in doing so provided researchers with model for interpreting the social world. The key point of distinction of Tarde's model, over other emerging sociological models such as the structural sociology of his compatriot Emile Durkheim, was that social organisation and the social relations describing that organisation were understood as a means for the more ultimate goal of propagation.

*"Self-propagation and not self-organisation is the prime demand of the social as well as of the vital thing. Organisation is but the means of which propagation, of which generative or imitative imitation, is the end." Tarde 1903/1962: 74*

By 'generative imitation', Tarde referred to a process of invention that he conceptualised in terms of the combination of existing imitations, that is, the "fruitful interference of repetitions", whilst he used the term 'imitative imitation' to denote the propagation of inventions and their imitations across time and space. It should be clear that such an inclusive understanding of imitation as denoting general products and processes of repetition in the social world is very different from the more restrictive use of the term adopted by some meme theorists (e.g. Blackmore 1998) who, following comparative psychologists, employ the term to denote only acts socially learned by observation that are then performed. Rather, Tarde wished to stretch the already more inclusive meaning of the French word imitation and imbibe it with a deeper meaning to refer to the general class of objects that were reproduced in society and the various processes that were instrumental in their reproduction, "whether willed or not willed, passive or active" (xiv).

Tarde's basic argument was that human history could be usefully interpreted as a "career" of imitations, trajectories of inherited inventions through populations that differentially survive an ongoing culling process of "counter-imitation", that is, rejection (154).

*"All resemblances of social origin in society are the direct or indirect fruit of the various forms of imitation, - custom-imitation or fashion-imitation, sympathy-imitation or obedience-imitation, precept-imitation or education-imitation, native imitation, deliberate imitation, etc. In this lies the excellence of the contemporaneous method of explaining doctrines and institutions through their history. It is a method that is certain to come into more general use." (14)*

So central and universal to the human condition was imitation thus conceived for Tarde that he went so far as to suggest that society itself could be defined as imitation insofar as societies could be described in terms of populations of individuals with common imitated traits who are apt to share imitations (68)

*"What is society? I have answered: Society is imitation". (74).*

The primary task of Tarde's proposed sociology was to identify the variables that appeared to influence whether an object of imitation would become successful, that is, get reproduced, in order to explain how some imitations came to be selected over others (110-111).

*"Our problem is to learn why, given one hundred different innovations conceived of at the same time - innovations in the form of words, in mythical ideas, in industrial processes etc.- ten will spread abroad, while ninety will be forgotten." (140)*

Tarde's answer to this question was to propose a number of general laws that seemed to pattern the differential reproduction of imitations, which he divided into those that had an apparent logical quality, and those of an extra-logical or less rational nature. Specifically, the "logical laws of imitation" (140-188) proposed by Tarde were that:

The origination of an invention involves the recombination of existing imitations, and this origination will be influenced by the social context and abilities of those involved with the recombination.

The success of an imitation in spreading geometrically from its point of origination will be a function of its fit, that is, compatibility, with the environment of existing imitations.

The selection, that is, adoption of an imitation occurs either through "substitution" involving a "logical dual" and "struggle" between two alternatives, or through "accumulation", a process entailing a logical union of imitations.

Tarde also noted that it was important to understand the success of competing imitations as a function of the power held by those with vested interests in the reproduction of those imitations (169). Although, he did not couch it in such terms, I have suggested that this insight may provide a useful rationale for interpreting the differential propagation of imitations in terms of relations to the means of reproduction (Marsden 1998).

Tarde's principal "extra-logical influences" (189-365) stated the following:

The reproduction of ends generally precedes the reproduction of the means to those ends (194-213). In other words, in Tarde's view, goals tend to be imitated before the actions that serve to attain them are adopted. For example, the imitation of a goal, say to become rich, will generally antecede the adoption of imitations employed to further this goal.

Imitations tend to propagate through a process of stratified diffusion from those perceived as superior to those who perceive them as superior (213-243). For example, traits originally associated with celebrities and otherwise privileged tend to trickle down to those who associate them with such traits.

Tarde also made several observations pertaining to what he suggested were general tendencies in the spread and selection of imitations. For instance, he suggested that in democratic populations

the voice of public opinion increasingly becomes the authority whose example is copied, whilst traditional and expert authority wanes (229). Similarly, he proposed that imitative activity, including generative imitation, tends to be proportional to population density (239), thereby allowing for the faster spread and development of imitations in cities than rural areas.

Additionally, Tarde noted a general shift in the mode of social reproduction from custom, that is, endemic horizontal transmission, to fashion - epidemic horizontal transmission (244-255).

In an observation that would be at home with Brodie's memetic reconceptualisation of social influence literature (Brodie 1996), Tarde also raised the possibility of engineering a successful imitation independently of any truth or utility that imitation may have. To engineer a successful imitation, or meme in today's parlance, Tarde suggested that it might suffice to present the invention (mutant imitation) as a descendent of an endemic part of culture into which it is to be introduced:

*"...assume the mask of the enemy and besiege existing custom by unearthing some ancient custom long since fallen into discredit and rejuvenated for the needs of her cause..."* (261)

Although Tarde's proposed research programme was somewhat eclipsed in sociology by Durkheim's more structural approach, his work comprising over 5000 published pages (Clark 1969) became the inspiration for innovation adoption research in applied sociology, particularly rural sociology (Ryan and Gross 1943) and medical sociology (Katz 1961). Likewise, Tarde's model underpins much that passes for diffusion research in marketing (Bass 1969, Rogers 1995), whilst in psychology his framework influenced the thinking of Baldwin (1894) and more indirectly Mead (1934).

For researchers in memetics, any potential contribution of Tarde has yet to be realised, but it is hoped that this brief introduction to one of the forefathers of memetics might serve as a creative stimulus for the 'generative imitation' that will come to define our enterprise.

## References

- Baldwin, J. M. (1894). Imitation: A chapter in the natural history of consciousness. *Mind*, **3**:25-55.
- Bass, F.M. (1969) "A new product growth model for consumer durables" *Management Science*, **13**(5):215-227.
- Blackmore, S. (1998) Imitation and the definition of a meme. *Journal of Memetics – Evolutionary Models of Information Transmission*, 2. [http://jom-emit.cfpm.org/1998/vol2/blackmore\\_s.html](http://jom-emit.cfpm.org/1998/vol2/blackmore_s.html)
- Brodie, R. (1996). *Viruses of the Mind: The New Science of the Meme*. Seattle, WA: Integral Press.
- Clark, T. (ed.) (1969) *Gabriel Tarde on Communication and Social Influence*, Chicago, University of Chicago Press.

Katz, E. (1961) The social itinerary of social change: Two studies on the diffusion of innovations" in W. Schramm (ed.) *Studies of Innovation and of Communication to the Public* Stanford, CA.

Marsden (1998) "Evolutionary sociology reinterpreted". Unpublished paper presented to the Graduate Research Centre in the Social Sciences, University of Sussex, June 1998.

Marsden (1999) "Review of Thought Contagion: How Beliefs Spread through Society by A.

Lynch" *Journal of Artificial Societies and Social Simulation*

<http://jasss.soc.surrey.ac.uk/2/2/review4.html>.

Mead, G.H. (1934) *Mind, Self, and Society*. Chicago: University of Chicago Press.

Price, I. & Shaw, R. (1998) *Shifting the Patterns: Breaching the Memetic Codes of Corporate Performance*, Charlford, Glos. Management Books 2000.

Rogers, E. M. (1995) *The Diffusion of Innovations* (4<sup>th</sup> ed). New York. The Free Press.

Ryan, B. & Gross, N.C. (1943) "The diffusion of hybrid seed corn in two Iowa communities" *Rural Sociology* 8:15-24.

Tarde, G. (1890) *Les Lois de l'Imitation*, Paris 3ème éd. revue et augmentée 1900.

Tarde, G. (1884) "Darwinisme naturel et Darwinisme social" *Revue Philosophique* XVII; 607.

Tarde, G. (1903) *The Laws of Imitation*, translated by E.C. Parsons with introduction by F.Giddings, New York, Henry, Holt and Co.

Tarde, G. (1962) *The Laws of Imitation*, translated by E.C. Parsons with introduction by F.Giddings, reprint, Gloucester, MA, Peter Smith.

# IS SUICIDE CONTAGIOUS? A CASE STUDY IN APPLIED MEMETICS

By

**Paul Marsden**

Visiting Research Fellow

Graduate Research Centre in the Social Sciences

University of Sussex

[paul@viralculture.com](mailto:paul@viralculture.com)

## Abstract

The phenomenon of suicide contagion is demonstrated experimentally. An interpretation of the results is proposed using an understanding of *memetics as contagion psychology informed by selectionist thinking*. Using the term 'meme' to denote an object of contagion and 'contagion' to denote a process of spread by exposure, a selectionist explanation of why certain people might be susceptible to a contagion of suicide is provided. Specifically, it is suggested that people who have become socially isolated and culturally disenfranchised, i.e. those with reduced residual cultural fitness (compromised access to the means of cultural reproduction), might be at particular risk from suicide contagion. Finally, public health policy implications of this memetic understanding of suicide are briefly outlined.

## 1 Introduction

In June 1962, the managers of a textile factory in Strongsville, USA, were obliged to close down their factory because of a 'mysterious sickness' that was affecting onsite workers.

On the evening of the closure, a news report described how at least ten women and one man had been admitted to hospital suffering from rashes and severe nausea. The news report also suggested that the cause of the sickness was a poisonous bite from an insect that had arrived in a shipment from England, and had taken up residence in the factory.

Several weeks later, with the plant still closed, a total of 62 workers had sought medical attention, having developed the symptoms of having been bitten by the bug. From initial rashes and nausea, the effects of the poisonous bite appeared to develop into chronic severe weakness punctuated with acute panic attacks.

The US Public Health Service 'Communicable Disease Center' with a team of entomologists was called in to investigate this 'June Bug', and took the expensive decision to fumigate and vacuum the huge textile plant.

After careful investigation, it was concluded that there had been no dangerous bug at all, English or otherwise. Rather, the episode had been an example of a process that social psychologists call 'social contagion', a generic label used to describe the apparent infectious spread of opinions, emotions and behaviour by exposure to similar opinions, emotions or behaviours.

Social contagion usually occurs in contexts of uncertainty and stress, where people make use of information in the actions of people around them to make sense of situations, resolve ambiguity and to inform their own responses (Colligan, Pennebaker and Murphy 1982). Sure enough, when two social psychologists, Kerckhoff and Back (1968), interviewed victims of the mysterious 'June Bug' illness they found that those who had been 'bitten' had been suffering from undiagnosed stress and alienation.

## 2 Memetics and Social Contagion: Two Sides of the Same Coin

I have argued in a previous paper that for applied memetics, an inclusive and pragmatic working definition of a meme is *the object of contagion* (Marsden 1998a). This is consistent with the popular understanding of memetics as the study of 'infectious' elements of culture. By 'infectious', what is meant is the quality of some acts, emotions and opinions, in certain contexts, to spread by exposure rather than by some deliberate attempt to influence (such as coercion or persuasion).

The purpose of this paper is not to engage in an abstract discussion of the merits of this or that conception of a meme or memetics; indeed the purpose of using the meme neologism to describe objects of contagion is wholly pragmatic; *to provide a rich, substantive research focus, and get on with doing memetics*. Interestingly, contagion psychology shows that a meme, thus conceived, may have no essential properties, and is always a meme in context, thereby defying useful abstract discussion. For example, an act may spread by exposure in one context (and thus be a meme), but spread by coercion in another (and therefore not be a meme). Of course, if the term 'meme' is used as a scientific-sounding basket label for all received ideas or socially learned behaviours, then the notion of a meme as an object of contagion, and *memetics as a selectionist interpretation of contagion psychology*, may seem unnecessarily restrictive. However, the advantage of the more modest conception of memetics is that it is useful *within* established social scientific models of human psychology, instead of being of interest only from *without*.

What follows is an example of how applied memetics might usefully proceed based on this understanding of a meme as an object of contagion, focusing on the peculiar phenomenon of suicide contagion.

## 3 Suicide Contagion - Infected by a Suicide Meme

Suicide contagion is said to occur when exposure to suicidal acts appears to trigger copycat suicidal acts. This peculiar phenomenon challenges rational interpretation perhaps as much as the notion of suicide, that is, deliberate self-destruction, itself. Now, sense can be made of many memetic phenomena within an established social learning framework; that we use information vicariously available to us in other people's experiences to inform our own adaptive responses to ambiguous situations. However, suicide contagion does not seem to fit this model insofar as suicide is clearly not adaptive from the suicidal individual's point of view. So bizarre is the phenomenon, that despite the consistent finding of elevated suicide levels following suicide publicity (Marsden 1998b), some researchers find the whole idea of suicide contagion incredible. Nevertheless, strict official guidelines do exist on how suicide should be publicly discussed (e.g. CDC 1994) based on the assumption that individuals contemplating suicide who see suicide or

suicidal individuals rewarded (in attention and/or positive evaluation) may infer that suicide is an appropriate response to their distress.

The suicide contagion hypothesis is that exposure to suicide is a suicide risk-factor for people experiencing an unresolved conflict as to whether suicide is an appropriate response to current unresolved distress. Whilst ethical considerations obviously preclude the direct experimental investigation of this hypothesis, one aspect of suicide contagion does lend itself to indirect experimental research. This is the impact of semantic priming.

#### **4 Priming your Mind with Suicide**

Priming refers to the idea that the interpretation of situations can be involuntarily patterned by recent and frequently experienced events (Fiske and Taylor 1991). To take a trivial example, a habitual walk home after seeing a frightening movie may result in an increased level of 'feeling spooked' because the themes evoked by the film remain salient in the mind, thereby producing contagion; the spookiness of the movie influences subsequent interpretations of events. More generally, priming suggests that the ways in which situations are interpreted can be influenced by ideas salient in memory that have been the focus of recent and/or frequent attention. Put simply, we tend to use the meanings we have to hand in interpreting situations. Further, because some ideas cue other ideas, this aspect of what is known as social cognition may result in 'spreading activation' whereby exposure to ideas cue not only those ideas but also related ideas.

*"Thoughts of which one is consciously aware send out radiating activation along associative pathways, thereby activating other related thoughts. In this way, ideas about aggression that are not identical to those observed in the media may be elicited later. In addition, thoughts are linked, along the same sort of associative lines, not only to other thoughts but also to emotional reactions and behavioural tendencies." (Geen and Thomas 1986: 12)*

For example, in one experiment it was found that people became more aggressive following exposure to images of weapons, as aggressive ideas around weapons were primed in the mind of the exposed to them (Leyens and Parke 1975). The idea is that when activated by becoming the focus of attention, a concept and related concepts in the semantic network of an individual's memory are easier to retrieve, that is, the mind is primed with these concepts and will tend to use them to interpret situations (Higgins 1989, Fiske and Taylor 1991, Berkowitz 1984, Jo and Berkowitz 1994).

Applied to suicide contagion, this interpretation would predict that exposure to suicide should have a short-term impact on how a situation is interpreted as being potentially suicidal. To test the plausibility of this model, the following experiment was conducted as part of a D.Phil research project in 1999 (Marsden 2000).

## **5 Testing the Model - Is Suicide Contagious?**

### **5.1 Research Objectives**

To assess the empirical plausibility of the idea of suicide contagion by priming; specifically that mediate exposure over the Internet to the concept of suicide can influence how a situation is subsequently interpreted as being potentially suicidal.

Specifically, the research hypothesis was as follows:

Priming participants with the idea of suicide, by informing them that the topic of research was 'young people and suicide', would result in an increased likelihood of interpreting a distressing and ambiguous situation described by a text as suicidal

### **5.2 Materials and Method**

Standard experimental design of controlled exposure to the concept of suicide over the web was followed by data capture in a self-completion web questionnaire. Piloted in April 1999, the research took place in May 1999. Participants were randomly assigned to either an experimental group or a control group, all invited to first read a short text about a distressed student at university, and then indicate on a standard 5-point Likert scale, the likelihood that they thought that the student would commit suicide (1 = not at all likely, 5 = very likely). The only difference in materials used between the two groups was how the research study was introduced to them; in the experimental group the welcome screen introduced the study by saying that it was about young people and suicide, and in the control group about young people and stress. The texts and HTML questionnaires were published to the Internet at using Microsoft FrontPage and a custom CGI script that randomly directed participants from an initial index welcome page to one of the two texts. The responses were captured in a simple CSV (comma separated variable) database sent from the HTML form.

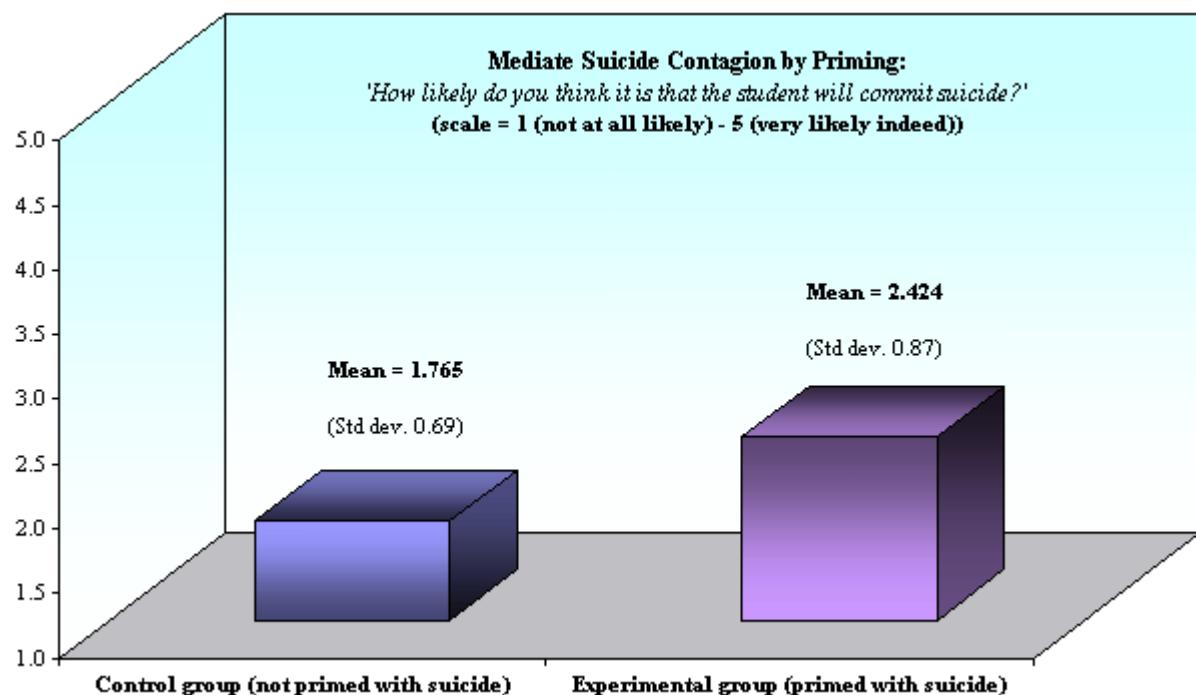
### **5.3 Participants and Recruitment**

Participants were UK Internet-users over the age of 18 recruited through chain email. Because of the potentially sensitive nature of the research, no incentives were offered, and instructions were given to those recommending individuals that only those over eighteen should be invited. Additionally, attention was called to a hyperlink provided on each page of the web questionnaire to an online support group, (Samaritans - although this was unnamed) to be used if participants found the exercise in any way distressing. Finally, participants were told that they could change their mind about participating, and withdraw from the exercise at any time.

### **5.4 Results**

67 UK Internet-users completed the task (36 male, 31 female, average age 22) over a two-week period in May 1999. The research hypotheses predicted that participants primed with the concept of suicide would be more likely to assess an ambiguous situation as more suicidal than participants not primed with the concept of suicide. An inspection of the mean results was

consistent with this, with those primed with suicide interpreting the likelihood that a distressed individual would commit suicide at 2.42, and this compared to 1.77 in the control group. A one-way ANOVA (analysis of variance) revealed that this difference was statistically significant at the 0.001 level ( $F(1,65) = 11.79$ ,  $p < .001$ ) indicating a strong patterning of the data in a way consistent with expectations of the model.



**Figure 1:** Suicide is Contagious

## 5.5 Discussion

The results of this study were consistent with the idea that exposure to the idea of suicide can prime the concept of suicide in peoples' minds and influence the interpretation of ambiguous situations by facilitating a perception of that situation as suicidal. Although the research hypothesis was confirmed by a strong patterning of the data, thus suggestive of such an influence for the group of participating adult UK Internet-users, the use of non-probability sampling would have precluded any possibility of generalising the results to a more general UK Internet-user population, even if this were considered a legitimate strategy. Further, although the results were continuous with expectations following from the model in terms of interpreting somebody else's situation as suicidal, this does not imply, other than extremely weakly, that participants might also be more likely to interpret their own situations as suicidal following exposure to suicide. Nevertheless, the results of this small study are consistent with such a view.

## 6 But is this Memetics?

The above experiment is a simple example of applied memetics operationalised as the study of infectious acts, opinions and emotions. Now, the critic could argue that this understanding of

memetics adds little to the already established field of social contagion research, apart from providing a useful label for the object of contagion. However, from the understanding of memetics proposed here, this is not a problem because memetic research is social contagion research.

Further, it may be possible to gain some leverage from the selectionist heritage that underpins Dawkins' neologism, and use it as a creative stimulus for generating hypotheses on why certain people appear to be susceptible to contagion. By way of example, I would like to conclude by proposing a selectionist understanding of why certain people, and not others, may be susceptible to suicide in general and suicide contagion in particular.

As noted above, suicide is problematic because it is clearly not an adaptive act from the perspective of the suicidal individual. This has led some to suggest that suicide might be enabled by a heritable genetic variation that remains in the gene pool over time because it enables suicide when the residual inclusive fitness of an individual is particularly low (e.g. de Catanzaro 1980, 1981).

Whilst a genetic 'enablement' of suicide may well be feasible, it is possible to apply an alternative cultural model using a similar logic. Specifically, if cultural variations, such as those that enable suicide, persist based on their likelihood of being adopted, then the inclusion of suicide in a culture as a meme is only viable over time if it has no systematically deleterious effect on its own reproduction. Now, one way that this could be possible is if those committing suicide were not significant contributors to the propagation of suicide themselves, so that their deaths would not negatively impact on the persistence of suicide. It is this point that provides a selectionist hypothesis for susceptibility to suicide contagion insofar as such susceptibility might be contingent on a reduced residual capacity to pass on culture. In such cases, suicide would not be maladaptive, because the suicidal individual would not be culturally 'viable'.

This model leads to an empirical prediction pertaining to those most at risk from suicide contagion. The model would predict that those susceptible to suicide contagion should be those with a reduced residual capacity to spread culture, that is, those who become socially isolated and culturally disenfranchised. Indeed, over and above the possibility that suicide may be used as a strategy for increasing a waning cultural fitness, the fact that cultures are shared means that the suicide of those whose capacity to reproduce their culture has become compromised could actually increase the overall capacity of cultural relatives to pass on shared culture, including any suicide meme. In fact, if an individual actually represented a cost to the overall reproductive potential of the shared culture to which suicide is a part, then suicide could actually have a positive effect on the likelihood that that some 'suicide culture' gets reproduced. In this model, differential ownership of the means of cultural reproduction, or simply put, marginality, would be a key variable in susceptibility to suicide contagion, and to suicide in general.

This memetic prediction is borne out by empirical research that has consistently demonstrated that low levels of social cohesion are one of the most significant risk factors in suicide (e.g. Durkheim 1970 [1897], Halbwachs 1978 [1930], Henry and Short 1954, Gibbs and Martin 1964, Maris 1981). In these cases, suicide contagion could be seen as a 'fortuitous' cultural mutation

that allows a culture to effectively rid itself of parasites that reduce its reproductive potential to maintain itself.

The tentative support for this memetic model might warrant formalisation for future research using, for instance, Hamilton's (1964) model of inclusive fitness in a cultural substrate. Specifically, as long as overall inclusive capacity of the set of culture that describes an individual is not reduced by suicide (or indeed self-sacrifice, as in the case of war), then any suggestion-by-exposure to suicide occurring within a constellation of pro-suicidal circumstances could be more likely to result in suicide. This is because the benefit of suicide to cultural relatives, in terms of enhancing their capacity to reproduce their shared culture multiplied by a degree of cultural relatedness to the suicidal individual, would be equal or superior to the direct cost to the suicidal individual of suicide in terms of any personal capacity to make such contributions.

Memetic Model of Suicide
$C_d < rB_r$

**Figure 2:** Hamilton's Rule for the Biological Communication of the Individually Maladaptive Traits. Here:  $C_d$  = Cost in cultural reproductive potential of suicide to individual,  $r$  = coefficient of relatedness, and  $B_r$  = Benefit in cultural reproductive potential of suicide to cultural relatives

With respect to practical public policy recommendations on reducing contagion, this interpretation provides an argument for curtailing any sensationalist media publicity around suicide, because such publicity might effectively increase the cultural fitness of the suicidal individual, that is, the capacity to pass on culture. Rather than merely refrain from the positive portrayal of suicide and suicide victims in the media as per current recommendations, this memetic model of suicide could be used to suggest that there could be a potential benefit (if reducing suicide levels is seen as beneficial) of not covering suicides stories at all, or at least reporting the deaths but not as suicides. Of course, there are important and controversial issues of censorship and free speech that are raised here, and it may be decided that suicide contagion is a lesser evil than what amounts to ideational eugenics.

## 7 Conclusion

This paper presents one vision of memetics, as an integrated part of applied social science investigating substantive issues of human experience. Understanding memetics as contagion psychology, using selectionist thinking to inform interpretation, is certainly not the only way to conceptualise the nascent discipline, but it is hoped that it is one that will allow memetics, after a quarter of a century of discussion, to start providing useful insight into real-world issues and problems.

## References

- Berkowitz, L., (1984). Some effects of thoughts on anti- and prosocial influence of media events: A cognitive neoassociationist analysis. *Psychological Bulletin*, **95**, 410-427.

CDC (1994) Centers for Disease Control and Prevention. Suicide Contagion and the Reporting of Suicide: Recommendations from a National Workshop. *Morbidity and Mortality Weekly Review*, **43**, (RR-6).

de Catanzaro, D., (1980). Human suicide: A biological perspective. *Behavioral and Brain Sciences*, **3**(2), 265-290.

de Catanzaro, D., (1981). *Suicide and Self-Damaging Behavior: A Sociobiological Perspective*. San Diego: Academic Press.

Colligan, M.J., Pennebaker, J.W., and Murphy, L.R., (1982) *Mass Psychogenic Illness: A Social Psychological Analysis*. New Jersey: Lawrence Erlbaum Associates.

Durkheim, E., (1970). *Suicide: A Study in Sociology*. Translation J. A. Spaulding & G. Simpson, (1951) London: Routledge & Kegan Paul. Originally published in French 1897.

Fiske, S. T. & Taylor, S. E., (1991). *Social Cognition*. New York: McGraw-Hill.

Geen, R., and Thomas, S., (1986) The immediate effects of media violence on behavior. *Journal of Social Issues*, **42**, 7-27.

Gibbs, J., and Martin, W.T. (1964). *Status Integration and Suicide*. Eugene, Oregon: University of Oregon Press.

Halbwachs, M., (1978) *The Causes of Suicide* (translation of 1930 Les Causes du Suicide) London: Routledge.

Hamilton, W., (1964). The genetical evolution of social behaviour. *Journal of Theoretical Biology*, **7**, 1-52.

Henry, A., and Short, J. (1954). *Suicide and Homicide*. New York: McGraw-Hill.

Higgins, E. T., (1989). Knowledge accessibility and activation: Subjectivity and suffering from unconscious sources. In J. S. Uleman & J. A. Bargh, (eds.), *Unintended Thought*, New York: Guildford Press, 75-123.

Jo, E. and Berkowitz, L., (1994). A priming effect analysis of media influences: An update. In J. Bryant, & D. Zillmann, (eds.). *Media Effects: Advances in Theory and Research*. Hillsdale, NJ: Erlbaum.

Kerckhoff, A. C. and Back, K. W., (1968). *The June Bug: A Study in Hysterical Contagion*. New York: Appleton-Century-Crofts.

Leyens, J. and Parke, R. D., (1975). Aggressive slides can induce a weapons effect. *European Journal of Social Psychology*, **5**, 229-239.

Maris, R., (1981). *Pathways to Suicide: A Survey of Self-Destructive Behaviors*. Baltimore: Johns Hopkins Press.

Marsden, P., (1998a). Memetics and Social Contagion: Two Sides of the Same Coin? *Journal of Memetics - Evolutionary Models of Information Transmission*, 2.  
[http://www.cpm.mmu.ac.uk/jom-emit/1998/vol2/marsden\\_p.html](http://www.cpm.mmu.ac.uk/jom-emit/1998/vol2/marsden_p.html)

Marsden, P., (1998b). Operationalising Memetics, Suicide, the Werther Effect and the work of D. P. Phillips. *Proceedings 15<sup>th</sup> International Congress on Cybernetics*. Namur, Belgium.  
<http://pespmc1.vub.ac.be/Conf/MemePap/Marsden.html>

Marsden, P., (2000) The Werther Effect: Suicide Contagion in the Internet Age - A Critical Evaluation, Theoretical Reconceptualisation and Empirical Investigation. D.Phil Thesis (Sussex). <http://www.viralculture.com/DPhilsummary.htm>

# THE INTERNET AND MEMETICS

By

**Garry Marshall**

School of Computing Science, Middlesex University, Trent Park, Bramley Road, London N14  
4XS, England.  
e-mail: Garry2@mdx.ac.uk

## Abstract

The functioning and usage of the Internet are examined in terms of memes and memetics. It is shown that memetic systems can be distinguished at various levels of Internet operation, and that these systems become increasingly simple as they move further from the user level. In this way, memetics provides a unified framework for examining the overall behaviour of the Internet and its users.

## 1. Introduction

Memetics provides a powerful new way to think about things such as, for example, creativity (Gabora, 1997). The aim of this paper is to use memetics as the basis for obtaining a better understanding of the Internet both in its operation and in the way it is used. A primary purpose of a network such as the Internet is to support communication. In this role, and especially as a carrier of e-mail messages, it serves to disperse memes, spreading them across the network rapidly and accurately. Recently, the principal use of the Internet has moved from e-mail to the World Wide Web. While e-mail is still widely used, the move to the World Wide Web has caused a shift in emphasis from carrying messages to the storage of interconnected documents (Heylighen, 1996). The results of this change are, among others, to ensure the longevity of memes and to link together related memes. The consequence of the availability of both e-mail and the World Wide Web is that the Internet is, for its users, an ideal medium for the spread, replication and storage of memes.

The Internet, like all computer networks, is designed and constructed in a layered fashion, with layers of software added to the basic hardware. The layering facilitates the synthesis of a complex and complicated system but it has always been difficult to describe the operation of the layers in a consistent way so that the activities in one layer could be related directly to those in another. More importantly, the basic rationale for layering is to facilitate meta-system construction, but it has never been entirely clear that this has been achieved in any coherent way. In any event, if a major purpose of the network is to communicate and store memes, one would expect to find memetic ideas present in its design and implementation, even though its designers may not have called on them explicitly.

In fact, as this paper shows in a limited way, the functioning of the various layers can be interpreted memetically. Such an interpretation is of value not only in linking the purpose and operation of the network but also in showing the extent to which the memetic processes the network is intended to support map onto the memetic processes carried out by the underlying

layers. A clear and direct mapping will indicate a well designed network, whereas the absence of any direct mapping will not only indicate shortcomings and inefficiencies in the way that the network operates but also help to identify their causes. This paper, then, is an attempt to expose the memetic systems and processes occurring at different levels of Internet operation. It also shows how the processes in adjacent layers relate to one another which, in turn, exposes the meta-system transitions.

## 2. Memetics at the operational level

We begin by considering the World Wide Web. Here, a system of connections, or links, is maintained in a data space so that the space can be navigated. But this data space, as with everything else on the Internet, is supported by the underlying network, and the connections of the network do not correspond directly to the connections of the data space. To ensure that the users of the virtual world provided by the data space are unaware of the underlying network, any connection that is to be followed in the data space must automatically be mapped onto a route in the network. This means that routing is an aspect of network operation that is vital to the success of the World Wide Web, just as it is for other Internet services.

Routing is, in essence, achieved with routing tables. Each node of the network has a routing table that determines from the destination of a message how that message should be forwarded from that node. A message is delivered to its destination after being forwarded in this way by a series of nodes. Internet routing, however, not only provides routes, but also provides them in a way that adapts to the state of the network (Parker, 1994). The aim of this adaptive routing is to provide all the users of the network with the best possible common level of service regardless of the pattern of traffic carried by the network. The adaptation is controlled by network management centres that collect status information from network nodes, use this information to construct a representation of the state of the network and then send out messages to change the routing tables as necessary.

This way of achieving adaptive routing can be interpreted as a memetic system. The messages carrying the changes to the routing tables can be seen as memes, since they are, in essence, information patterns. The patterns can be interpreted by nodes in such a way as to bring about the required changes to the routing tables and so to the behaviour of the network. They spread across the network by being transmitted, and can be replicated as necessary, for example, if the same changes are needed at different nodes. A network management centre causes the memes to be modified in accordance with its representation of the state of the network. The aim of the modification is to cause the overall network to remain fit for its purpose despite its changing environment. The system can be redescribed in terms of a society. Nodes can be seen as the members of a 'network society'. Their capability is rather limited in that, as far as routing is concerned, they merely forward the messages that come to them. All the same, they co-operate with each other in an altruistic way to support routes across the network. In fact, a node will have no sense of the overall route of a message, merely passing it on to the next hop of its journey. Further, it will have no knowledge of messages in other parts of the network, and yet the other messages will be delivered to the same performance specification as the ones it handles. The purpose of the memes is to ensure that the routing tables are maintained in such a way that the 'network society' meets its goal no matter how the state of the network changes. The claim that

this is a memetic system must be accompanied by the rider that it is rather a limited one. The memes do not mutate or recombine, but are modified by a central controller in response to its perception of what is necessary. In the terminology used by Wilkins (1998), the memes are instructed.

### 3. Memetics at the service level

Agent technology is becoming ever more widely used on the Internet as a means of supporting services, notably search engines and 'push' services.

An agent can be defined as a message that can be sent into the network to perform some task and to return with the result. So, to give an example, Jango (to be found at <http://www.jango.com/>) is the name of an agent that can search on-line Internet stores for product availability and price information. Clearly, this agent must be able to locate sites of on-line stores, determine the products available at each site, and find their prices. In other words, it must have access to a program that tells it how to perform the task required of it. As an information pattern that can be replicated and spread across a network, an agent is a meme. Since it carries a program for performing a task, it provides a way of disseminating this program and, when there are different agents for the same task, they provide a basis for the selection of the best way of performing that task. Now consider the situation when several agents arrive at a node to carry out their various tasks. They could simply queue to carry out their tasks in succession so that eventually all the tasks would be done. It would, however, make better use of the scarce resource at their disposal if they co-operated with each other by, for example, sharing common tasks and ordering the tasks to facilitate the execution of the overall set of tasks. In this case, the set of agents can be seen as a meme complex and, when they co-operate, they will, if not exactly recombine, at least re-program one another. In terms of a society, the agents can be seen as the members, each having a single, possibly complex, capability, namely to perform the task required of them. The overall aim of the society is to perform the set of tasks contributed by its members.

The question that arises is: How should the agents co-operate? Memes for co-operation are needed! Co-operation can take place in many ways, including those that are benevolent or selfish, disruptive or co-operative, aggressive or submissive. Agents, as representatives of their dispatchers, should presumably co-operate in the same way as would their dispatchers. This makes it hard to envisage a uniform style of co-operation for agents, and an agent that is disposed to co-operate in a particular style must be able to react appropriately to an agent that attempts co-operation in a different fashion. Further, just as norms and laws govern co-operation in the real world, so, in the field of multi-agent systems, it is appreciated that social norms and laws are required for agent co-operation (Conte and Falcone, 1997). To summarise this, we can say that when services are implemented using agent technology, one agent can be seen as a meme, and a group of agents gathered in the same place as a meme complex. In this context, a meme is, again in the terminology used by Wilkins (1998), an interactor. The individual memes of a meme complex interact with each other, modifying themselves as required to achieve the overall aim of the group, which is to achieve their set of tasks in a way that is to the general benefit.

#### **4. Memetics at the user level**

The Internet, as mentioned earlier, supports the rapid and accurate world-wide transmission of the memes of its users. This is a supplement to, and not a replacement for, the traditional means of spreading memes. It is also, clearly, a means which makes use of a different medium. To the extent that the characteristics of the medium differ from those of the traditional medium, so the result of dispersing memes is, in turn, different.

It has been argued that the Internet's capability to spread memes across the world both accurately and instantaneously supports a tendency towards homogeneity in world culture (Heylighen, 1996). Memes can appear at much the same time in different parts of the world regardless of geographical and cultural boundaries to exert their effects. It has also been argued that the speed of transmission, and the resulting rapid cascade of memes across the Internet, makes it more difficult to distinguish between the more and less valuable memes (Taylor, 1996). There is a premium on short, catchy memes as opposed to more complex memes such as lengthy stories. Infectiousness assumes an importance far greater than that of attributes that may well have greater long-term value such as utility and authority.

With these insights, it is possible to sketch the essential differences between virtual and real-world communities. Virtual communities are not structured in the same way as real-world communities. Constraints of geography and status do not come into play: what matters is a common interest. But a diverse collection of people, perhaps drawn from all parts of the world and united by only a common interest, needs to construct its own culture. The network facilitates this with communication and the spread of memes. But, by comparison with the real world, memes are spread rapidly and accurately. This causes virtual communities to develop cultures that are narrow, often extreme and, in consequence, rather precarious (Marshall, 1998). Their precarious nature is re-inforced by the favouring of infectious memes over memes that might bring greater benefits in the long term, in that their adoption can result in a gap between the conceptions of the virtual world and those of the real world that is so great as to become unsustainable (Umpleby, 1996).

The increased use of agents brings another aspect to the situation in that they can be used to control the memes to which their dispatchers are exposed. Agents are a key element in the so-called 'push' technology with which users can specify the sort of information they want to receive, after which the network will 'push' that, and only that, information to them. As with any form of 'narrowcasting' this leads to the reinforcement of existing beliefs and the avoidance of the uncertainty associated with opposing, conflicting or even just different ideas (Salem and Gratz, 1997). This kind of selective attention produces individuals who are unaware of, or even afraid of, other views and of groups holding these views. In turn, this leads to social fragmentation and the production of incompatible social segments.

As a final point, the gulf between those with access to the network and those who do not is amplified by the fact that these groups exchange different memes in different ways. As Internet culture develops and moves farther from real-world culture, it becomes harder, not necessarily to gain access, but to join in effectively once access is gained. Besides this, there is the very real possibility of those with access to the network becoming cut off from their real-world history.

They may have no interest in it: those who do may not share the technical and cultural interests of those with access. The consequence can be a separation of those who spend most of their time on the network from their own history, which can only serve to widen further the gulf between those with access and those without.

## 5. Conclusions

This paper has shown that activities taking place at various levels in the organisation of the Internet can be interpreted in terms of memes and memetics. This provides a unified framework for examining the overall functioning and behaviour of the Internet. Given that the Internet is widely used for the communication of memes, this unified analysis has the real benefit of providing an account of the operation of the network that is coherent with the usage of the network. To the extent that the Internet is an archetypal network, these conclusions apply equally to any computer network.

The levels have been examined, starting from the bottom (hardware) level and working towards the top (user) level, to reveal their embedded memetic systems. The findings of the previous sections show that the simple memetic system supported at the operational level, in which the memes are instructed, is less complex than the system supported at the service level, where the memes are interactors. Again, the system at the service level is less complex than that at the user level which, with certain reservations, is a full-blown memetic system. Thus, the memetic system supported at a particular level is always less complex, and exerts greater control, than that supported by the level above. This finding is in tune with the rationale for adopting a layered approach to the design of a complex system, which is that the addition of a layer adds further capabilities to the overall system by building on those that already exist. In this way, the network supports the capability that users require by progressively building up this capability in a layered way with clear meta-system transitions. The memetic approach adopted here has made it possible to reveal this consistency of design and operation in a coherent way.

## 6. References

Conte, R. and Falcone, R., 'ICMAS '96: Norms, Obligations, and conventions', AI Magazine, Winter 1997, pp 145-147.

Gabora, L. M., 'Memes: the creative spark', Wired, June 1997, pp 110-112.

Heylighen, F., 'Evolution of memes on the network: from chain letters to global brain', Ars Electronica Catalogue, Ingrid Fischer (ed.), (Springer, Vienna/New York, 1996).

Marshall, G. J., 'The nature and culture of cyberspace', Cybernetics and Systems '98, Robert Trappl (ed.), (Austrian Society for Cybernetic Studies, Vienna, 1998), pp 293-298.

Parker, T., 'TCP/IP', (Sams Publishing, Indianapolis, 1994). Salem, P. and Gratz, R., 'Social development and electronic communication technologies: a comparison and contrast of developed and developing countries', Proc. World Multiconference on Systems, Cybernetics and Informatics, vol. 3, N. Callaos et al, (ed.), (IIIS, 1997), pp 475-481.

Taylor M. A., 'Fiction as artificial life: exploring the ideoosphere', Cybernetics and Systems '96, Robert Trappl (ed.), (Austrian Society for Cybernetic Studies, Vienna, 1996), pp 893-896.

Umpleby, S. A., 'Four models from cybernetics to guide our understanding of cyberspace', Cybernetics and Systems '96, Robert Trappl (ed.), (Austrian Society for Cybernetic Studies, Vienna, 1996), pp 868-870.

Wilkins, J. S., 'What's in a meme?: reflections from the history and philosophy of evolutionary biology.' Journal of Memetics, 2(1), 1998.

# THOUGHTS AS TOOLS: THE MEME IN DANIEL DENNETT'S WORK

By  
Kelby Mason

Most people interested in memes know of Daniel Dennett, at least by name; he seems to be the resident philosopher-mascot and is often mentioned in the context of memetics. It is therefore worthwhile to examine more closely his work in philosophy, to see how it relates to his memetics. Originally, I was going to discuss some of the objections to memetics raised by orthodox biologists, and how Dennett's ontology might deal with them. This proved to be beyond the scope of this paper, however, and instead I have chosen to focus specifically on the claim that memetics is reductionistic. I shall examine first Dennett's naturalism, then how this interacts with the alleged reductionism. Next I discuss the apparent threat of memetics to humanity's self-image, and finally some genuine problems posed by Dennett's treatment of memetics.

## THREE NATURALIST SLOGANS

Daniel Dennett loves stories, metaphors and thought-experiments, and so it is suitably Dennettian to start this essay with a cartoon. In *Consciousness Explained* Dennett reproduces an illustration from *American Scientist*: two men stand before a complicated mathematical proof on a blackboard. At either end of the proof are the cartoonist's idea of arcane mathematical symbols, but in the middle the words "Then a Miracle Occurs". One of the two gentlemen suggests to the other, "I think you should be more explicit here in step two".

A little further on in the same book, the very first ground rule for Dennett's explanation of consciousness is "No Wonder Tissue allowed"; and a third slogan is provided in *Darwin's Dangerous Idea: Skyhooks or Cranes?* Skyhooks are miraculous forces postulated simply because of a failure to imagine how natural, material forces could produce the object of study. Cranes, on the other hand, are just such natural forces, and thus thoroughly respectable when used in explanation.

The three are equivalent: No Wonder Tissue Allowed, No Miracles, and No Skyhooks. In short, to co-opt a more familiar and less cute slogan, *Natura saltus non facit*; and therefore *nobis saltus facere non licet* in our explanations. Dennett's world-view is explicitly naturalistic; everything, including human consciousness, is to be explained in natural, not supernatural, terms. It may well turn out that science (or, if you prefer, Science) cannot explain everything. Until this is actually *shown* to be the case, however, Dennett will keep working away at his scientific edifice, forgoing miracles and skyhooks for slower yet more reliable cranes.

It is this "Standard Scientific Epistemology and Metaphysics" that Dennett believes to have led to his views on consciousness, intentionality and free will [2]. He supposes that he is simply working out the implications of "everyday science", what any materialist is led to believe. More importantly for present purposes, these implications include the *memetic* stance, as the best non-miraculous explanation of consciousness and culture.

It is also this "scientism" (better called naturalism) that has raised strong opposition to his whole enterprise. Dennett himself puts it well:

Those who see themselves as outside the gates of the scientific culture (which of course includes not just scientists, but all the science-literate, science-friendly people) see it as approximately as threatening to their own sense of power as Martians with advanced technology beyond our ken would seem to us. [9]

Or, as Bo Dahlbom says,

When peddled second-hand...Dennett comes off as something of a villain: a science-crazed mechanist, with a bleak and inhumane world-view, a rather superficial engineer compared to deep and serious thinkers like Thomas Nagel and John Searle" [2]

Indeed Dennett has concocted some disturbing descriptions of the memetic picture. The brain, he says, is "a sort of dung heap in which the larvae of other people's ideas renew themselves" [5]. Now of course the brain is not *really* a dung heap (at least *Dan's* isn't), but the image remains. If Dennett himself, memophile as he is, finds memetics at first "unsettling, even appalling" (*ibid.*), how will memophobes react?

Predictably enough, not well. Consider, for instance, the following from Richard Braddock: The meme theory is yet another attack on human subjectivity. Our complex social development is first simplified into technological progress, then reduced to culture, and finally explained away through the biology of memes. Our creativity and imagination as humans is once again denied. [1]

We can take Braddock as spokesman of a type of opponent not just to memes, but to any materialist account of humanity. Dennett, following Owen Flanagan, calls them "the New Mysterians" [5] because they are effectively arguing that some of the universe (specifically, the bit called "us") will *always* be inexplicable by science. The aforementioned Searle and Nagel are themselves leading advocates of this anti-materialism; and both are long-time foes of Dennett.

As shown by Braddock, the new mysterians generally have two concerns about memetics. The first is epistemological: memetics "explains away" culture and consciousness, and thus is not really an adequate theory. The second is moral/aesthetic: even if memetics were right, it would have such terrible consequences for our sense of creativity and freedom that we are better off leaving that Pandora's Box closed. Both concerns can be evaluated by with reference to Dennett's work.

## **WHAT'S WRONG WITH REDUCTIONISM, ANYWAY?**

Every field has its own peculiar four-letter words that are substituted for genuine argument; once applied to an opponent, they leave no apparent need for further dispute. The list of dirty words in biology includes Lamarckism, vitalism, essentialism and, increasingly, reductionism. Dawkins remarks [4] that "a kind of "holister than thou" self-righteousness has become fashionable" in

evolutionary theory. Dawkins should know, having been called a reductionist more often than anyone except perhaps Skinner, and Dennett himself.

In *Darwin's Dangerous Idea*, Dennett introduces a distinction between "good" and "greedy" reductionism. Good reductionism is basically materialism, a belief that material causes are necessary and sufficient for any phenomena. This sort of ontological reductionism does not, however, entail a commitment to theory reduction, the view that we should replace biology by physics. The difference between the two is well documented in philosophy of science.

Greedy reductionism, on the other hand, is when explaining is replaced by explaining away, "when overzealousness leads to falsification of the phenomena" and a denial of "the existence of real levels, real complexities, real phenomena" [7]. The explanatory paradigm is *too* simple and, accordingly, the data have to be squeezed and mangled beyond recognition to fit. This is the marrow of Dennett's earlier criticism of Skinner [6], that he mistakes good reductionism for greedy, and reductionism *qua* materialism for theory reduction.

In a sense, the real sin of greedy reductionism is against what Dahlbom calls "Dennett's very modern, very American, belief in hard work" [2]. The greedy reductionist is just lazy, glossing over complications and inconsistencies for the sake of her one simple, better yet simplistic, idea. With these important distinctions in mind, then, we can now ask what is meant by calling memetics reductionistic.

It is fairly clear what is meant by those other biological insults, all of them conjuring up the bad old pre-Darwinian days. Reductionism cannot, however, be such an insult, for while neo-Darwinism is definitely opposed to Lamarckism, vitalism and essentialism, it is by no means so obviously opposed to "reductionism". Neo-Darwinism itself could easily be labelled reductionist, insofar as it "explains away" complex adaptations by "mere" changes of gene frequencies in populations; and Dennett calls "Darwin's dangerous idea reductionism incarnate".

It cannot be therefore that when memetics is called "reductionist", what is meant is simply that it is materialist. We are all, I hope, materialists here, and if anyone wants to argue that we are begging the question by supposing that *some* materialist explanation of culture and mind is possible, then I agree. *Nolo contendere*: just as Dennett begs the question by calling good reductionism *good*, let us beg the question of materialism.

What is interesting is not the extreme claim that materialism itself can never explain mind/culture but the milder claim that *memetics* is not an adequate materialist explanation, that it is *greedy* reductionism. This is at the heart of Braddock's criticism, for he sees memetics as explaining away culture as the result of biology. If this were the case, then we should expect memetic papers such as "The Reproductive Advantages of Sonnet XVIII".

Yet memetics is not that sort of "biological reductionism", and must not be confused with socio-biology in the style of Edward O. Wilson. Rather, it seeks to explain human behaviour largely as the product of a new evolutionary domain which is not biological, although it builds on the biological. Precisely because it involves memes, it does not explain everything as due to biological foundations. Behaviour is explained by its relation, not as in socio-biology to genetic

reproduction, but to memetic reproduction; a crucial message of memetics is that the interests of these new replicators do not always coincide with those of the old.

It is not necessary, however, to confound memetics with socio-biology, to claim that it is greedy reductionism. One still might maintain that the mechanism of natural selection of memes is not powerful enough to explain mind/culture. How *could* memes possibly explain all the diversity of culture, as well as the subjectivity of consciousness?

Dennett sees this sort of argument as simply intellectually bankrupt. It is something he has long fought in philosophy; he calls it "Philosopher's Syndrome: mistaking a failure of imagination into an insight into necessity" [5]. To those who cannot understand how complexity can emerge from a simple process, Dennett replies: "Try harder". His work in philosophy of mind is largely devoted to showing that we can imagine such things, if only we try hard enough. Similarly, in much of *Darwin's Dangerous Idea*, Dennett tries to show that evolution by natural selection can explain natural history, without appealing for divine design.

Dennett likes to compare the mind to a computer, even calling it a "serial virtual machine implemented on the parallel hardware of the brain" [5]. Such comparisons are widespread in cognitive science, and John Searle parodies them by reminding us that similar comparisons were made to hydraulic systems and telephone switchboards when those technologies were new and fashionable [12]. Yet Searle has missed an important point: in all cases, the comparisons have not arisen because the Artificial Intelligentsia see everyone around them as zombie-like robots. Rather, it is precisely because they appreciate just how complex and impressive that technology can be, while having a thoroughly material basis.

Similarly, Dennett does not describe the mind as a meme complex because of a dehumanised view of humanity, but because he recognises the potential of the memetic stance to explain seemingly non-adaptive human behaviour. More importantly, he has learnt from Darwin that simple algorithms can create complex results, or in his own terms, that enough Cranes can simulate a Skyhook. Anyone suggesting that *au contraire* a memetic approach "reduces" or "explains away" mind/culture obviously has both an impoverished sense for the richness of biology and evolutionary theory, and a poor imagination.

## IS MEMETICS AN ANTI-HUMANISM?

All this should suffice to show that memetics is not *a priori* greedy reductionism. Whether or not a workable science of memetics is possible is something only to be determined by the efforts of pioneer memeticists. This still leaves unresolved the other concern for the mysterians, that we should be worse off if memetics were successful. They argue that the more our "manifest image" of ourselves, as experiencing and effective agents, is affected by scientific investigation, the more our "creativity and imagination" is denied.

Now this sort of concern was widely voiced against Dawkins' *Selfish Gene*, and prompted a detailed response in Chapter 2 of *The Extended Phenotype*. There Dawkins criticises the "belief that genes are somehow super-deterministic". Even without considering this response, we can see that both Dawkins and Dennett identify with their memes, not their genes. Dawkins says

"We are built as gene machines and cultured as meme machines, but we have the power to turn against our creators. We, alone on earth, can rebel against the tyranny of the selfish replicators." [3]

and Dennett

there is, in the basement, a persisting tension between the biological imperative of the genes and the imperatives of the genes, but we would be foolish to side with' our genes [5]

If we are more closely related to our memes than our genes, then, we might ask with Dennett "Have we broken the tyranny of the selfish genes, only to be taken over by the selfish memes?" [7]

The problem with this question is that it misplaces our selves, where the "we" apparently controlled by our memes is supposed to be. For according to Dennett, "our *selves* have been created out of the interplay of memes" and it cannot be "memes versus us," because earlier infestations of memes have already played a major role in determining who or what we are. The "independent" mind struggling to protect itself from alien and dangerous memes is a myth. Finally, "what makes a person the person he or she is are the coalitions of memes that govern---that play the long-term roles in determining which decisions are made along the way". [7]

Thus to think that in memetics we hand over our self-control to our memes instead, who then act as the Sinister Puppet Master, is to create an independent self that never has existed. It is as if we were to say "Who is responsible for what I do, is it myself? How, then, can *I* also be responsible for my actions, when myself is?" We *are* our memes, and the memes are running the show; therefore, so are we.

Moreover, the memetic stance is just another aspect of Dennett's broader strategy. In *Elbow Room* Dennett describes how as a student he had intended to be a sculptor and that he has "never abandoned the methods [he] developed in the studio, but simply changed media". He is making a different point, but this is also exactly how he sees ideas, as something to work at and something to work *with*. Ideas are tools and material for those tools. In fact, that whole book deals with the question of free will by studying our intuitions and how they mislead us, or rather how philosophers mislead us and each other by their "intuition pumps".

Dennett is in a general sense pursuing the Wittgensteinian project, revealing how we are trapped by "word games": the title of one chapter in *Consciousness Explained* is "How Words Do Things With Us". Dennett is trying to gain an understanding of ourselves, and to dissolve constructed mysteries. His is also similar to the Freudian project: "our therapeutic aim...[is] to restore the ego...to give it back the command over the id which it has lost" [11].

Far from crippling us with a sense of our own impotence, then, memetics should empower us. For it is not just as a science of thought that Dennett is interested in it, but also as a *technology* of thought. This conception of thoughts is one that encourages us to take control of them, and use them to their full advantage: if we can now get a science (in the form of memetics) of the ideas

that rule us then, according to Dennett, we will have that much more power over ourselves and our environment.

## TWO DEBATABLE ISMS

Before concluding, I should like to say a few words on the possible limitations of Dennett's work, or at least indicate the underlying biases that colour his discussion of memes. For in spite of his claims, mentioned above, to be just a believer in traditional science, Dennett holds some unorthodox views on evolution. There is, for instance, what he himself admits to be a "shocking view", that we must apply the intentional stance to evolution in order to understand it; in other words, we need a personified Mother Nature.

Those of us particularly who are not biologists, and thus learn their evolutionary theory second-hand, should be careful not to take Dennett's orthodox credentials at face value. My questioning the two following aspects of Dennett's approach to memetics should not be interpreted as disagreement, but merely as a note of caution. Dennett does not *necessarily* have the right framework, and we need not only to decide for ourselves, but also to realise that there is a decision to be made in the first place.

**I. Adaptationism** This is neither the time nor place for a dissertation on the pros and cons of adaptationism. It suffices to point out that the debate over it is still a live one. Dennett's book is called Darwin's Dangerous Idea, not evolution's dangerous idea. He has correctly identified the central tenet of Darwinism as natural selection; but we must recognise that natural selection is not all there is to evolution. There may be other important mechanisms of memetic evolution, memetic drift, epimemetic constraints; or perhaps all memetic explanations will be just-so stories. Acknowledging the possibility of the former, however, does not mean abandoning materialism; the recognition of *other* Cranes does not equate to holding out for Skyhooks.

**II. Atomism** Similarly, gene selectionism is by no means universally accepted in biology. We need not therefore expect that the only fruitful research strategy in memetics will be an equally atomistic one that focuses on individual units of memes. Organism-centred selection worked for a long time in evolutionary biology; by analogy, a memetics centred on meme-complexes rather than single memes might be equally successful, not to mention easier to manage.

There is much that the memeticist can better understand from a familiarity with Dennett's work. In the first place, the firm situation of memetics within the naturalist tradition. Secondly, exactly how memetics is "reductionistic", what this entails, and what it does not. Memetics need not "explain away" anything, or threaten our moral ideas of self-control. While there may also be dangers in reading Dennett as gospel, these are not insuperable and may be overcome. Dahlbom calls Dennett an optimist [2] and ultimately his memetics too is optimistic. Via the technology of memetics, we should be even more able to take control of our selves, the encouragement of which has been Dennett's project all along.

## Bibliography

1. Richard Braddock "Never Mind the Cyberbollocks..."

<http://web.aec.at/meme/symp/panel/msg00076.html>

2. Bo Dahlbom (ed.) *Dennett and His Critics* Blackwell 1993
3. Richard Dawkins *The Selfish Gene* 2nd edition Oxford University Press 1989
4. Richard Dawkins *The Extended Phenotype* Oxford University Press 1983
5. Daniel Dennett *Consciousness Explained* Penguin Books 1993
6. Daniel Dennett *Brainstorms* Bradford Books 1978
7. Daniel Dennett *Darwin's Dangerous Idea* Penguin Books 1995
8. Daniel Dennett *Elbow Room* MIT Press 1984
9. Daniel Dennett Review of Burkert for *The Sciences* (forthcoming)
10. David J. Depew and Bruce Weber *Darwinism Evolving* MIT Press 1996
11. Sigmund Freud *The Question of Lay Analysis* 1926, in *The Penguin Freud Library* 1986
12. John Searle *Mind, Brains and Science* BBC Publications 1984
13. Kim Sterelny & Paul Griffiths *Sex and Death: An Introduction to Philosophy of Biology* University of Chicago Press (Forthcoming February 1999)

## THE FUTURE OF MAN

By  
Sir Peter Medawar  
(1995)

In this last lecture, I shall discuss the origin in human beings of a new, a non-genetical, system of heredity and evolution based upon certain properties and activities of the brain. The existence of this non-genetical system of heredity is something you are perfectly well aware of. It was not biologists who first revealed to an incredulous world that human beings have brains; that having brains makes a lot of difference; and that a man may influence posterity by other than genetic means. Yet much of what I have read in the writings of biologists seems to say no more than this. I feel a biologist should contribute something towards our *understanding* of the distant origins of human tradition and behavior, and this is what I shall now attempt. The attempt must be based upon hard thinking, as opposed to soft thinking; I mean, it must be thinking that covers ground and is based upon particulars, as opposed to that which finds its outlet in the moping or exaltation of poetistic prose.

It will make my argument clearer if I build it upon an analogy. I should like you to consider an important difference between a juke-box and a gramophone --- or, if you like, between a barrel-organ and a tape-recorder. A juke-box is an instrument which contains one or more gramophone records, one of which will play whatever is recorded upon it if a particular button is pressed. The act of pressing the button I shall describe as the ``stimulus.'' The stimulus is specific: to each button there corresponds one record, and vice versa, so that there is a one-to-one relationship between stimulus and response. By pressing a button --- any button --- I am, in a sense, instructing the juke-box to play music; by pressing this button and not that, I am instructing it to play one piece of music and not another. But --- I am not giving the juke-box *musical* instructions. The musical instructions are inscribed upon records that are part of the juke-box, not parts of its environment: what a juke-box or barrel-organ can play on any one occasion depends upon structural or inbuilt properties of its own. I shall follow Professor Joshua Lederberg in using the word ``elective'' to describe the relationship between what the juke-box plays and the stimulus that impinges upon it from the outside world.

Now contrast this with a gramophone or any other reproducing apparatus. I have a gramophone and one or more records somewhere in the environment outside it. To hear a particular piece of music, I go through certain motions with switches, and put a gramophone record on. As with the juke-box I am, in a sense, instructing the gramophone to play music, and a particular piece of music. But I am doing more than that: I am giving it musical instructions, inscribed in the grooves of the record I make it play. The gramophone itself contains no source of musical information. My relationship to the gramophone --- again following Lederberg --- I shall describe as ``instructive''; for, in a sense, I *taught* it what to play. With the juke-box, then --- and the same goes for a musical-box or barrel-organ --- the musical instructions are part of the system that responds to stimuli, and the stimuli are elective: they draw upon the inbuilt capabilities of the instrument. With a gramophone, and still more obviously with a tape recorder, the stimuli and instructive: they endow it with musical capabilities; they import into it musical information from the world outside.

It is we ourselves who have made juke-boxes and gramophones, and who decide what, if anything, they are to play. These facts are irrelevant to the analogy I have in mind, and can be forgotten from now on. Consider only that organism on the one hand --- juke-box or gramophone; and on the other hand, stimuli which impinge upon that organism from the world about it.

During the past ten years, biologists have come to realize that, by and large, organisms are very much more like juke-boxes than gramophones. Most of those reactions of organisms which we were formerly content to regard as instructive are in fact elective. The instructions an organism contains are not musical instructions inscribed in the grooves of a gramophone record, but *genetical* instructions embodied in chromosomes and nucleic acids. Let me give examples of what I mean.

The oldest example, and the most familiar, concerns the change that comes over a population of organisms when it undergoes an evolution. How should we classify the environmental stimuli that cause organisms to evolve? The Lamarckian theory, the theory that acquired characters can be inherited, is, in its most general form, an *instructive* theory of evolution. It declares that the environment can somehow issue genetical instructions to living organisms --- instructions which, duly assimilated, can be passed on from one generation to the next. The blacksmith who is usually called upon to testify on these occasions gets mightily strong arms from forging; somehow this affects the cells that manufacture his spermatozoa, so that his children start life especially well able to develop strong arms. I have no time to explain our tremendous psychological inducement to believe in an instructive or Lamarckian theory of evolution, though in a somewhat more sophisticated form than this. I shall only say that every analysis of what has appeared to be a Lamarckian style of heredity has shown it to be *non*-Lamarckian. So far as we know, the relationship between organism and environment in the evolutionary process is an elective relationship. The environment does *not* imprint genetical instructions upon living things.

Another example: bacteriologists have known for years that if bacteria are forced to live upon some new unfamiliar kind of foodstuff or are exposed to the actions of an anti-bacterial drug, they acquire the ability to make use of that new food, or to make the drug harmless to them by breaking it down. The treatment was at one time referred to as the *training* of bacteria --- with the clear implication that the new food or drug *taught* the bacteria how to manufacture the new ferments upon which their new behavior depends. But it turns out that the process of training belies its name: it is not instructive. A bacterium can synthesize only those ferments it is genetically entitled to synthesize. The process of training merely brings out or exploits or develops an innate potentiality of the bacterial population, a potentiality underwritten or subsidized by the particular genetic make-up of one or another of its members.

The same argument probably applies to what goes on when animals develop. At one time there was great argument between ``preformationists'' and those who believed in epigenesis. The preformationists declared that all development was an unfolding of something already there; the older extremists, whom we now laugh at, believed that a sperm was simply a miniature man. The doctrine of epigenesis, in an equally extreme form, declared that all organisms begin in a homogeneous state, with no apparent or actual structure; and that the embryo is moulded into its adult form solely by stimuli impinging upon it from outside. The truth lies somewhere between

these two extreme conceptions. The genetic instructions are preformed, in the sense that they are already there, but their fulfilment is epigenetic ---- an interpretation that comes close to an elective theory of embryonic development. The environment brings out potentialities present in the embryo in a way which (as with the buttons on a juke-box) is exact and discriminating and specific; but it does not **instruct** the developing embryo in the manufacture of its particular ferments or proteins or whatever else it is made of. Those instructions are already embodied in the embryo: the environment causes them to be carried out.

Until a year or two ago we all felt sure that **one** kind of behavior indulged in by higher organisms did indeed depend on the environment as a teacher or instructor. The entry or injection of a foreign substance into the tissues of an animal brings about an immunological reaction. The organism manufactures a specific protein, an "antibody," which reacts upon the foreign substance, often in such a way as to prevent its doing harm. The formation of antibodies has a great deal to do with resistance to infectious disease. The relationship between a foreign substance and the particular antibody it evokes is exquisitely discriminating and specific; one human being can manufacture hundreds --- conceivably thousands --- of distinguishable antibodies, even against substances which have only recently been invented, like some of the synthetic chemicals used in industry or in the home. Is the reaction instructive or elective? --- surely, we all felt, instructive. The organism learns from the chemical pattern of the invading substance just how a particular antibody should be assembled in an appropriate and distinctive way. Self-evident though this interpretation seems, many students of the matter are beginning to doubt it. They hold that the process of forming antibodies is probably elective in character. The information which directs the synthesis of particular antibodies is part of the inbuilt genetical information of the cells that make them; the invading foreign substance exploits that information and brings it out. It is the juke-box over again. I believe this theory is somewhere near the right one, though I do not accept some of the special constructions that have been put upon it.

So in spite of all will to believe otherwise, and for all that it seems to go against common sense, the picture we are forming of the organism is a juke-box picture --- a juke-box containing genetical instructions inscribed upon chromosomes and nucleic acids in much the same kind of way as musical instructions are inscribed upon gramophone records. But what a triumph it would be if an organism could accept information from the environment --- if the environment could be made to act in an instructive, not merely an elective, way! A few hundred million years ago a knowing visitor from another universe might have said: "It's a splendid idea, and I see the point of it perfectly: it would solve --- or could solve --- the problems of adaptation, and make it possible for organism to evolve in a much more efficient way than by natural selection. But it's far too difficult: it simply can't be done."

But you know that it has been done, and that there is just one organ which can accept instruction from the environment: the brain. We know very little about it, but that in itself is evidence of how immensely complicated it is. The evolution of a brain was a feat of fantastic difficulty --- the most spectacular enterprise since the origin of life itself. Yet the brain began, I suppose, as a device for responding to elective stimuli. **Instinctive** behavior is behavior in which the environment acts electively. If male sex hormones are deliberately injected into a hen, the hen will start behaving in male-like ways. The potentiality for behaving in a male-like manner must therefore have been present in the female; and by pressing (or, as students of behavior usually

say, ``releasing") the right button the environment can bring it out. But the higher parts of the brain respond to instructive stimuli: we **learn**.

Now let me carry the argument forward. It was a splendid idea to evolve into the possession of an organ that can respond to instructive stimuli, but the idea does not go far enough. If that were the whole story, we human beings might indeed live more successfully than other animals; but when we died, a new generation would have to start again from scratch. Let us go back for a moment to genetical instructions. A child at conception receives certain genetical instructions from its parents about how its growth and development are to proceed. Among these instructions there must be some which provide for the issue of further instructions; I mean, a child grows up in such a way that it, too, can eventually have children, and convey genetical instructions to them in turn. We are dealing here with a very special system of communication: a **hereditary** system. There are many examples of systems of this kind. A chain letter is perhaps the simplest: we receive a letter from a correspondent who asks us to write to a third party, asking him in turn to write a letter of the same kind to a fourth, and so on --- a hereditary system. The most complicated example is provided by the human brain itself; for it does indeed act as intermediary in a hereditary system of its own. We do more than learn: we teach and hand on; tradition accumulates; we record information and wisdom in books.

Just as a hereditary system is a special kind of system of communication --- one in which the instructions provide for the issue of further instructions --- so there is a specially important kind of hereditary system: one in which the instructions passed on from one individual to another change in some systematic way in the course of time. A hereditary system with this property may be said to be conducting or undergoing an **evolution**. Genetic systems of heredity often transact evolutionary changes; so also does the hereditary system that is mediated through the brain. I think it is most important to distinguish between four stages in the evolution of a brain. The nervous system began, perhaps, as an organ which responded only to elective stimuli from the environment; the animal that possessed it reacted instinctively or by rote, if at all. There then arose a brain which could begin to accept instructive stimuli from the outside world; the brain in this sense has dim and hesitant beginnings going far back in geological time. The third stage, entirely distinguishable, was the evolution of a non-genetical system of heredity, founded upon the fact that the most complicated brains can do more than merely receive instructions; in one way or another they make it possible for the instructions to be handed on. The existence of this system of heredity --- of tradition, in its most general sense --- is a defining characteristic of human beings, and it has been important for, perhaps, 500,000 years. In the fourth stage, not clearly distinguishable from the third, there came about a systematic change in the nature of the instructions passed on from generation to generation --- an evolution, therefore, and one which has been going on at a great pace in the past 200 years. I shall borrow two words used for a slightly different purpose by the great demographer Alfred Lotka to distinguish between the two systems of heredity enjoyed by man: **endosomatic** or internal heredity for the ordinary or genetical heredity we have in common with animals; and **exosomatic** or external heredity for the non-genetic heredity that is peculiarly our own --- the heredity that is mediated through tradition, by which I mean the transfer of information through non-genetic channels from one generation to the next.

I am, of course, saying something utterly obvious: society changes; we pass on knowledge and skills and understanding from one person to another and from one generation to the next; a man can indeed influence posterity by other than genetic means. But I wanted to put the matter in a way which shows that we must not distinguish a strictly biological evolution from a social, cultural or technological evolution: **both** are biological evolutions: the distinction between them is that the one is genetical and the other is not.

What, then, is to be inferred from all this? What lessons are to be learned from the similarities and correspondences between the two systems of biological heredity possessed by human beings? The answer is important, and I shall now try to justify it: the answer, I believe, is almost none.

It is true that a number of amusing (but in one respect highly dangerous) parallels can be drawn between our two forms of heredity and evolution. Just as biologists speak in a kind of shorthand about the ``evolution'' of hearts or ears or legs --- it is too clumsy and long-winded to say every time that these organs participate in evolution, or are outward expressions of the course of evolution --- so we can speak of the evolution of bicycles or wireless sets or aircraft with the same qualification in mind: they do not really evolve, but they are appendages, exosomatic organs if you like, that evolve with us. And there are many correspondences between the two kinds of evolution. Both are gradual if we take the long view; but on closer inspection we shall find that novelties arise, not everywhere simultaneously -- pneumatic tires did not suddenly appear in the whole population of bicycles -- but in a few members of the population: and if these novelties confer economic fitness, or fitness in some more ordinary and obvious sense, then the objects that possess them will spread through the population as a whole and become the prevailing types. In both styles of evolution we can witness an adaptive radiation, a deployment into different environments: there are wireless sets not only for the home, but for us in motor-cars or for carrying about. Some great dynasties die out ---airships, for instance, in common with the dinosaurs they were so often likened to; others become fixed and stable: toothbrushes retained the same design and construction for more than a hundred years. And, no matter what the cause of it, we can see in our exosomatic appendages something equivalent to vestigial organs: how else should we describe those functionless buttons on the cuffs of men's coats?

All this sounds harmless enough: why should I have called it dangerous? The danger is that by calling attention to the similarities, which are not profound, we may forget the **differences** between our two styles of heredity and evolution; and the differences between them are indeed profound. In their hunger for synthesis and systematization, the evolutionary philosophers of the nineteenth century and some of their modern counterparts have missed the point: they thought that great lessons were to be learnt from the similarities between Darwinian and social evolution; but it is from the differences that all the great lessons are to be learnt. For one thing, our newer style of evolution is Lamarckian in nature. The environment cannot imprint genetical information upon us, but it can and does imprint non-genetical information which we can and do pass on. Acquired characters are indeed inherited. The blacksmith was under an illusion if he supposed that his habits of life could impress themselves upon the genetic make-up of his children; but there is no doubting his ability to teach his children his trade, so that they can grow up to be as stalwart and skillful as himself. It is because this newer evolution is so obviously Lamarckian in character that we are under psychological pressure to believe that genetical

evolution must be so too. But although one or two biologists are still feebly trying to graft a Lamarckian or instructive interpretation upon ordinary genetical evolution, they are not nearly so foolish or dangerous as those who have attempted to graft a Darwinian or purely elective interpretation upon the newer, non-genetical, evolution of mankind.

The conception I have just outlined is, I think, a liberating conception. It means that we can jettison all reasoning based upon the idea that changes in society happen in the style and under the pressures of ordinary genetic evolution; abandon any idea that the direction of social change is governed by laws other than laws which have at some time been subject of human decisions or acts of mind. That competition between one man and another is a necessary part of the texture of society; that societies are organisms which grow and must inevitably die; that division of labor within a society is akin to what we can see in colonies of insects; that the laws of genetics have an overriding authority; that social devolution has a direction forcibly imposed upon it by agencies beyond man's control --- all these are biological judgments; but, I do assure you, bad judgments based upon a bad biology. In these lectures you will have noticed that I advocate a "humane" solution to the problems of eugenics, particularly of the problems of those who have been handicapped by one or another manifestation of the ineptitude of nature. I have not claimed, and do not now claim, that humaneness is an attitude of mind enforced or authorized by some deep inner law of exosomal heredity: there are technical reasons for supposing that no such laws can exist. I am not warning you against quack biology in order to set myself up as a rival peddler of patent medicines. What I do say is that our policies and intentions are not to be based upon the supposition that Nature knows best; that we are at the mercy of natural laws, and flout them at our peril.

It is a profound truth --- realized in the nineteenth century by only a handful of astute biologists and by philosophers hardly at all (indeed, most of those who held and views on the matter held a contrary opinion) --- a profound truth that Nature does **not** know best; that genetical evolution, if we choose to look at it liverishly instead of with fatuous good humor, is a story of waste, makeshift, compromise and blunder.

I could give a dozen illustrations of this judgment, but shall content myself with one. You will remember my referring to the immunological defenses of the body, the reactions that are set in train by the invasion of the tissues by foreign substances. Reactions of this kind are more than important: they are essential. We can be sure of this because some unfortunate children almost completely lack the biochemical aptitude for making antibodies, the defensive substances upon which so much of resistance to infectious disease depends. Until a few years ago these children died, because only antibiotics like penicillin can keep them alive; for that reason, and because the chemical methods of identifying it have only recently been discovered, the disease I am referring to was only recognized in 1952. The existence of this disease confirms us in our belief that the immunological defenses are vitally important; but this does not mean that they are wonders of adaptation, as they are so often supposed to be. Our immunological defenses are also an important source of injury, even of mortal injury.

For example: vertebrate animals evolved into the possession of immunological defenses long before the coming of the mammals. Mammals are viviparous: the young are nourished for some time within the body of the mother: and this (in some ways) admirable device raised for the first

time in evolution the possibility that a mother might react immunologically upon her unborn children --- might treat them as foreign bodies or as foreign grafts. The haemolytic disease that occurs in about one new-born child in 150 is an error of judgment of just this kind: it is, in effect, an immunological repudiation by the mother of her unborn child. Thus the existence of immunological reactions has not been fully reconciled with viviparity; and this is a blunder --- the kind of blunder which, in human affairs, calls for a question in the House, or even a strongly worded letter to *The Times*.

But this is only a fraction of the tale of woe. Anaphylactic shock, allergy, and hypersensitivity are all aberrations or miscarriages of the immunological process. Some infectious diseases are dangerous to us not because the body fails to defend itself against them but --- paradoxically --- because it does defend itself: in a sense, the remedy *is* the disease. And within the past few years a new class of diseases has been identified, diseases which have it in common that the body can sometimes react upon its own constituents as if they were foreign to itself. Some diseases of the thyroid gland and some inflammatory diseases of nervous tissue belong to this category; rheumatoid arthritis, lupus erythematosus, and sclerodema may conceivably do so too. [They do. PM] I say nothing about the accidents that used to occur in blood transfusions, immunological accidents; nor about the barriers, immunological barriers, that prevent our grafting skin from one person to another, useful though it would so often be; for transfusion and grafting are artificial processes, and, as I said in an earlier lecture, natural evolution cannot be reproached for failing to foresee what human beings might get up to. All I am concerned to show is that natural devices and dispositions are highly fallible. The immunological defenses are dedicated to the proposition that anything foreign must be harmful; and this formula is ground out in a totally undiscriminating fashion with results that are sometimes irritating, sometimes harmful, and sometimes mortally harmful. It is far better to have immunological defenses than not to have them; but this does not mean that we are to marvel at them as evidences of a high and wise design.

We can, then, improve upon nature but the possibility of our doing so depends, very obviously, upon our continuing to explore into nature and to enlarge our knowledge and understanding of what is going on. If I were to argue the scientists' case, that case that exploration is a wise and sensible thing to do, I should try to convince you of it by particular reasoning and particular examples, each one of which could be discussed and weighed up; some, perhaps, to be found faulty. I should not say: Man is driven onwards by an exploratory instinct, and can only fulfil himself and his destiny by a ceaseless quest for Truth. As a matter of fact, animals do have what might be loosely called an inquisitiveness, an exploratory instinct; but even if it were highly developed and extremely powerful, it would still not be binding upon us. We should not be *driven* to explore.

Contrariwise, if someone were to plead the virtues of an intellectually pastoral existence, not merely quite but acquiescent, and with no more than a pensive regret for not understanding what could have been understood; then I believe I could listen to his arguments and, if they were good ones, might even be convinced. But if he were to say that this course of action or inaction was the life that was authorized by Nature; that this was the life Nature provided for and intended us to lead; then I should tell him that he had no proper conception of Nature. People who brandish naturalistic principles at us are usually up to mischief. Think only of what we have suffered from

a belief in the existence and overriding authority of a fighting instinct; from the doctrines of racial superiority and the metaphysics of blood and soil; from the belief that warfare between men or classes of man or nations represents a fulfilment of historical laws. These are all excuses of one kind or another, and pretty thin excuses. The inference we can draw from an analytical study of the differences between ourselves and other animals is surely this: that the bells which toll for mankind are ---- most of them, anyway --- like the bells on Alpine cattle; they are attached to our own necks, and it must be *our* fault if they do not make a cheerful and harmonious sound.

**EVOLUTIONARY PARADIGM FOR GLOBAL POLITICS**  
**INTERNATIONAL STUDIES QUARTERLY**  
**Vol. 40(3), September 1996, pp.321-342**

By  
**George Modelska**  
Department of Political Science  
University of Washington

**Abstract:** The evolutionary paradigm for global politics here presented consists of four key propositions: (1) The global political system is a population of policies or strategies; (2) Global politics constitutes a complex system that evolves in specifiable conditions; (3) Accounting for global political evolution is a four-phased learning process whose key operators are variation (innovation), cooperation, selection, and reinforcement; (4) Global politics coevolves with global economics, community, and opinion etc. The evolutionary paradigm sheds light on two processes in particular: the formation of institutions at the global level, and the rise and decline of world powers (the long cycle).

Two propositions are central to this paper: (1) the institutions of world politics evolve, that is they undergo change subject to identifiable evolutionary processes, and (2) the rise and decline of world powers (the long cycle) is a mechanism of global political evolution.

By institutions of world politics we mean constitutive and widely accepted arrangements in respect of war and peace, nation-states, alliances, and international organization, and global leadership and international law. If we consider these arrangements in a sufficiently long perspective, say over the span of the past millennium, we cannot but help noticing significant changes that have occurred in relation to these, that continue to affect them, and that therefore need to be understood and explained. We need a structural-historical theory of world politics. The rise and decline of world powers, that has been the lead story of the past few centuries of world politics, also needs to be understood in a wider framework. It is not the case of some eternal struggle for power but rather that of a mechanism that in the recent past has mediated major changes in world political and social organization. We need to see the long cycle not in isolation but as a feature of world institutional growth.

That is why, to better understand world politics in its time dimension in particular, we require an evolutionary framework. What might be the salient features of such a paradigm?

## **EVOLUTIONARY PARADIGMS**

### **Types of evolution**

What is a paradigm? A paradigm is an exemplary pattern that identifies the key questions and the fundamental variables; more generally it is a set of canons for the statement of problems of general significance that is espoused, or shared, by a research community.

What does the pattern exemplify? It represents a conception of the natural order of things, and specifies what, in a particular realm, is to be normally expected. Explanations then amount to showing why actual events diverge from 'normal', that is 'reasonable', expectations. A paradigm defines, for a class of events, what stands to reason.

An evolutionary paradigm is one such pattern. Kenneth Boulding (1981:9, 25), a social scientist, described it as "a pattern of the universe", "a pattern in space-time". Like Herbert Spencer before him /1/, Boulding saw this pattern characterizing not only the biological world, but also the physical universe, viewed on the very large scale of cosmology, and also the social world. If, following Spencer, culture is added to that list as a separate realm, also subject to these considerations, four types of evolution emerge: physical, biological, social, and cultural that can be ranked along a time axis according to the period of their evolutionary processes. Thus cosmology and geology operate on a time scale of billions of years; the story of life, and the animal kingdom on earth is reckoned over tens and hundreds of millions years. Social organization, for humans in particular, does not extend much over hundreds of thousand years, into some million, and culture has an even shorter time span. All of these major processes together make up evolution, but they all need to be kept distinct, too, if only because their periods vary.

That makes evolution a pervasive pattern of some considerable power and generality. It is indispensable in grasping long-range processes even though it is not claimed to be a pattern of everything and does not explain every class of event or process. But does that also mean that the study of evolution necessarily involves a hierarchy of sciences, and that biology, as the science of life, has an inherent or overriding claim on priority of insight in evolutionary theory? Not necessarily. In the evolution of evolutionary theory, all branches of knowledge (that are fields of culture) have participated in the past and continue doing so. In mid-nineteenth century, biologist Charles Darwin captured public attention with an account of the origin of the species, and his name came to be completely identified with the concept of evolution by natural selection. But his own insights and discoveries were profoundly shaped not only by earlier work in geology, such as that of Charles Lyell, whose principle of "uniformitarianism" (that past changes are to be accounted by processes still in operation) is basic to Darwin's work, but also by that formative work in early demography, Thomas Malthus' "Essay on the Principle of Population". What is more, sociologist Herbert Spencer developed his own concept of evolution even before Darwin's work had appeared in print.

That is why each type of evolution may also be regarded as occupying its own domain, and privileged time-space. The knowledge of each contributes to evolutionary theory, and researches mutually cross fertilize work on all others. While evolutionary conceptions have recently had particularly full elaboration in biological theory, no one should forget the role of geology and astronomy in pioneering the understanding of deep time. Claims to priority or one-way influence by any one field of evolutionary study need to be treated with caution.

## Social science evolutionary paradigms

The present study is therefore located in social evolution, and in the social sciences, and concerns global politics as one sub-type of social evolutionary process. What are the essentials of

an evolutionary paradigm for the social sciences? /2 / According to R.C Lewontin writing in the International Encyclopedia of the Social Sciences (1968:203): "There is a hierarchy of principles in the evolutionary world view: *change, order, direction, progress, and perfectibility*. Evolutionary theories are distinguished by how many of these are successively included as essential."

These five principles serve as a convenient framework for discussion. The argument will be that, for present purposes, the essential ingredients of the evolutionary paradigm for the social sciences are only two: change, and direction. Order, progress , and perfectibility are not essential parts of such a framework.

## Change

The most basic evolutionary considerations center around "change". An evolutionary perspective represents "a commitment to the instability of the present order as well as the past. In its simplest and irreducible form evolutionism is the doctrine that change of state is an unvarying characteristic of natural systems and human institutions and that such change follows immutable laws" (Lewontin, ibid.).

Change of state in societies means change in the economic, political, societal, and cultural structures that constitute them. Structural change is to be distinguished from routines that characterize all social life; (this distinction is central e.g. to the Nelson and Winter (1982) analysis of economic development). Structural change commonly represents innovation that is a departure from standard operating procedures, and that is why the story of social evolution is a record of innovation. But structural change, or the diffusion of innovation, also takes time, and that is why the observation of change in human institutions invariably requires a long perspective.

Such an approach is clearly structuralist, in that it proposes that the persisting clusters of social behavior that are subject to social evolution form emergent social structures whose properties cannot be deduced from the parts composing them, and that it focuses on transformations of these structures. It emphasizes change, rather than "evolutionary stages" that are often seen as the principal products of evolution analysis, and often reified, or even personified, as in "capitalism". It is not functionalist in that it does not inquire into the functions of persisting structures, but it does search for explanations of change in these structures.

Distinguishing "structural" from "routine" change helps to get over the problem that Lewontin raises, of separating "real" change from a stasis that has only the appearance of change. But to assert that evolution is structural change does not necessarily imply the statement that order is the natural outcome of evolutionary processes. A more modest proposition would stipulate that such processes are concerned with adaptation; that is, they might cope with a set of identifiable problems, in relation to which they may, or may not, be adaptive.

## Directionality

Lewontin's third principle (1968:204), direction, also is a basic one. "By direction in evolution we mean the concept that there is some natural linear order of states of the system and that an evolutionary process can be described as passing through successive states in that order" in a line that is always ascending or descending.

Is the scale on which such directionality can be measured differentiation, or complexity, as Herbert Spencer first proposed it? Over very long periods, we can observe both differentiation, and increase in complexity in social systems. In the evolution of the world system of the past millennium, a prominent instance of differentiation has been the formation of the global and national systems, where previously there was only regional and local organization. Arguably, too, the world system was less complex ca. 1000 than it is going to be in the year 2000. But is such differentiation, and higher complexity due to evolution, and if so how?

Directionality does not posit the existence of a design, or blueprint; "genes are much more like a recipe than like a blueprint" (Dawkins 1987:296), and it is not teleological, in the sense of implying search for evidence of such design in nature. It does not lead to "laws of history" proclaiming developmental sequences familiar in, and rejected by, the social sciences such as that of "feudalism ---> capitalism" leading to the "final goal" of socialism".

But directionality does imply that evolution is not random and that it is a cumulative process, whereby a succession of small changes can bring about great transformations (Dawkins 1987:Ch.3). A recipe is a set of ingredients, and a set of instructions. Instructions organize the process in time thus giving it a temporal structure; evolution might be thought of as involving some such instructions. Ingredients compose the conditions that induce evolution; they define the spatial aspect of that process.

Directionality can be made more tractable at shorter time frames if it understood to be the product of learning. J.W. Pringle (1951) has shown that learning viewed as increased complexity of behavior over time, may be thought of as equivalent to organic evolution, usually thought of as increased structural complexity (in space); hence social evolution is basically about learning new behavior (see also Campbell 1969, Schull 1991). Successful social learning produces structural change; a learning process, furthermore, is inherently phased, has a distinct time-structure, hence direction. It might therefore be argued that it is learning that gives directionality to evolutionary processes. It could be seen as a "natural" process of trial and error, as if the unwinding of an internal logic or program of adaptation, one that does not require the postulation of a grand design or purposeful intention, but does call for an explanation.

Such an approach is neither deterministic, nor does it assume randomness; it is probabilistic. /3/ It favors directionality without projecting for it a fixed content or a finite purpose, and does not require the teleological assumption of a final goal or destination. It requires adjustment to changing conditions of the world system, including environment, population, urbanization, technology, and wealth, but does not require the assumption of a grand design. All it says is that the system under study proceeds on its way in a certain manner, and according to codes, or programs, that need to be explicated. It obeys a set of rules in that it plays out in a given natural

and social environment that includes other processes and policies each carrying out their own programs. All such a model does is to postulate an "inner logic", that is the formal-logical requirement that the processes evince a time-structure that constrains them.

## **Progress and perfectibility**

There is a tradition of long standing (cf. review in Ginsberg 1961) that regards progress as an essential characteristic of evolution. Progress is not identical with evolution, but is linked to it: it is evolution in a direction that satisfies certain criteria of value. The prominent formulations of the idea of progress date from the era of the Enlightenment. The most famous among these formulations were those of Condorcet, and his criteria, that included equality, and peace, do not seem as utopian as they were once made out to be. But whether the study of social evolution can by itself provide the relevant criteria of value is a matter for debate. And even if the criteria were agreed, there is room for disagreement as to whether or to what degree the record of human history, be it of the 20th century, of the last millennium, or of the past 10,000 years, shows progress in the human condition. Some social evolutionists maintain that any such claim is misplaced and that no such progress has in fact occurred.

These are matters still open to discussion, both on empirical and on theoretical grounds. But they make it clear that it would be unwise to include progress, movement in a "good" direction, among the essential traits of social evolution. Biologists have had some difficulty specifying precisely what might be meant by biological progress. Suffice it to say that a "learning" conception of the directionality of evolution keeps open the possibility of progress, but leaves the determination of the precise characteristics of that progress to the analysis of cases. A "learning" conception of social evolution leaves room not only for "materialist" components (of wealth and power) but also for "idealist" elements (of truth, and love) that make for a well-rounded analysis. Lewontin's fifth principle, perfectibility, is even more stringent. This again, is a criterion of Condorcet (a mathematician), who saw it as a limit toward which the process might be moving, without ever attaining it. To-day evolution is more often viewed as an endless process with no ultimate goal or destination, and even a "learning" conception stresses in the first place adaptation to current problems rather than final purpose. But as Lewontin points out (1968:206) if there is directionality on some criterion, then perfectibility cannot be altogether ignored.

Indeed, for social evolution it raises the problem of a possible life cycle for the human species. Could it be programmed to die out at the end of such a cycle?

## **Macro- and micro-evolution**

This discussion leaves us with two main evolutionary principles: structural change, and Formal-logical directionality. It remains to point out that these are broadly equivalent to the two major divisions of biological theory: macro- and microevolution. These are very basic distinctions, even if the dividing line between macro and micro, description and explanation is not as sharp as it might be thought to be.

Biological theorists (such as Ayala 1982, Pollard 1984) commonly now divide evolutionary (or the synthetic) theory into two areas, macroevolution and microevolution. Macroevolution,

meaning the evolution of all living groups, considers the question whether evolution has occurred and by what pathways. It was called the theory of descent by Darwin (who defined evolution as descent with modification), and sometimes is called the fact of evolution, with a strong descriptive element. Darwin's "tree of life" is the most general graphic representation of the observed facts of evolutionary change.

Microevolution is the study of the mechanism of evolution. Darwin suggested that natural selection was a chief mechanism that explains the non-random aspects of evolution, and thus supplied a principal explanation for the observed variety of life forms, but we now think of it as one among such mechanisms. These mechanisms could be thought of as supplying the directionality of evolution.

The discussion of Lewontin's principles has now brought the subject down essentially to where biological theory also finds itself. But do we wish to draw, in our own analysis, a sharp distinction between micro- and macro-evolution, between fact and explanation? Probably not, because the boundaries between description and explanation are not really that sharp; good description and classification imply a good theory, and convincing explanations need to be tested against data collected on the basis of a theoretical scheme. What is more, macro-analysis requires reliable knowledge of micro-conditions, and vice versa.

How do such distinctions apply to the study of social evolution? Are conditions such that an "evolutionary analogy" is in fact justified?

### **Comte-Spencer v. Darwin**

Since the second half of the 19th century, two conceptions of evolutionary theory have existed side by side: the Comte-Spencerian, and the Darwinian.

Auguste Comte, and after him, Herbert Spencer, proposed that human evolution passed through major stages of social development. Also referred to as "evolutionism", this view emphasized major stages that might be manifested in the history of humanity, and could therefore be regarded as a form of "macro-evolutionary" analysis.

The Darwinian model elucidated a central causal mechanism of evolution to explain continuity and change in populations, but avoided the temptation for quick explanations of socio-historical processes. It centered on the analysis of selection, and for that reason has also been referred to as selectionism. In the social sciences, Darwinian selectionism is a form of "micro-evolutionary" analysis.

Over time, the Comte-Spencerian program fell into disuse, even if the problems it was intended to tackle, understanding large-scale change in human affairs, has not disappeared. But in mid-20th century, Darwinian theory experienced a strong revival and reinvigoration through a "modern synthesis" (Huxley 1942, 3rd ed. 1974) that followed the revolution in genetics and was followed by the discovery of DNA, that in turn has been subject to much critical analysis (e.g. Pollard 1984), and that also exerted much influence on the social sciences.

Our project here is to combine these two conceptions. Darwinian micro-mechanisms of search and selection, as adapted to the social context, that have now become an accepted part of social science (cf. Nelson and Winter 1982, or Elster's treatment of selection, 1989) are also strongly represented in our model, albeit in novel forms. It is the "macro-evolutionary" component of this project, that raises larger questions and deserves closer scrutiny, because it paints, in broad strokes, a "big picture" of global political evolution in the Comte-Spencerian manner, but with new concepts and in a way that also requires Darwinism for its validity.

All in all, this does not imply that evolutionary biology (as well as physics, and cultural theory) and the social sciences must have an identical evolutionary theory. The question is not: how much like biological systems are social systems, or how much social behavior is rooted in biology. Nor is that question part of the debates that have centered around Social Darwinism (cf. Campbell 1969), even though it is fair to assume, (without taking on a commitment to biological determinism), that social behavior has a significant biological component.

Instead the question is: given that the theoretical basis of biology is micro- and macro-evolution, in the sense that both change, and directionality are essential components of evolutionary theory, what additional useful analogies might there be for the social sciences, if the social system of the human species, *mutatis mutandis*, is viewed as subject to evolutionary processes. There are important differences between biological organisms (not to mention the physical universe, and culture) and societies, and they argue for keeping, at this stage, the several realms analytically distinct.

There is also the consideration that the structures that undergo change are not unitary but differentiated, in the case of the social sciences, at least into those of economics, politics, society, and culture. Accordingly, there is not only evolution of each of these structures to be considered in their own right, but also the relationship among these evolving structures, each proceeding at their own pace, but sensitive, and adaptive to, developments in each of the other processes. That is not a problem (how the evolution of one species affects that of others) that is apparently much or commonly studied in biology. Hence the problem of coevolution: how change in one set of social structures and its direction relates to the others. That is why coevolution may be added to macro- and microevolution.

## Rational choice v. evolutionary paradigm

Evolutionary paradigms for social sciences embody no claim to universal solution for every problem. They are fitted to deal with some important problems, but are not necessarily the prime remedy for many others. Ultimately, and in the long run, the various paradigms should be mutually compatible, being attempts to study the enormous elephant of social life from a number of different directions. But in the shorter run, such compatibility may not be immediately attainable.

In the social sciences to-day one important paradigm is that represented by rational choice theories (cf. Elster 1989). It has an excellent pedigree, and a fine methodology, contemporary neo-classical economics being one of its successful incarnations. How does it contrast with evolutionary theory?

Most basically, rational choice is the study of decision, that is, of actions to be taken in the light of given preferences, and constraints. Its time perspectives are those of the rational decision-maker, which might often be severely discounted by lack of knowledge or limited by short time horizons. Even more fundamentally, rational choice theories offer no purchase on social structure or structural change, and have difficulty with processes that have time and directionality as their essential characteristics. Neo-classical economists have not had much success in dealing with long-term economic development, structural unemployment, and turn-over in leading sectors. In political science, neo-realists have had some trouble coping with structural change at the global level.

In other words, rational choice theory might yield better insights in the analysis of individual decisions and policy choices, but evolutionary approaches should give superior results in the study of long-range social processes, and of structural change in particular. Evolutionary approaches do not require the postulate of rationality; they allow for the possibility of trial and error solutions by social selection. But they also supply the context within which policy decisions must be formulated, just as micro-choices provide a sound basis for the understanding of great movements.

Such an approach places both "ends" and "means" together at the center of analysis. Both values and ideas, and power, both idealist (agendas, free societies) and materialist (politico-strategic, and economic) components of social processes are equally implicated in this analysis. This is not just a case of trying to have it both ways, in the manner of eclecticism, but rather a deliberate strategy to cover the four dimensions of enduring social experience /5/.

## **What needs explaining?**

What problems in the study of world politics are particularly suited to an evolutionary approach? Global politics evolution might be defined as the theory accounting for the appearance of political organization at the global level, and the processes by which global political structures have acquired their present form. Therefore the short answer is: structural change at the global level which changes, via mechanisms to be determined, is the global political system, and the process is shown in Table 1 as a sub-set of (world) political evolution. /6/ Not nation-states, not countries, but the global political sector of the world system.

Moreover, if a distinction is drawn between institutions (the rules of the game), and organizations (the actors or players in that game, pursuing strategies) (as does e.g. North 1990), two kinds of change will be recognized: in major actors, and in basic institutions of global politics. A most prominent instance of change in actors over the most recent 500 years have been the rise and selection of successive world powers (the most recent being that of the United States) that has provided the basic pulse of global politics.

The other, longer-range process has been that of institutional change. A superficial glance at world politics suggests to some onlookers nothing but a chaos of perpetual 'coming and going' but in fact there is a pattern to it: the succession of world powers not only involved successful copings with a parade of global problems, but also powered basic institutional change. For in the past one thousand years, global politics has moved steadily on a path toward greater global

organization, beginning with failed attempts at world empire that overshadowed the earlier centuries of world system history, continuing with the rise of the nation-state system, and moving increasingly into forays toward world organization that are also likely to extend considerably into the future. In other words, in taking a long view of world politics, we can perceive not just changes in factors, but also changes in basic institutions.

The outline of that process portrayed shows a succession of leading powers, central to global politics of their time, and grouped into blocs of four, each representing periods with characteristic forms of global organization. The first period, Eurasian-centered, is that of the establishment of preconditions, and its defining feature, the failure of the Mongol design for world empire. The second period, West-European in spatial orientation, shows the laying down of a nucleus of an emerging global system in that area, with world-wide repercussions. The third period, opening about 1850, is the start of global political organization, seriously taking off after 1945. Each period marks a transition to new institutions, and new rules of political organization.

Students of this field therefore confront the fact that global politics is subject not just to routine processes, but also to substantial and continuous structural change at more than one level and the basic question becomes: what explains such change in the past millennium so that the process can be projected into the near future? For in that millennium (a) the form and substance of global organization has changed substantially (i.a.) from a condition of minimal or non-existent structure and low connectivity, to one of substantial structure in conditions of substantially higher connectivity to-day. Moreover, that development has been not merely one of change but has also shown (b) directionality (rather than randomness) in that the change has embodied search for innovative forms of organization (appropriate i.a. to an expanding population), and has also traced an orderly path in space, and exhibited a temporal structure.

Put differently, significant aspects of world politics have been about institutional innovation, and about mechanisms and agents of such innovation. Because global politics has been subject to an evolutionary process, an evolutionary paradigm is likely to afford the best answers to questions about that process, and justify some confidence for projections into the future. Such a paradigm accommodates and privileges (diachronic) studies across time, as in long range processes, but also requires (synchronic) at-one-point-in-time studies of conditions that favor and govern evolution.

All this represents the claim that theoretical (Darwinian) biology and the theoretical social sciences are "equivalent, albeit different examples of the use of one and the same general theoretical calculus (or model), the theoretical structure of which remains the same" (Schmid 1987:82). Arguably, therefore, this same model aids in understanding social processes and structures in general, but in the formulation here attempted, the basic propositions of evolutionary theory will be specified as applicable to global political structures. It is these evolving global structures, and not individual polities or societies, that are the subjects (or unit) of the evolutionary process.

## **EVOLUTIONARY GLOBAL POLITICS**

Let us propose here, in respect of global politics, a process model of evolutionary change. A process is a sequence or string of events, and we fit the global political process into the causal structure of the world that we take to be an evolutionary one.

It is a process model because it singles out for emphasis changes over time rather than static "stages of development" that such changes might be said to be bringing about. But allowances are also made for a multilevel analysis both of change of actors, and of change in basic institutions of world politics.

As applied to political events at the global level, the process model consists of four sets of basic expectations:

1. The global political system is a population of strategies.
2. Global politics is a complex system that evolves in specifiable conditions.
3. Accounting for global political evolution is a learning process of which the key operators are variation (innovation), cooperation, selection, and reinforcement.
4. Global politics coevolves with global economics, community and opinion (etc).

These are the basic propositions that comprise the present evolutionary paradigm; they might be called the "hard core", in Imre Lakatos' sense, of this evolutionary research program. Taken together, these four propositions offer a framework for the explanation of structural change in world politics.

### **1. Global political system as a population of strategies**

The starting point for evolutionary analysis is the global political system viewed as a set of policies (or strategies) for the (collective) management of global problems. These policies may (conceptually) be carried by a variety of actors or agents: world empires, city-states or nation-states exercising, or aspiring to, global leadership, alliances and coalitions, international regimes, and world organizations. But the emphasis at this point is not on actors (that afford the ingredients for policy) but on the policies themselves viewed as sets of instructions, or programs of global potential.

The instructions embodied in global policies provide the basis for the standard operating rules, or routines, of the global political system. Such routines reproduce themselves through processes of socialization and training. Variation and innovation in these routines is the material for global political evolution, and occurs as generations of policies succeed each other.

That which experiences global political evolution is the social organization of the human species. Which makes it clear that nation-states are not the basic units of world politics, even though certain nation-states may at times be carriers of global policies. That also means that policies of nation-states such as the United States, or Japan became principal units of evolutionary analysis only in so far as vehicles or indicators of global political change. Their evolutionary fitness, if any, is a component of the global process. Similar considerations apply to evolution viewed regionally, or locally, whose operation should be specified in relation to the global level. It

might be supposed that the working of evolutionary processes is most marked at the species level.

## 2. Complex systems evolve

Most basically this argument rests upon the conjecture that the global political system is a complex system, and therefore it evolves. The explanation of the evolution of global politics rests upon the global political system belonging to a larger class of phenomena, that of complex systems, all potential subjects of evolution. The argument has two aspects: global politics evolves because it is a complex system; it evolves when "necessary conditions" are best satisfied.

### Complexity

The directionality of evolutionary politics is that which Davies (1984:239-240) calls "organized complexity". The basic distinction is between complex systems, and others that may be either orderly, or chaotic. Ordered (or equilibrium) systems follow a fixed pattern and have no flexibility or capacity for change; chaotic systems are disordered and unpredictable. Complex systems stand at the "edge of chaos" but are not themselves chaotic; they have sufficient capacity for change to adapt to new conditions. That is, in the present case, it is the argument that world politics is neither an equilibrium (or near-equilibrium) system, as postulated in traditional "balance-of-power" accounts, nor is it anarchic, in the sense of being chaotic, but is in fact fluid, far from equilibrium, and flexible, one in which order arises through fluctuations /7/.

Complex organization of living organisms can be shown to arise spontaneously given the existence of an *ensemble* that is a large collection of similar systems. Complexity has been defined as the ability to make transitions, that is to evolve. According to Murray Gell-Mann, a "complex adaptive system" is a collection of simple parts that interact to form a complex whole capable of learning about, and reacting to, the outside world. (8)

In the present case, the relevant collection is the population of strategies or policies, past, present, and future. That way experiments will occur with alternative strategies until, in favourable conditions, innovation comes along that is selected out and then cumulates through amplification. The accumulation of countless innovations, large and small, leads to systems as intricate as modern market economies or free democratic communities.

To show directionality, or future-orientation, or "naturalness" there is no need to embrace determinism or assume "progress", as in "evolutionism"; but to postulate only that the evolutionary process unfolds in accordance with an inner logic and/or sequential structure, in that each phase creates the conditions for the next, always responding to new conditions in the environment. Such a process requires some capacity to anticipate the future but no greater motivation other than "search for a better life", or as Adam Smith put it, when accounting for what prompts humanity to save, the ever-present "desire for bettering our condition".

A complementary assumption is that of "sensitive dependence on initial conditions": that the beginning forms have an important effect on the course of development, in that they help cumulate the results of early changes, a basic reason why to examine carefully the time path of

structural change that we also think of as path-dependent. David (1988:18) describes processes whose outcomes are path-dependent as those dynamic processes in which the position and motion of the system, and its constituent sub-systems, are "sensitive to initial conditions". They are characterized by non-ergodicity (they do **not** pass through all the states compatible with its energy in the course of time) and irreversibility /9/. Both path-dependency, and future-orientation are ways of saying that the processes under study have a temporal structure.

## Conditions of evolution

If it is now established that complex systems are both path-dependent, and future-oriented, the questions then becomes: what are the optimum conditions required for the occurrence of evolution?

Even though there is no need to invoke the postulate of progress, there is no reason to believe that evolution is a random process, a matter of lucky accident, or "manna from heaven". Rather we suppose that in the presence of certain conditions political evolution (or political learning) will tend to occur. For Charles Darwin, these conditions included Malthusian population growth, and certain environmental conditions. According to Ervin Laszlo (1991:110) "evolution is not an accident but occurs necessarily whenever certain parametric requirements have been fulfilled". These are for complex systems: openness to energy flows, diversity in components, catalytic cycles, and feedback. A chief characteristic of complex systems is heterogeneity.

What might these requirements be for social systems? In respect of global politics it has been argued (Modelska 1987, 1995a) that the necessary and sufficient policy conditions for the rise, that is selection, of one nation-state to leadership in the global political system in the past half-millennium have been: politico-military organization of global reach; lead economy; a cooperative society; and openness and responsiveness to global problems. These are (local) conditions that made it possible for some nation-states to get ahead of others in the competition for global status in a heterogeneous world system. To the degree that long cycles have been the drivers of political evolution, these have also been the conditions of institutional development. Thus they also serve as guides to policy making, and to institution-building.

More broadly speaking they are also conditions that apply to the evolution of world institutions, and to global political evolution in the next century; the nearer the global system approximates, not uniformly but at least in some of its regions, such conditions, suitably generalized, as adequate political organization, matched by a world economic infrastructure of a maturing world market hospitable to the rise of new lead industrial sectors, in the context of an emerging global democratic community, and while responding to global problems, the more likely is political evolution to proceed at a smooth and measured pace. Even more generally these might be argued to be the conditions of evolutionary development in all dimensions of society.

Thus it is important to specify these conditions with some care. All in all there is no denying the most basic postulate of this approach that political evolution **necessarily occurs** in conditions of complexity. That also means that this is not an attempt to explain evolution as the generation of complexity, nor seek to measure its progress by that yardstick; rather to take social evolution to be the property of complex systems that, as occasions and conditions demand, may grow more or

less complex, but that flourishes best in certain specifiable conditions. It is the elucidation of conditions hospitable to evolution that is a first priority of the evolutionary theory.

## Questions

What expectations does this set up in respect to our field? Most basically, that global politics is subject to an evolutionary process: that it is subject to evolution, and capable of evolution. Global politics is not, and has never been since its inception, frozen into a stasis of a system populated by "preformed" states of unclear origin and uncertain future. Rather it is the constantly changing global political system of the human species with some parts of it evolving at faster rates than others. Within it, distinct policy lineages or lines of descent can be distinguished that carry the evolutionary process. An evolutionary tree of world political development might clarify such concepts; it would assert the common descent of populations of policies, and their common origin via a branching process (of differentiation). That is why we might view the evolution of the entire world system world as one process (Modelski 1995b).

Pre-evolutionary biologists regarded each individual species as created "preformed", that species were immutable productions, and had been separately created (Aristotle having originally posed the question whether the embryo was preformed, a miniature individual, or differentiated from an amorphous initial state). The preformation doctrine was widely held until early in the 19th century when epigenesis (development involving gradual differentiation of an initially undifferentiated entity) was finally demonstrated by Karl von Baer. It now "stands to reason" that students of world politics can no longer start with "preformed" states as basic units, but must fine-tune their analysis in light of the long-term variability of political institutions by refining the conceptual equipment that bundles change.

### 3. Evolutionary mechanisms

The third proposition proceeds to conceptualizing the working of this process over time; it asks how and why, what are the mechanisms, and replies, in terms that reach beyond Darwinian analysis, that the basic evolutionary operators are four: not only the Darwinian staples of selection and variation, but also cooperation, and amplification, that jointly constitutes a coherent set. They do not operate haphazardly but appear in sequences that extend over time and constitute learning processes. /10/

Selection is, of course, the classical Darwinian mechanism, so much so that some discussions such as Elster's (1989) convey the impression that it is the only evolutionary mechanism that counts. But Darwin himself gave equal billing to variation, being that which sets the entire process in motion, and Walter Bagehot's (1872) early social and political interpretation of Darwinism certainly gave it a prominent place. However, genetic mutation being also seen as an apparently random process, it often occupied a less conspicuous role, less subject to control than selection. To-day, innovation is generally seen as the source of variation in social evolution, and its independent importance is now getting to be widely appreciated.

Emphasis on selection and what some saw as necessary concomitants, "survival of the fittest", conflict and war, jarred many and led to arguments, one of the early ones being Peter Kropotkin's

MUTUAL AID: A FACTOR IN EVOLUTION (1914), placing cooperation at the center of the evolutionary scheme. Robert Axelrod's (1984) simulations showed cooperation not only to be possible among rational egoists, but also to be subject to evolutionary change. Studies of synergy (Corning 1983) have pointed in the same direction. More broadly we might regard "self-organization" as a fundamental attribute of all evolutionary processes, and regard it, as Stuart Kauffman (1995) has argued, as equal in importance to selection. We therefore regard cooperation as one of the necessary conditions of evolutionary fitness, and survival.

Finally, also implicit in Darwin's framework was the notion of selective advantage, or differential survival, the reinforcement and amplification experienced by "selected" programs, as maintained e.g. by Donald Campbell (1969:73) for "retention": "a mechanism for the preservation, duplication or propagation of the positively selected variety". "Operant conditioning ... is clearly an evolutionary mechanism ... a mechanism of reinforcement whereby a habit is selected 'within' a particular organism" (Van Parijs 1981:96). Reinforcement takes place where there is capacity to learn, and is essential to all learning, and so it is to social evolution.

By synthesizing these classic conceptions, the present account suggests not only that the social evolutionary process utilizes all four mechanisms and that all four are its necessary components. /11/ It also proposes that these four mechanisms are closely and meaningfully related to one another, and may be best understood if seen working in sequences, starting with variation, through cooperation and selection, to reinforcement and retention. But even though they are interdependent, their close relationship is often obscured by the fact that each of these mechanisms falls within the domain of a different social scientific discipline. Very broadly speaking, variation "belongs" to cultural studies, cooperation to several social sciences, including sociology, selection raises problems of politics, and reinforcement, of economics. "The attempted integration of these diverse mechanisms in the study of evolutionary processes represents an ambitious and risky synthesis" (Andersen 1994:14).

## Global political evolution

The basic units of analysis in respect of global political processes, as previously argued, are not nation-states, empires, or world organization, but persistent global political strategies (or policy routines). We regard forms of global organization, such as world empire, or global leadership, as carriers of tight clusters (or populations) of strategies that may or may not experience change.

Students of international politics are conditioned to think of states, and non-state actors, as the fundamental units of analysis. Such a view might serve as a first approximation, and often suffices as a shorthand expression. But on close analysis it cannot withstand scrutiny.

Organizations such as states use individuals, but they act through policies. The interplay of global policies constitutes the global political system. It is changes in these policies, changes that alter standard operating procedures that need to be subjected to the greatest scrutiny.

The essentials of an evolutionary learning process, or "calculus" /12/ might therefore be formulated as follows:

- a) The starting point is a population of global political strategies (or policies) that persist (that is, successfully reproduce themselves). "Persistence" (or reproduction) means the transmission of a

program, or code, or set of generating rules, to the next generation of strategies. Persistence of strategies need not be, in and of itself, problematic for an evolutionary theory, for it is accounted for by the basic inertia of all social systems.

- b) Over time, some if not most of these strategies will be reproduced in a routine fashion, by copying; but others will undergo change, e.g. by experiment or chance mutation, or will be proposed as innovations by policy entrepreneurs in response to demands for the solution of global problems. These are the sources of VARIATION that introduce innovation into the population of strategies.
- c) In complex evolving systems, innovations will engender cooperative, combinatorial or synergistic (cf. Corning 1983) effects. Strategies that become the focus of effective alliances have a better chance of surviving. Such COOPERATION, combinations and coalitions are more probable in free societies, and are not random.
- d) The political and social environment of this population of strategies (and not, in the first place, "nature", or the "natural environment") might then be regarded as comprising a selective factor or mechanism that helps to determine causally which parts of the program will persist, and which policies shall be substituted for by new programs. In global politics, SELECTION has been most directly the product of macro decision, and in the past, those of global war; more generally, for all political systems, elections are the selective mechanism *par excellence*. In global economics, there is the competitive environment of the world market; in global community formation, the contest of ideologies for the building of "model societies".
- e) This completes the process of revising the code, and all that remains is REINFORCEMENT, (that is reward, combined with punishment for non-selection) such that the result is a set of revised strategies that are then diffused, via mechanisms of amplification, and transmitted via a system of inheritance, in successive generations of policies.

One example of an evolutionary learning process exhibiting these characteristics is the long cycle of global politics (Modelski 1990, 1995a; Modelski and Thompson 1996). This (learning) cycle is a non-linear process that comprises four phases, each one of which manifests, and brings to the fore, the operation of one evolutionary mechanism. The first of these phases is "Agenda-change," that through "variation" brings policy innovation onto the global agenda. "Coalitioning" manifests the importance of cooperative action in global affairs. "Macrodecision" is the selection process of that system, and "Execution" is an opportunity for amplification, diffusion of innovations, and the building of memory.

But the turnover of lead powers is no more than a mechanism of world institutional change. Such change at the global political level may be reckoned in four long cycles, each cycle (driven by one global power, competing against others) representing a characteristic evolutionary phase of the global politics process, and four long cycles constituting a distinct period of global political evolution. The West European period of global politics comprised of the Portuguese, Dutch, and two British long cycles, might be thought of as having been formative both of the nucleus of the global system, and of the basic elements of the nation-state system, defined by rules of international law (forming what some call the Westphalian system). Each of these cycles might

also be regarded as having activated successive evolutionary mechanisms at the global level: the Portuguese, those of variation through discoveries, the Dutch those of cooperation in creating, with Britain, the nucleus of the global system, and the two British cycles first selecting the new international system, based on balance-of-power strategies, and then amplifying it through industrial and trade expansion. In other words, we observe (as in Table 3) the working of the same four evolutionary mechanisms both on the scale of one world power rising to global leadership, and of the institutional complex of the world polity in formation.

Hence we argue that evolutionary processes involve the same mechanisms in different settings and in different time frames.

That is, the process of changing policies, or institutions, is not homogeneous but passes through a number (four) of distinct and sequenced phases of a learning experience, each strongly linked to one of the evolutionary mechanisms. Such phases are (in one form or another) the property of all social evolutionary processes because they are all learning processes.

## **Periodicity**

A generation is a key temporal unit of evolution, because evolutionary processes are measured in terms of such generations, and generational turnover seems a basic source of periodicity. The measured generations in question are those of policies or strategies. But it might also be assumed that structural change is closely linked to generational turnover, or the coming and going of generations of organisms, and of humans. If a generation is reckoned as the replacement period, that is the interval during which a generation replaces itself, then in (human) social systems the interval is of the order of 25 to 30 years, and is the basic temporal unit of (social) evolutionary processes. Generations, moreover, are basic units of social learning processes. It is these phases, and generational sequences of phases, that constitute the most important material for the analysis of policy and institutional change.

The fact of regularity of the long cycle, and of related global processes, is now quite well documented. What are the explanations of that regularity? In and of itself, it may not be altogether surprising. According to Paul Davies (1984:241, 57) "periodic motion, or oscillation, is perhaps the most widespread example of order in physics"; indeed, "physical systems which display exponential behavior are also likely to display periodic 'sinusoidal' behavior".

The fact of periodicity or oscillation, or better still, a constant rate of evolutionary change, in turn, accounts for the synchrony that can be observed in co-evolution. Both global politics and the global economy can be called "oscillators" because, as shown, they execute periodic behavior. In physics, as in biology, it is now taken for granted that "coupled oscillators tend toward synchrony". As there are reasons to believe that the global political and economic systems are significantly coupled, they might also be expected to synchronize their behavior, and by extension the same argument applies to processes by which cooperation evolves at the global level, that is the processes of global community formation.

All this poses the question as to the mechanism of such regularity, and suggests the intriguing hypothesis of a social-evolutionary clock. A lead in this matter is provided by biologists who

recently found evidence of evolutionary molecular clocks (in Dobzhansky et al. 1977:308-313, Ayala 1984). Regular changes in molecular structure have been found to pace evolutionary processes over very long time spans (millions of years).

The assumption of a constant rate of evolutionary change (for a given process) is crucial to this analysis. But it is not confined only to molecular biology. That assumption was recently employed with good results in the massive study of the genetic history and geography of the human species, by Luca Cavalli-Sforza and his associates (1994:33). What is more, as they point out, studies of "genetic distance", used to infer when two populations shared a common ancestor, show a close correspondence with, and are supported by the results of, work on linguistic evolution. Indeed, linguists have used the same (though independently derived) logic of a "linguistic clock" (Jones 1993:111) to unravel the origins of the world's languages, the pioneering work being that of Morris Swadesh (1971) who based "glottochronology" (history of the differentiation of languages, on the "relatively constant rhythm of substitution" in a basic vocabulary.

A postulated social-evolutionary clock might be stochastic in character, governed by a constant probability of a certain amount of mutation; at the level of social processes, why not conceive of innovations that might cluster in particular time periods. Such bunchings have been noted in the literature on innovations, and have had so far no clear explanation. They might possibly be linked to the organization, in time, of social processes.

Such a clock could also be metronomic, that is timing such change. The determination of calendars has been basic to the emergence of civilization, and also highly politicized. For instance, the succession of dynasties, as in ancient Egypt, or in China, have been clear markers of world time. For centuries, history was written as political history, and as a story of political regimes. The more recent trends in historiography that emphasize social and economic trends, enrich our understanding of the past, but do not negate the "time-keeper" or "time-setter" effect of political processes. Possibly some evolutionary processes, such as the long cycle, might serve to time others.

## Nesting

Basic to this analysis is the insight that evolutionary processes and mechanisms operate simultaneously, though at different speeds, at more than one level. As previously argued (also Modelska 1987, 1995a), in respect of global politics two such levels might be distinguished: one is the actor level, at which, in each long cycle, a new actor has been selected for global leadership, or its equivalent. The other is the institutional level, where a global polity process operates, each period of which represents the cooperative search and amplified selection of a revised institutional framework, the adoption of new rules, and the reordering of the constraints defining the system.

That is why the third long cycle shows three periods of the global polity, those of Eurasian Transition, West European, and Post-West European, each of which represents the search for a new set of basic rules, and an evolutionary process of a restructured kind. The nesting of these four-fold evolutionary mechanisms and in particular how each long cycle, of the rise of one

world power, is composed of four phases, and how four long cycles in turn add up to one period of institutional evolution at the global level. At a yet more inclusive level, these three periods of the Global Polity might each be seen as exhibiting the working of the same evolutionary mechanisms at the world system level as eras of world order.

These two global levels are, of course, analytically distinct /13/, but they are also substantially related in a "nesting" fashion on a basis of self-similarity /14/. That is, these processes are structurally similar, but differ in scale and duration. What is more, the greater (institutional) process might be seen to enfold the smaller (long cycles), while on the other hand, the smaller, while nesting within the larger process, might also be seen to drive it.

It is the specification of the calendar of events (or sequences of events) at the two levels of the global process that is the other top priority of evolutionary theory.

## Questions

What additional questions and expectations does this discussion place on the agenda of scholarship? First, that world politics is subject to learning processes of determinate structure that are steadily changing it. Second, it tells us that time is a cardinal dimension of that field, and that no event or policy may be considered in ignorance of that factor because events and policies have duration, and sequential order, and call for coordination. Attention to such "temporality" highlights the constancy and ubiquity of evolutionary change, and the importance of synchronization.

### 4. The co-evolution of global structures

We have shown so far that global political processes can be studied, in the first instance as endogenous, but it is also clear that the conditions that favour political evolution in turn depend on other evolutionary processes that are exogenous to it. In so far as leadership in global politics depends in part on economic leadership, then the lead condition of a candidate economy is a function of its ability to produce global leading sectors (in originating a K-wave); in turn, leading sector expansion nests in yet other exogenous processes (that is, in the same example, in the evolution of the entire world economy, and of the world system). Concretely, British political leadership ca.1700-1850 was the product of the leading condition of the British economy, etc. The picture is complex indeed.

"Co-evolution" is a term referring to "diachronic changes in two or more interacting objects or systems" and Lumsden and Wilson (1981:367) have extended it to include the reciprocal effects of genetic and cultural evolution. In populations of policies we might speak not only of co-evolution of strategies in global politics and economics, but also of policy lineages. The relationship of coevolution also partakes the character of nesting based on self-similarity. We have previously determined a set of conditions that are "necessary" for selection to global leadership, and that will determine the shape of global organization. As just noted, in the case of the politico-strategic structure the relevant process is, of course, endogenous; that organization rises as part of the long cycle. But for the three other conditions recourse needs to be had to a set of conditions that while evolutionary too, are exogenous to the global political process. These

might also be called "interaction effects" because students of international political economy habitually lay much stress on the interaction of politico-military with economic factors. In respect of the lead economy, as just mentioned, we need to consult developments in the global economic system, and inquire into the conditions that are likely to foster new global economic sectors in particular. The "co-evolution of global politics, economics, community, and opinion, shown in Table 2 is a schematic (descriptive) representation of four processes. The third column, previously noted, shows long cycles that mark the rise (and implicitly also the decline) of world powers, and their antecedents in the early modern period. The fourth column shows K-waves that are coordinate with long cycles and that chart the rise and decline of leading sectors of the global economy; these sectors have been both industrial, and mercantile, representing innovative spurts (or Schumpeterian shocks) in economic and commercial organization, centered on the lead economy. Just as the long cycle might be viewed as the mechanism propelling the global political process, K-waves can be seen as the moving element of the global economy; jointly they activate the international political economy (full documentation in Modelska and Thompson 1996).

In respect of social organization, there is need to look into the rise and decline of "model societies", the shifting fortunes of social movements, and the prospects for a global democratic community; the second column of Table 2 shows the "model societies" that (each accounting for one period of that process) have been successively prominent in cumulating for the formation of a community at the global level, and have thus constituted the outline of what might be called the democratic lineage. Each entry also reports the "opposition", social forces in resistance to which emerged the successive members of the democratic lineage, viewed initially as experiments. This process is in effect one of global community formation, community being defined as a framework or pattern of enduring cooperation. /15/ It is community formation because the global community that is yet to emerge, may be expected to do so around a nucleus formed by the experimental societies of recent experience, among whom a democratic lineage might be traced; equally important this community formation is also coordinate with the political and economic processes just discussed. At the most general level, the evolution of the entire global system appears to be paced (as shown in the first column) by changes in the conceptual and intellectual currents that shape the media, global opinion and education.

Table 2 therefore depicts global politics of the past millennium as a co-evolutionary process activated by successive world powers, interacting with parallel developments in economy and community, and also displaying fundamental change at the institutional level. The elements of a global system arose in Sung China after the 10th century, and generated the Mongols' attempted world empire in the Eurasian context. As that attempt failed, the process then moved, from small beginnings of trial, error, and experiment, through the consolidation of a global nucleus in Western Europe, toward a condition of greater world organization in its current post-West-European period. In turn, that process is shown to be supported, at different but symmetric time scales, by related movements in the global economy, and in the structure of the global community. In these respects it also suggests explanations for the collapse of Communism and of the Soviet bloc in Eastern Europe and in Russia in 1989-91.

So much for co-evolution of global structural processes. For an even more complete picture, reference must also be made to developments at regional and national levels. All in all, a

complex task, that opens broad and ever widening vistas for evolutionary theory. But it might be rendered more tractable because the conditions hospitable to evolution (specified in section 2 above) that define and characterise complex systems may be regarded as the initial proxies for these more wide-ranging ramifications.

## Questions

How does this discussion change expectations about the normal course of global politics? It makes it clear that structural change at the global level is not confined to global politics but needs to be thought of as in close coordination with global economics, community, and opinion, all of them hanging, albeit at different rates, and with different "ingredients". No longer is it possible to subsume such efforts under the umbrella of "international political economy" alone. Transition crises in the several processes synchronize, and need to be tackled comprehensively. Questions also arise that reach beyond global politics. This entire discussion implicitly proposes, that in its basics and *mutatis mutandis*), that framework might also hold for global evolutionary processes in the realm of the economy, society, and opinion. Not in respect of all the problems that might be encountered, but specifically in respect of those of long-range import and transformational character.

Hence certain general statements about global politics might also be relevant for other social sciences and seen as basic: that for purposes of evolutionary inquiry, the basic unit of analysis is the human species in its political interactions; that global politics, as well as global economy, immunity, and opinion all evolve, and flourish in certain specifiable conditions, those involving variation, selection, cooperation, and amplification; that the mechanisms of that evolution are (four-phased) learning processes of determinate periodicity; and that at the global level these processes systematically co-evolve through coaction and synchronization.

## A research program

So much for the "hard core" (in the Lakatosian sense) of the social evolutionary research program in respect of global politics. It is actuated by the conception of the world system as subject to evolutionary process, in particular in regard to world politics, and one that is driven by the long cycle understood as a (phased and therefore also periodic) selection process. This is the heart of the "evolutionary analogy" (see also Pringle 1951). "Analogies are, of course, not sources of proof, but sources of hypotheses" (Campbell 1969:73).

The "hard core" of the program may not be directly testable, but predictions generated from within it may be, and have given rise to a sustained research program. It began with a demonstration of the existence of long cycles as historical-structural regularities in world politics (Thompson 1988). The theory predicted that at the close of each of the past five global wars one power would emerge with a monopoly of sea power, a key element in the politico-strategic organization for global reach that is a necessary condition of global leadership. Research and measurement has confirmed that prediction (Modelski and Thompson 1988). Other research and documentary work confirmed that prediction with qualitative material, and yet other work related to developments at the regional level.

The next stage was the search for explanation, and the advancement of the thesis that long cycles were an instance of an evolutionary process (Modelska 1990, 1995a). Such work has shown that the rise of world powers may be understood as a phased learning process, as predicted by the "evolutionary analogy", and flourishing best in conditions that may be interpreted as those favouring evolution.

A new stage was entered with the demonstration that the global political process (driven by long cycles) is synchronous with the evolution of the global economy (Modelska and Thompson 1996). There are grounds for thinking that co-evolution of a similar kind can be shown with the global community process (as proposed in the second column of Table 2), and evolution of the global system as a whole (in the first column). Another challenge for the evolutionary paradigm is the construction of a calendar of global politics for the 21st century.

# **AN EVOLUTIONARY THEORY OF CULTURE?**

## **A Commentary on Rose's paper: Controversies in Meme Theory**

**By**  
**George Modelska**

Nick Rose asks the question: if we can intentionally design memes, why do we need an evolutionary theory of culture?

Let us discuss this problem with the help of an example. As a political scientist with a special interest in world politics, I think of International Law as the current stock of norms (that are memes) regulating the behavior of states. That is, International Law is the meme-type carrying instructions as to how e.g. an embassy is to be established, and run, or how an international organization is to be founded, or operated. International Law can be thought of as a subset of world culture (the world stock of memes).

If the norms comprising International Law are subject to a force of inertia they might also be seen as self-replicating: in as much as states generally tend to go on behaving in a predictable fashion, in the way established by customary state practice and relevant precedents. Law practice and law schools will transmit the meme-type to a new generation of state operatives. It is useful to think of forms in which International Law is codified, such as treaty conventions, or authoritative manuals or textbooks, as forming the "material" meme-type.

The first premise therefore proposes that, in a 'natural' but 'social' process, the meme-type will tend to replicate itself. But then how does International Law change, adopting new norms and dropping others? The short answer is: by processes of social, including political, selection.

At times, the process can be a massive one. At the end of World War II, an entire new set of rules establishing norms in fields as diverse as the maintenance of peace, regional institutions, the competence of international organizations, trade, finance and investment, and even health and civil aviation, were adopted by the international community. Viewed overall, this selection was determined by the outcome of the global war that had just come to a conclusive ending. Viewed as a macro decision, the global war was the primary selection mechanism for a new set of international norms, and it was the winning coalition of 1945 that jointly selected this new set of norms. Had the opposing coalition won, a different meme-type, meme-type B, would surely have been put in place.

At more 'normal' times, the selection process might focus on one meme at a time. Most recently, the prohibition on the use of land mines has been under discussion world-wide. An international conference that convened to discuss this subject decided to propose a treaty banning such weapons, and the treaty might soon come into effect. The movement that launched this new rule of International Law was surely innovative, and must have involved deliberate foresight, but the process by which such potential innovation was being transformed into international practice was one of extended social selection, involving information campaigns, social organization and coalition-building, conference bargaining and negotiation, and has and will be followed by

voting in national assemblies, and ultimately, in national elections (an election campaign is the paradigmatic social-political selection process).

More generally, proposed social innovations (sooner or later) take the form of memes that are launched on the sea of society so that they either replicate, associate with fellow memes, and brave the challenges of a turbulent environment, or else they sink to the bottom of social concerns and are forgotten. That is, they are either socially selected in, or selected out. The fact of origin of such memes in individual or collective experience does not preclude the operation of social selection processes that ultimately add to, or subtract from, the world stock of memes.

In other words, I see no inherent contradiction between the conscious adoption of memes, and the need for an evolutionary theory that treats the social selection processes by which memes come to establish themselves in society, thus becoming part of culture. That also means that an evolutionary theory of culture has its centre of gravity less in individual consciousness and more in events leading to social-structural reorganization.

# **AN ALTERNATIVE APPROACH TO GENDER AND CONSUMER BEHAVIOR: MEMETICS**

**By  
TJ Olney**

Western Washington University  
Gender and Consumer Behavior Conference  
June 25, 1998, San Francisco

## **Abstract**

Memetics, a relatively new discipline, provides an exciting tool for the study of social phenomena. It combines the cognitive sciences with our understanding of the fields of evolutionary biology and epidemiology. The memetic approach to a subject area allows a non-traditional view as to how that subject has evolved and continues to evolve. Gender provides a marvelous application of the memetic concepts as it is a phenomenon clearly originally rooted in Darwinian survival issues, but one which has just as clearly socially outgrown prehistoric imperatives.

This paper is a Mind Virus that will attempt to infect the recipient with the meta-meme of memes and additionally to suggest how this conceptual framework can provide insight into problems of gender and consumer behavior.

## **What is Memetics?**

"Personally you did not create even the smallest microscopic fragment of the materials out of which your opinion is made; and personally you cannot claim even the slender merit of putting the borrowed materials together. That was done automatically -- by your mental machinery, in strict accordance with the law of that machinery's construction. And you not only did not make that machinery yourself, but you have not even any command over it." -- Samuel Clemens, 1917

## **Replicating Information Patterns**

Memes live in minds as replicating patterns of information. They are thoughts that act as if they have a life of their own. Essential to understanding how memetics differs from other ways of studying human thought and behavior is the notion that it makes sense to say that a meme has its host. The word meme was coined by Richard Dawkins (1976) in his book called "The Selfish Gene". Memes are analogs to genes but of the mental and cultural world rather than the biological world. The study of memes, their mutations, replications, and extinctions closely parallels the study of evolutionary genetics. Thus, it makes sense to talk about successful memes as those that spread to the most minds with the greatest fidelity. The spread of memes within an existing population follows the patterns of epidemiology.

Just as genes can be bundled into packages that infect biological hosts together as viruses or bacteria, memes can bundle into packages known as mind viruses. Mind viruses are composed of mutually dependent memes, analogous to symbiotic relationships in biological settings. Cults provide outstanding and easy to appreciate examples of mind viruses, spreading by proselytizing from cult member to new recruit. Harder to acknowledge and apprehend are the mind virus configurations of enormous complexity that represent such cultural institutions as consumer capitalism, governments, universities, or any organized religion.

If you, in reading this paper, or having listened to a presentation, further investigate the meme of memes and pass it along to someone else, then the meta-meme of memetics will have successfully replicated one more time. Not all memes are as complex as the meta-meme of memes, proverbs and maxims provide a good counter example.

Another characteristic of memes is that they are typically capable of causing behavior changes in their hosts, even if the only behavior changed is to pass along the meme to someone else. So, a joke would be a meme that gets passed along, doing little to change the behavior of its various hosts except that they re-tell the joke. Other memes can have much more profound effects on behavior. For example, a woman holding the meme that "a woman should marry a man who can support her" might choose to marry someone she really did not like because of that belief.

One counter-intuitive result of memetic theory is that a successful meme need not be beneficial to its host. It only needs to be configured such that its host will pass it on. The highly successful memes of astrology and astrological prediction provide an interesting case where there is arguably no tangible benefit to the host, and arguably some harm. Some memes are so virulent that they cause their own self-destruction within the span of a human lifetime. The meme complex of Nazi Germany provides an interesting example (Lynch 1997)

While host benefits are not necessary, most long term successful memes confer some benefits to their hosts, either through enabling or encouraging greater genetic reproduction and subsequent transmission of the memes to offspring, or through preventing the death or destruction of the host by helping them to behave in a life assuring manner. Agricultural techniques provide a simple example of the former case while the "learn to swim" meme might provide an example of the latter. For a meme to have long term success, replicating in many hosts, it cannot have severe negative consequences to its host. For example, the memes that defined Shakerism have essentially died out. Central to the set of Shaker memes was celibacy and no children. Consequently, over time, the rate of adoption of the meme was slower than the deaths of the hosts.

## **Mutations of the patterns**

Like genes, memes do not always replicate with perfect fidelity. Hence, any meme is subject to change as it passes from one individual to another. The joke that changes subtly from telling to retelling exemplifies this phenomenon. The mutation may be accidental as when the quality of the transmission or storage of the original meme fails to be adequate. Or, memes can be intentionally mutated in that are consciously designed by self interested parties. Such memes are often subtle mutations of original memes that had other social agendas in mind. For example, the

cooption by the fashion industry of counter-cultural symbols usually involves some mutation of the original fashion to make them acceptable to a wider audience. The grunge phenomenon provides an ideal example, as the only thing required of the fashion industry was to make the clothes new, and put high prices on them. The pricing alone served to sanitize the dumpster diver look into high fashion.

## **Categories of Memes**

Memetic theory has applied some of the standard memes of positivism to the meme of memes.

One such applied meme categorizes memes into types based on their typical functions in the realities of the hosts.

### **Distinction Memes**

These are the memes that we use to divide up reality into the organization that we know. This part of memetic theory is a straightforward outgrowth of years of categorization literature, particularly illuminated by differences in categorization across cultures. The meme of dividing things into categories is itself a distinction meme.

### **Association Memes**

These are the linking ideas that we hold to say that for instance, "After work is Miller Time." They express the relationships between two or more memes. Much copy writing attempts to engineer association memes between existing memes that people already hold and a new meme that causes the person to purchase a product or service.

### **Strategy Memes**

Strategy memes tell us how to behave in a given context to achieve some result. For instance, the strategy meme that says "Go with the flow of traffic, and you won't get stopped for speeding." might inform our driving behavior. Or, "never go to bed with a man on a first date" might inform our sexual behavior.

### **Resistance memes**

Resistance memes are thought complexes that reduce the probability that a host will acquire new memes that could threaten the memes already held by the host. Because of their key importance in meme transmission, these will be discussed in greater detail below.

Meme transmission and replication: the spreading of memes.

Memes spread by communication. That's the short answer. But the communication of memes can take many forms.

### **Observation**

When we see a new thing, or behavior, for example a new fashion, the meme that entails that fashion has been communicated to us. If we adopt the new fashion, or tell others about it, then the meme has been replicated successfully.

## **Word of Mouth**

When someone tells us that something is so, they have transmitted a meme to us that we may or may not accept. To the extent that the meme gets past our memetic defenses, we will become hosts to the meme.

## **Parental Transmission**

Parents dominate the meme acquisition of their children. The children's minds start out devoid of memes, and fill first with the memes of the caretakers closest to them. (Nanny users beware.) For many cultural memes, things that appear only vaguely or not at all, but rather seem to form the texture of reality, parental transmission provides the route. Empirically, it can be demonstrated that religions that promote many children grow faster than those that are silent on the subject, parent to child being the dominant transmission mode for religion.

## **Outside the family**

Peers, bosses, teachers, caretakers, and authority figures with whom people make direct contact also provide a ready avenue for meme transmission by word of mouth. In this area, there already exists a substantial literature on best communication methods for intentional transmission, but not much is said about the characteristics of the messages that will tend to be successfully copied and passed along to others.

## **Mass media (include books)**

The rise of mass media, along with increasing population densities, has enabled extremely rapid transmission of new memes and variants of older memes. In particular, radio has been singled out as a particularly effective medium for transmission of political memes because the radio audience often consists of people listening together at work. (Lynch 1997)

## **Evangelism**

A meme can be considered to be evangelic if it includes *as part of itself*, the idea that it is important to tell other people. Hence, when the meme of an impending hurricane reaches a certain place, all residents will be likely to evangelically spread the word to any who might have been out of the normal information loop. When someone enjoys a novel kind of sexual behavior, and attempts to convince others to participate in that behavior with them, evangelism is at work. Similarly, people who consume some brands (e.g. MacIntosh) will often evangelically attempt to convince others to consume the same brand. Crisis memes, religious memes, and commercial memes exemplify the class of memes that spread evangelically.

## **Crisis**

Depending on the severity of the crisis, people who accept the crisis's meme are likely to feel more or less inclined actively and effortfully spread the meme. Contrast the impending hurricane to the crisis of ecological disaster following years after the harvesting of a rainforest.

## **Religious**

These are the memes we generally think of as evangelic, hence the term. The admonition to go out and save others with this meme or mind virus is part and parcel to the message.

## **Commercial**

To the extent that a commercial establishment can successfully generate either positive word of mouth about its offering, or negative word of mouth (about the competition or failure to adopt the offering), that commercial interest has designed an evangelic meme.

Issues in meme spreading

## **Rate of Spread**

As with germs like viruses and bacteria, memes spread by communication from one host to another. If the meme contains a sub-part that says that everyone should be sure to pass it along, it acts the same as the sneeze does for the common cold virus. Density of population also affects spread rates, as well as type and availability of mass media.

## **Intention**

An aspect of memetic transmission just beginning to be explored surrounds the question of purpose or intention of the host. This perhaps provides the most direct link to marketing, if not consumer behavior, as producers of products use engineered memes in the form of advertising to attempt to consciously increase the adoption rate of the meme represented in the product or service being advertised. In a very real sense, the offering of a company, that includes its advertising is a set of symbiotic memes that have been intentionally designed to spread. The skill of the designer will affect the transmission rate, the repeat rate and the meme adoption rate. These will of course be affected by the size of the promotional budget and the effectiveness of public relations team in securing additional exposure for the memes.

## **Indirect meme transmission**

There are times and kinds of memes that pass from host to host via rather indirect means, either lying latent in the minds of their hosts, or lying fallow as a hidden text. One such case, that of generation skipping, provides food for speculation about how other such memes might propagate. The example given in several sources cites a tendency for lenient child rearing practices in one generation to be followed by strict child rearing practices in the next generation and a subsequent reversion to lenient practices in the third.

Fallow ideas in forgotten texts, such as the genetic ideas of Gregor Mendel which were ignored until many years after his death, or religious writings of the Druids, can be "rediscovered" and lead memetically rich new lives.

## **Susceptibility of host**

The acceptance of new memes by any one individual depends to a large degree on their susceptibility to the new meme. Susceptibility depends not only on the nature of the meme, but also on the state of the individual in terms of currently held memes. Just as people's biological immune system provides mechanisms for resisting many infectious agents, a mind develops a memetic immune system that protects it from invasion by "foreign" memes. Hence, most mature minds resist colonization by most new memes, but some memes have characteristics that make them more virulent and better able to take residence in any mind, regardless of immunity.

### **Preexisting Meme set**

The degree of compatibility with an existing set of memes strongly effects whether or not we adopt them. This phenomenon has much empirical support from the persuasion literature.

### **Resistance memes**

There are two memes, faith and skepticism, that most of us acquire at some time, that by their nature, act as a sort of memetic immune system for the rest of the memes that we hold against any potentially disruptive memes that we encounter. A third category of resistance meme, the taboo, appears much more explicitly in some cultures than in our own, but taboos do exist at many levels of analysis, from individual, through family and social group, to general cultural taboos. Adversative memes form a fourth type of resistance meme, using intimidation as a primary means of suppressing competing memes.

### **Faith**

Faith represents any meme that you believe without question. Because you accept it as reality, it is very resistant to attack or modification by existing memes.

### **Skepticism**

Skepticism forms a shield against new memes that can be just as potent as faith. With skepticism built into our meme set, we take nothing at face value, but seek further verification of new memes and their potential value. Cleverly composed memes, however, can and do slide right past our skeptical barriers. The scientific method forms perhaps the penultimate skeptical meme set, providing a strong shield against adoption of new memes without adequate evidence, much to the dismay of aspiring doctoral candidates.

### **Taboo**

Any subject that generates a response like, "We really don't like to talk about that." represents at least a mild taboo. The existence of taboo memes provides a passive form of resistance whereby not talking about something makes it harder to pass along memes related to the taboo, hence the first memes to colonize a mind, usually transmitted from parents, have a survival advantage. By forbidding discussion of certain topics, each taboo acts as an antibody against a small set of memes. Often, taboos find very little positive expression, but rather are propagated by negative sanctions surrounding discussion of the taboo subject. While the taboo does not directly counter the memes that it is set against by replacement or competition, the taboo generally slows the spread of the circumscribed memes. Taboo, however, may have interesting side effects, planned or desired by the agencies setting them up as taboo. For example, a taboo against talking about sex and sexuality, including birth control, tends to be associated with a higher incidence of unplanned pregnancies. (Lynch, 1997)

### **Adversarial/Intimidation**

The last form of resistance meme, the adversative meme, sets up an automatic adversarial position in its own host against competing memes. In doing this the host of competing memes becomes intimidated by the holder of the adversarial meme, thereby suppressing any communication of the competing meme to the holder of the adversative meme, and in extreme cases, suppressing any communication of the competing meme to any other potential hosts. Most memes that advocate the incarceration of a particular group use this adversative form to gain

reproductive advantage. The advantage comes not from better replication of the adversative meme, but from the suppression of the replication of competing memes through mechanisms of intimidation. Usually, the espouser of memes competing with the adversative meme finds themselves in the uncomfortable position of being accused of being exactly what the adversative meme is trying to eradicate, be it communists, feminists, gays, lesbians, or intellectuals.

### **Hot Buttons**

One controversial assertion made by some authors writing on memetics contends that we have certain hard wired and early wired kinds of memes that have at their roots survival of the species and of individuals within social group. (Brodie, 1996)

#### **Primary Hot Buttons: the four F's**

Any meme will command at least a modicum of attention if it somehow deals with one of biology's four f's, food, fighting, fleeing, and ... finding a mate. As a species we have learned, perhaps have it genetically/chemically encoded that these are important things to attend. They all involve states of emotional arousal related to hunger, anger, fear, and lust. Because of this, memes that use one or more of the four f's as integral parts of themselves will have a greater chance of being received by a new host.

#### **Secondary Hot Buttons**

In addition to the four f's, memes that tap into the realms of belonging, identity, approval by others – including those in authority, and caring tend to transmit easily and replicate faster. This in particular forms a rich area for research that links the persuasion literature with the memetic way of viewing problems.

#### **Physical and emotional state**

The classic case of enhancing susceptibility to new memes by manipulation of emotional states comes in cult and military brainwashing, where by means of deprivation, and sometimes pleasure/pain techniques, individuals are "programmed" with the desired new meme. To a lesser degree, anyone can be caught with their guard down, and find themselves agreeing to memes and even replicating them without having so intended.

#### **Value of a Memetic Approach to Gender and Consumer Behavior**

Aside from being an exciting way to think about thinking and behavior, memetics provides some additional, more rational incentives to its use as a conceptual framework for dealing with gender issues..

#### **Breaking out of our box of using SEX not gender**

Perhaps the single most important contribution that memetics offers the study of gender is a way to break out of our box of thinking about gender in terms of sex. With a few notable exceptions (Fischer and Arnold, 1994) most research directly involved with consumer behavior and gender measures sex and not gender. By examining the various memes of gender, and how they are interrelated, we gain a perspective that truly examines gender as a ***social phenomenon first***.

#### **Understanding how gender ideas spread**

Memetic approaches must necessarily include the ***arbitrariness*** of a particular meme, and look

not only at its value to individuals, but also at its *fitness as a meme*. Does it resist mutation? Is it passed along to offspring? Is it passed along evangelically? Memetic approaches can begin to answer questions about what makes some ideas anathema to some, while being quite acceptable to others.

### **Explaining how gender ideas change.**

It is clear that people hold many differing ideas of gender, of gender roles, and hold differing attitudes toward gendered subjects. How those ideas came to hold sway in one individual, how certain groups have adopted one set of ideas over another, or how one group comes over time to split off from its parent group are all questions amenable to memetic analysis. When a random change in a genetic sequence occurs that enhances the survivability of the whole gene sequence, we say that there is a positive mutation. Similarly, when a random (either accidental or intentional) change in a set of memes occurs which causes either faster spread or better resistance to competing memes, we have a memetic mutation. Not all memetic mutations will increase the survival of the base meme set. However, those that do will come to dominate the original meme set. The precise mechanism for mutation of memes can be intentional, motivated by biological or economic imperatives, or it might be accidental, caused by a misunderstanding or faulty replication of the parent gene.

A possible example of such a faulty transmission of gender role memes would be the individual who has at once embraced "traditional" gender roles that dominated during the 1950's and also embraced newer feminist ideas of equality and interchangeability of gender roles. The former memes were transmitted to them by parents, while the latter memes have been spread by evangelism through friends and teachers. If the person is a woman, for example, she might have taken on the masculine role of providing income to support a household and found a stay at home man to care for their children. The man, to participate in the arrangement, will have adopted similar memes. Now, the reality that the woman still feels compelled to do the housework and to spend a great deal of her non-work time caring for the children and the man's feelings of inadequacy for not being a breadwinner, represent a faulty transmission of the feminist equality memes to both parties. (c.f. Mahoney 1996)

### **Explaining resistance to change cybernetically**

Feedback loops between existing gender notions, commercial interests, family interests, and gender memes that people choose to adopt or pass along *reinforce* existing memes to the disadvantage of new gender memes. Advertising in mainstream "women's" magazines provides the perfect example of this reinforcement phenomenon. Women who read the magazines pick them up because they believe there will be value in looking at the magazine and because they think that the particular magazine fits with their sense of who they are. The magazine then reinforces certain ideas of gender roles and behaviors, to the exclusion of other ideas. The feedback loop continues if friends talk about the ads or articles and agree with the advice or fashion ideas.

### **Facilitating multiple conceptualizations of gender**

As a way to allow true multiplicity in definitions of gender, gender roles, and gender orientation, memetics has no peer. The biological analogy of species provides us with different gender species that really don't have much to do with one another except to see the other as a threat or as

a possible symbiotic partner.

## Possible pitfalls of a Memetic Approach to Consumer Behavior

### The Nothing New Syndrome

As is often the case with new ideas, they are mutations of already existing ideas. Memetics suffers this and doing work based around memetic themes exposes the researcher to the skeptical resistance memes of journal editors and department chairs. The biggest danger of these resistance memes stems from the inherent bias to defend existing (held by me the reviewer) memes against all comers. To overcome this, we need to acknowledge that for Consumer Behavior, these ideas have their roots in existing social science theories. Simultaneously, we need to emphasize the benefits to be gained from the "new" parts, those drawn from genetics and epidemiology.

### The "Not Science" syndrome

Memetics, as an infant discipline, and one which seeks to be a meta-discipline, threatens many in social sciences as a usurper. The understandable reaction of many has been, and will continue to be dismissal out of hand, as not science. The greatest difficulties lie with definitions and with a lack of generally accepted methodologies for empirical research. More than most paradigms, memetics is fraught with possibilities of indeterminacy, as the act of observation changes the state of the system.

### Holds itself as a purely positivist paradigm

Even while subject to criticism from those calling it not science, people working in memetic areas seem to be predominantly positivists, bent on objective reductionist explanations for phenomena. This seems to extend to the fact that most will not acknowledge science or the scientific method as memes in their own right. This typical semiotic difficulty of framing one's own paradigm as natural makes unlikely any cooperative collaboration between memeticists of a positivist stripe, and memeticists more comfortable with a phenomenological approach to explanation.

## Avenues for research in gender and consumer behavior using memetics

### The Memes of Gender

Gender memes quite simply constitute the set of beliefs that anyone holds that cause them to behave differently because they are xx or xy and to behave differently toward xx or xy people. Also included in the set are memes about the roles of people other than one's self that one deems as appropriate because of the genotype of the other person. Notice that this does not preclude masculine xx's or feminine xy's, but merely puts a large boundary around gendered memes. "Appropriate" means of course, in the mind of the person holding the meme, and does not cast any social value judgment on anyone's behaviors. Each individual will filter the behavior of others and their own behavior through their own gender memes. Included are all "appropriate" behaviors, all "appropriate" possessions, and all relevant strategy memes as to how they should behave to accomplish their own ends in the context of a world constituted of xx's and xy's.

The description and mapping of gender memes, and more particularly consumption memes with gender components, provides the most immediate application of the concepts of memetics in the

field of consumer behavior. No list that included here could possibly be complete, (What are you wearing today?), but a few examples might illustrate the potential of memetic research into gendered consumer behavior.

**Investigating Gendered Distinction Memes.** The simple questions of what constitutes appropriate behavior given xx or xy chromosomes and what possessions and consumption patterns are deemed appropriate have caused endless debate among scholars and lay persons alike. When viewed memetically, these questions yield to some fairly straightforward analysis. Some analyses like this have already appeared in the literature (Brodie, 1996, Lynch 1997) but there remain many gendered consumer behaviors that have yet to be so analyzed. Furthermore, existing analyses practically invite refutation and rebuttal. The questions answered here will relate to how people divide up the world into male/female masculine/feminine and how those divisions change over time in response to mutations of these distinction memes

**Cataloging Gendered Association Memes** These memes of gender probably cause the most heated debates amongst otherwise likeminded people, as they are the ones that link or associate xx and xy genotypes to what we commonly have called attitudes. So, when a meme spreads around that "boys will be boys" or that "girls are sweet" or any other common cliché, we have a gendered association meme. The extent of these kinds of memes in the population should be documented and tracked over time.

**Cataloging Gendered Strategy Memes** Probably the most problematic of issues that face people in a world of ambiguous sex roles and diverse definitions of gender are those that deal with how to get what they want vis a vis mating and survival. Strategy memes are those that provide rules for behavior under certain conditions. The strategy memes surrounding, getting all the sex you desire, the children you want, and the standard of living you desire have changed dramatically over the years, especially since the invention of reliable birth control in the form of the pill. Where once "The Rules" (Fein and Schneider, 1996), might have been well accepted, they are now rejected by most.

Examples of gendered strategy memes that bear investigating follow:

In order to get the kind of life that you want, find a \_\_\_\_\_ to \_\_\_\_\_ you.

Follow the "rules."

Men should never \_\_\_\_\_.

To get what they want women should \_\_\_\_\_.

Your clothing should accentuate your masculinity/femininity.

**Study of mutations of gender ideas.** Memetics provides a venue to examine how variants of successful gender memes either die out or become more successful than the original memes. For example,

**Study of socialization to gender.** Here fruitful areas of inquiry for consumer research would revolve around such things as how shopping methods and procedures are differentially communicated by parents based on the sex of the child.

**Investigations of how consciously mutating gender memes can work to change behavior.** This now takes us into the realm of investigating memetic engineering. An area that those in the field of memetics seem to think we should be intimately familiar with as consumer researchers are often associated with marketing departments. Study in this area can provide us with empirical evidence as to which characteristics of memes do the most to facilitate accurate transmission of memes. We also need to study which characteristics of memes facilitate rapid transmission of memes.

## Summary

Memetics provides a useful reshuffling of our cognitive deck. Viewing ideas as having a life of their own, and characteristics that improve their chances of repeated accurate transmission from mind to mind gives us a new window through which to view consumer behavior problems. In particular, memetics gives us a slightly less emotionally loaded way to speak about inherently sensitive subjects like gender and the role of gender in consumer behavior.

## References

Bateson, Catherine and Bateson Gregory, (1987) "Angels Fear: Toward an Epistemology of the Sacred", New York, Macmillan.

Brodie, Richard, (1996) "Virus of the Mind: The New Science of the Meme", Integral Press.

Clemens, Samuel L., (1917) "What is man? and other essays," by Mark Twain (pseud.) New York, London, Harper & brothers <http://sunsite.unc.edu/pub/docs/books/gutenberg/etext93/wman10.txt>

Dawkins, Richard, (1976) "The Selfish Gene", Oxford University Press.

Fischer, Eileen and Arnold, Stephen J., (1994) "Sex, Gender Identity, Gender Role Attitudes and Consumer Behavior," Psychology and Marketing, Vol. 11(2) (March./April)

Fein, Ellen, Schneider, Sherrie, (1996) "The Rules: Time Tested Secrets for Capturing the Heart of Mr. Right ,," Mass Market Paperback.

Henson, H. Keith (1989) "Memetics: The Science of Information Viruses" Whole Earth Review, (Winter, #57).

Lynch, Aaron, (1996) "Thought Contagion: How Belief Spreads Through Society," BasicBooks.

Lynch, Aaron, (1997) "Gedankeninfektion Wie Überzeugungen Menschen Finnen gdi-impuls #3 September, pp 42-54. English translation

Mahoney, Rhona, (1996) "Kidding Ourselves: Breadwinning, Babies, and Bargaining Power", BasicBooks. On-line Introduction

Westoby, Adam, (1994)"The Ecology of Intentions" unfinished manuscript posthumously available on the World Wide Web

# **CREATIVITY, EVOLUTION, AND MENTAL ILLNESS**

**By**

**Antonio Preti and Paola Miotto**

CMG, Psychiatry branch

via Costantinopoli 42, 09129 Cagliari, Italy

tel + 39 70 480922, fax + 39 70 499149

[apreti@tin.it](mailto:apreti@tin.it)

## **Abstract**

Mental representations, or memes, transform the space in which they evolve. Their survival is dependent on the survival of the individuals and the groups hosting them. Creativity - the production of new and useful ideas - is closely linked to the social dynamics of the individuals expressing creative ideas: without social confrontation new memes cannot become diffuse. Creative individuals tend to be emotionally unstable, and many are affected by mental disorders. Studies on the link between creativity and mental illnesses show that it is exactly the characteristics of the mental disorder which also confer some advantage on afflicted individuals. These advantages extend to the groups to which the creative, mentally ill individuals belong. The group comprising the most creative personalities will therefore acquire an adaptive advantage which maintains the integrity of the group as a whole, in spite of the vulnerability of the individual.

**Key words:** creativity, evolution, meme, mental representation, mental disorders, bipolar disorders, schizophrenia, natural selection, information processing

## **1 Introduction**

Mental representations, or memes, transform the space in which they evolve. In a recent paper on this topic, Liane Gabora stated: "*culture not only affects biological fitness through its effects on behaviour... but it dramatically modifies the biological world*" [7]. Creativity - the production of new and useful ideas - is closely linked to the social dynamics of the individuals expressing creative ideas: without social confrontation new memes cannot become diffuse: "*creativity is a collective affair*" [7].

The majority of studies on creativity have focused on the characteristics which favour it at an individual level, but all researchers are conscious of the social nature of the creative process. It is not easy to define exactly what creativity actually is. To create, per se, implies the production of something new and original. However the qualities which make an individual able to produce new entities are not well understood. As is also the case with the concept known as "intelligence", it is unlikely there is a general creativity factor unevenly distributed across the population. Rather, creativity can be conceived as a complex of qualities that allow some people more easily than others to produce new objects or ideas. As a consequence creativity is often described in relation with the end product, in a way that can occasionally be tautological:

creativity is often conceived as that which has lead to results considered creative by general consensus.

Measuring creativity is not an easy task: the methods used in the evaluating of creative aptitude and ability are numerous and as ingenious the argument investigated demands them to be. Summarizing the different criteria used for measuring creativity Dennis Hocevar [10] reviewed ten main categories:

- Tests of divergent thinking
- Attitude and Interest inventories
- Personality inventories
- Biographical inventories
- Teacher nominations
- Peer nominations
- Supervisor ratings
- Judgment of products
- Eminence
- Self-reported creative activities and achievements

Most of these techniques are based on third person rating. Although inter-rater agreement is generally obtained, the problems of "*who judges the judges*", and what the judges should be looking for, remain unresolved.

A minority of methods rest on independent tests of measurement: these tests, like the tests of divergent thinking, are based on specific models of what creativity is, so any consideration of their results is biased by the personal view of the researcher who reported them. However, as in the case of "intelligence", there is probably no way to avoid this dilemma. Ultimately, any discourse on creativity, or any other conceptual construct, must involve a description of the boundaries of the concept discussed.

The description of creativity that is now generally accepted identifies creativity as the ability to create products or ideas which are original and which posses a strong social usefulness. Frank Barron [2], one of the most authoritative researchers in this field, notes that a creative product or idea should be new or original and be acknowledged as such by social consensus: a criterion of usefulness is implied by, although not essential to this definition. In this view "Creations" are products which appear new and are also considered valuable by consensus. To introduce an innovation requires the comparison of the new idea or product with the pre-existing alternative, and the evaluation of the innovative product with respect to existing needs. A creative result can serve to resolve a problem or to satisfy an unconscious aesthetic drive: although this can be experienced at an individual level, in the majority of cases the problem or the aesthetic impulse operates on a social level, and the creative solution must be communicated to others if it is to produce some effect.

Many factors contributing to the expression of creative potential are a reflection of this essential dimension of the creative process: i.e. the need for social testing of the creative product. The proper elaboration of the primary creative effort requires the development and the testing of the

new idea against scientific and aesthetic standards, which are, of course, of a social nature [15]. There are many studies which explore the link between creativity itself and the psychological characteristics of creative individuals, and which, in so doing, shed light on this argument [2,12,13,21].

Generally, creative people tend to emerge as both sociable and popular with their peers; they appear to welcome social contact and show interest in social activities. Creatively gifted people more than their peers also show risk-taking and novelty-seeking tendencies. While this grants them greater independence and unconventionality, it also leads to emotional instability, and in fact one of the most surprising psychological characteristics of creative individuals is their proneness to mental illnesses [10, 11, 12, 13].

## 2 Creativity and Mental Illnesses

Most studies show that there is a link between creative ability and the risk of mental disorder: in fact the prevalence of mental problems among creatively gifted people is significantly higher than in the general population [2, 11, 12, 13, 15, 21]. A link between eminence in the arts and science, leadership qualities and mental problems was even suggested by the author of the fragment known as "*Problemata XXX*", now part of the aristotelian canon. In this text, many behavioural characteristics are indicated which are peculiar to patients suffering from manic-depressive illness. This would suggest that eminence as result of creative aptitude and madness are linked by a non-casual link.

In the positivistic era, the Italian Cesare Lombroso illustrated this thesis in a circumstantial way referring to the "genius" and the "lunatic" related, in his opinion, by a shared genetic basis [14]. Most studies performed in the positivistic era, in order to either confirm or refute Lombroso's hypothesis, rest on biographical evidence, raising the suspicion that these studies claiming a higher prevalence of psychopathologies among creative or eminent people were biased by overexposure. However even the later studies, performed using methods applying specific nosographic categories and direct confrontation with the candidate through interviews and inventories yielded similar results, with a higher prevalence of mental disorders among gifted people than among the general population [12, 13]. In the majority of studies disorders with an affective (mood) component, in particular those characterized by melancholy (severe depression), were associated with creative achievement, just as the author of the *Problemata XXX* asserted [1, 11].

The disorders most implicated are those with a bipolar trend: i.e. disorders characterized by the alternation of depressive episodes with other episodes with an opposite euphoric mood. In the period between episodes those who are suffering from these disorders have a highly adaptive life style, with generally very high levels of social functioning, thanks to the tenacity, tirelessness and social ease that are typical of these individuals [8]. Nevertheless the disorders can be highly impairing: bipolar patients are often inconstant at work and their affectionate and social relationships tend to be stormy; their divorce rates are higher than those of the general population, and there is a very high prevalence among them of misuse of alcohol and other drugs, which bipolar patients may take either as auto-therapy to lessen anxiety and depression, or during impulsive euphoric periods [8].

Studies on the relationship between creativity and mental illnesses suggest that it is the same characteristics of the disorder, in their less severe manifestations, which confer some advantage on afflicted individuals and their relatives. Ruth Richards, who extensively studied this aspect of bipolar disorders, found a high propensity among bipolar patients and their relatives for the development and expression of creative potential in every field [19]. These abilities were evident mainly among individuals with less impairing or subclinical forms of the disorder.

Among the characteristics that favour creative achievement among bipolar individuals there are some that are specifically related to ability in social intercourse. People with bipolar mood disorders tend to be more emotionally reactive, which gives them greater sensitivity and acuteness. A lack of inhibition permits them unrestrained and unconventional forms of expressions, less limited by accepted norms and customs. This makes them more open to experimentation and risk-taking behaviour, and, as a consequence, more assertive and resourceful than the mean. Sensitivity and lack of inhibition make these subjects warmer and more friendly in social intercourse. Both aptitudes also represent a clear advantage on a professional level, particularly when competition is greater. Being more sociable and less inhibited in expressing themselves, individuals prone to bipolar disorders may manage to spread their ideas with greater ease, allowing their ideas to prevail over others competing for predominance.

### 3 Creativity, Memes and Evolution

These observations show the importance of the social side of creativity. Creative ideas, like other memes, reflect the dynamics of the individuals who host them. New ideas can therefore find success only when their hosts, and the groups to which these individuals belong, are also successful. In his studies Cavalli-Sforza showed that the diffusion of grain cultivation was directly linked with the spread of the population which elaborated the technique [5]. Another example is the heroin abuse epidemic: where there are no drug abusers there is no heroin to be sold.

As Liane Gabora asserts in her paper: "*the bottleneck in cultural evolution is the capacity for innovation*". Groups whose members possess the ability to make innovative mental associations can take advantage from this, even if the ability is linked to maladjusted behaviour at the individual level. Examples deriving from the studies on the relationship between creativity and bipolar disorders are suggestive. The high suicide rate (10% of patients) and the low fertility of bipolar patients, who tend to marry less than the general population and have less children than the mean, suggest that, both now and in the past, the carriers of the genetic burden of the disorder must have some compensatory advantage in order for a relatively high percentage of affected individuals (more than 1%) to be maintained in the general population [24]). Bipolar disorders have a strong genetic component [16]: people who suffer from them, and their relatives, share some genetic factor which is transmitted from one generation to the next.

Propensity to develop a bipolar disorder can be conceived as a hitch-hiker allele which confers no compensatory advantage per se, but which endures because it is linked to creative abilities important for survival. The bipolar gene may contribute to its own survival in the gene pool by virtue of the resistance, tenacity and energy it confers on the individual, or may favour behaviour

which contributes to its replication. If creative individuals are more likely to assume leading positions in various fields, then hitch-hiker alleles that contribute to creativity would also spread. There are many creative individuals with associated mental problems in the artistic field [11]. Poets and writers seem to be particularly prone to developing mental problems, generally of a depressive type, and it has even been asserted that one cannot write with success without being "exposed to the Dark Sun of Melancholy" [15]. Joseph Schildkraut, one of the fathers of the biological approach to the study of mental disorders, put forward a very suggestive hypothesis attributing a decisive role to artistic symbolism in the favouring of social cohesion of groups.

Thus the group containing the most creative personalities will acquire an adaptive advantage which maintains the integrity of the group as a whole, in spite of the vulnerability of the individual who is subject to depressive breakdowns [21]. This explanation derives from the concept of inclusive fitness, according to which a behaviour pattern under genetic control is selected for its ability to favour the production of successful offspring both by the individuals displaying the behaviour pattern and by their genetic relatives. Group selection has often been thought to favour traits that are individually disadvantageous but evolve because they benefit the wider group [22, 23].

The same results could also be associated with memes (such as respect of agreements or the stifling of revenge) that may be detrimental to the individual but clearly offer an advantage to communal life. Well known example are the prisoner's dilemma and the TIT for TAT strategies\*1.

## 4 Creativity, Schizophrenia and Information-processing

Studies on the relationship between creativity and mental illnesses also throw light on the role of the gating systems involved in information-processing. The creation of new ideas requires combinations and transformations of old ones: we can only hypothesize about what we do not yet know. However, information processing is not a fixed, mechanical procedure. The flow of inputs is guided by centralized systems which regulate the amplitude and length of oscillations in the canals that transfer the information.

Like schizophrenic patients, creative individuals often report odd sensory and perceptual experiences, feelings of restlessness and the inclination towards impulsive outbursts in association with rejection of common social values [4, 6, 25].

Highly creative "normals" also tend towards over inclusive or "allusive" thinking and, as pointed out by Albert Rothenberg [20], demonstrate a capacity to conceive and utilize two or more opposite or contradictory ideas or concepts simultaneously, without being disturbed by this simultaneity of opposition, as is also the case with schizophrenics. It seems that creative individuals, like schizophrenics, are subject to a widening of selective attention, which makes them more aware of and receptive to experience, with more intensive sampling of environmental stimuli [9, 19]. In fact ideational fluency and a preference for complex and asymmetrical designs, two of the main factors contributing to creativity, could derive from higher levels of arousal and faster stimulation of discrete cerebral areas. Schizophrenic thought processes tend to allow unusual associations which result in over inclusive thinking, with many irrelevant elements

being included in reasoning: this peculiar style of thought is assumed to derive from a failure in the filtering of stimuli by dysfunctional gating systems [3, 18]. Creative individuals, conversely, may gain advantage from higher levels of associative thinking, since they are capable of effectively processing these increased inputs without the risk of cognitive overload. Since to create consists essentially of the making of new combinations of associative elements [3], any ability which serves to bring together otherwise remote ideas will facilitate a creative solution [9, 17]. The favouring of associations implies an extended knowledge of the argument under study (memory of ideas to be associated) and a restriction of inhibitory influences on stimulation of remote cerebral areas. The more associations evoked by an element, the more likely it is that another element will be combined with it in a manageable form. Since inhibition or suppression (by anxiety or other more powerful competitive stimuli) would limit awareness and openness to both internal and external stimuli, freedom from these forces would favour associative thinking, and so creativity.

## 5 Conclusion

One should not ignore that the stereotype of the eccentric artist or of the mad scientist also plays a protective role in the collective imagination. It acts against the fear and suspicion that excellence and the diversity of others always engenders in the majority. The image of madness linked to genius has been repeatedly expressed in the history of the western world. It was codified during the Renaissance in the figure of the melancholic genius afflicted by Saturnian acedia, and underwent a resurgence during Romanticism in the figure of the tormented artist. Creativity is, after all, a challenge to an existing meme pool, since it leads to the creation of competing memes. The stigma attached to diversity may lead individuals to assume even more eccentric behaviour: the definition of such behaviour in terms of "mental illness" may result from "strategies" defensive towards existing memes. In fact it is the pool of existing memes that decides what is good and healthy, i.e. coherent with its own criteria. In addition, some temperamental traits, like eccentricity, uneasiness or propensity to excess and experimentation, which are widespread among creative people, could be a reflection not only of an underlying mental disorder, but also, and above all, of the tolerance on the part of society of the behaviour of famous individuals. In these cases deviant behaviour is tolerated because it allows the expression, by third parties, of dissenting demands which the majority of people are not able to express.

Many issues concerning the "dark side" of creativity remain unresolved. Equally, the same contribution of new memes to their own diffusion is only partially understood. A particular set of productive memes, for example, may exert a selective influence on the appropriate neuronal groups which in turn lead a particularly creative individual to produce creative things. We hope this paper will achieve this same end, conferring a competitive advantage towards an understanding of the creative process and cultural evolution on the memes detailed herein.

## Acknowledgements

The authors would like to thank two anonymous referees for their advice and comments on a early version of this paper, and Mr. Thomas Eagle for help with the revision of the English translation.

## References

- [1] Andreasen NC, Glick ID (1988). Bipolar affective disorder and creativity: implications and clinical management. *Comprehensive Psychiatry*, 29: 207-217
- [2] Barron F, Harrington DM (1981). Creativity, Intelligence and personality. *Annual Reviews of Psychology*, 32: 439-476
- [3] Braff DL and Geyer MA (1990). Sensorimotor gating and Schizophrenia. *Archive of General Psychiatry*, 47: 181-188
- [4] Cattell RB, Drevdahl JE (1955). A comparison of the personality profile (16 P.F.) of eminent researchers with that of eminent teachers and administrators, and of the general population. *British Journal of Educational Psychology*, 46: 248-261
- [5] Cavalli-Sforza LL, Feldman MW (1981). *Cultural transmission and evolution: a quantitative approach*, Princeton University Press
- [6] Dykes M, McGhie A (1976). A comparative study of attentional strategies of schizophrenic and highly creative normal subjects. *British Journal of Psychiatry*, 128: 50-56
- [7] Gabora L (1997). The origin and evolution of culture and creativity. *Journal of Memetics - Evolutionary Models of Information Transmission*, 1, [http://jom-emit.cfpm.org/voll/gabora\\_1.html](http://jom-emit.cfpm.org/voll/gabora_1.html)
- [8] Goodwin F, Jamison KR (1990). *Manic- Depressive illness*. Goodwin F & Jamison KR(eds), Oxford University Press, New York
- [9] Hansefus N and Magaro P (1976). Creativity and schizophrenia: an equality of empirical constructs. *British Journal of Psychiatry*, 129: 346-349
- [10] Hocevar D: Measurement of creativity: Review and critique. *Journal of Personality Assessment*, 1981; 45: 450-464
- [11] Jamison KR (1993). *Touched with fire*. The Free Press, Macmillan, USA
- [12] Juda A (1949). The relationship between highest mental capacity and psychic abnormalities. *American Journal of Psychiatry*, 106: 296-307
- [13] Karlsson JL (1978). *Inheritance of creative intelligence*. Nelson-Hall, Chicago
- [14] Lombroso C (1891). *The Man of Genius*. (original edition in Italian) Walter Scott, London

- [15] Ludwig AM (1989). Reflections on creativity and madness. *American Journal of Psychotherapy*, 43: 4-14
- [16] MacKinnon DF, Jamison KR, DePaulo JR (1997). Genetics of manic depressive illness. *Annual Reviews of Neuroscience*, 20: 355-373
- [17] Mednick SA (1968). The associative basis of the creative process. *Psychological Review*, 69: 220-232
- [18] Preti A (1995). The role of dopaminergic pathway in the neurobiology of schizophrenia. Critical issues in the treatment of schizophrenia, Florence, 102-103.
- [19] Richards R, Kinney DK, Lunde I, Benet M, Merzel APC (1988). Creativity in manic-depressives, cyclothymes, their normal relatives and control subjects. *Journal of Abnormal Psychology*, 97: 281-288
- [20] Rothenberg A (1971). The process of Tunisian thinking in creativity. *Archive of General Psychiatry*, 24: 195-205
- [21] Schildkraut JJ, Hirshfeld AJ, Murphy JM (1994): Mind and mood in modern art. Depressive disorders, spirituality and early deaths in the Abstract Expressionist Artists of the New York school. *American Journal of Psychiatry*, 151: 482-488
- [22] Wilson DS (1993): Group selection. In: Keller EF and Lloyd EA (eds), *Key words in evolutionary biology*. Cambridge Mass, Harvard University Press.
- [23] Wilson DS (1983). The group selection controversy. History and current status. *Annual Reviews of Ecology and Systematics*, 14: 159-187.
- [24] Wilson DR (1994). Evolutionary epidemiology. *Integrative Psychiatry*, 10: 6-12.
- [25] Woody E, Claridge G (1977). Psychologising and thinking. *British Journal of Social and Clinical Psychology*, 16: 241-248.

**THE LEARNING ORGANISATION MEME: EMERGENCE OF A  
MANAGEMENT REPLICATOR**  
**Or**  
**Parrots, Patterns and Performance**

**By**  
**If Price**

The Harrow Partnership, Pewley Fort, Pewley Hill, Guildford GU1 3SP UK  
[and Unit for Facilities Management Research, Sheffield Hallam University]

01483 574692

I.price@shu.ac.uk

and

**Ray Shaw**

The Harrow Partnership, 28 Rushworth Road, Reigate RH2 0PE UK  
01737 242895

*In Campbell . and Duperret Than V. Eds. 1996 Proceedings of the third ECLO Conference,  
Copenhagen May 1996*

Organisations and organisms are self-maintaining systems which spontaneously seek to preserve an evolved order. Both are enabled by replicators: memes or genes respectively. Whereas genes are the units of transmission of our biological inheritance memes are the units of transmission of our cultural inheritance. They cause organisations to settle into patterns, routines and habits of behaviour: manifestations of a particular memetic inheritance. These patterns enable the organisation but simultaneously limit its performance.

Both systems share the evolutionary dynamic of adaptive radiation followed by stabilisation. Memetic examples include new markets, new technologies and new business ideas. Business theories and their derivative, managerial fads, are a class of memes. This paper illustrates the increasing returns dynamic in the evolution of management recipes by contrasting Business Process Re-engineering and the Learning Organisation. It ends with a plea for the Learning Organisations to retain memetic diversity rather than be trapped in sterile competitions to define an LO. **The power of the Learning Organisation movement may, paradoxically, be that we are not stuck with what it is.**

## **IS ECLO ALIVE?**

This paper ultimately concerns the patterns of language and thinking which stop organisations, and individuals, learning as fast as they might, thus stunting their growth. Interrupting conventional language patterns facilitates interruptions of conventional thinking and we, intentionally, choose here a different and interrupting style of writing: hence the starting question.

Our hope is that we can extend a different invitation to the reader. We would rather stimulate dialogue and enquiry into that which limits, in any given context, than into the rightness or wrongness of a theory. When so-called learning is no more, or is perceived as no more, than an exercise in intellectual one-upmanship, or when it is a debate among spectators unconcerned

with helping themselves, and others, make a difference, then it is, for us, limiting. Please read not only to enjoy but also with a view to making a difference.

Consider, for a minute, the definition of life offered by the English biologist Richard Dawkins (9)

**[Life is]...a property of improbable complexity possessed by an entity that works to keep itself out of equilibrium with its environment.**

To exapt Dawkins most graphic example a dead parrot thrown into the air obeys the laws of physics, describes a perfect parabola, and then falls back to earth. A live one disappears over the county boundary; its component parts working together to maintain their collective entity against the force of gravity.

Now - using that definition what is different about a company, or any other form of organisation, including the ECLO? All organisations maintain an improbable complexity to achieve some purpose that their component people and parts, working separately could not achieve. [See (14) (27) (28) (36) for further discussion of theoretical comparisons].

Professor Herman Van den Broek started the ECLO conference series by reminding us ***how quickly people adapt to tacit agreements and recreate recognisable settings*** (35). We operate to keep ourselves in our own equilibrium rather than succumb to the random decay of the second law of thermodynamics. Yet, in the process, our organisations and sometimes even we ourselves fail to learn. How can we be so stuck in the rich tapestry of our collective achievement? Nothing if not ambitious, we offer here a very simple answer to the riddle.

## **PARROTS & PATTERNS OF PRESERVATION**

Why do we all tacitly participate in structures that we know are not yielding us the result to which we aspire?

Why do we all conspire in corporate behaviours that each of us may admit, privately, are suicidal?

Herman van den Broek's enquiry was one of many which have pointed out how organisations settle into patterns of established behaviour. In a spirit of European integration let us acknowledge that Machiavelli made the same point 500 years ago. Organisations coalesce around rules, habits, traditions, customs, codes, mental models, paradigms, language: shared patterns of taken for granted being-in-the-world. They maintain these patterns as surely as a live parrot maintains the long tradition of parrot-hood. The parrot ultimately has no choice in the matter. It is the creation of a tradition encoded in countless generations of parrot genes. It owes its being in the world to those genes relentless search for their own replication (8) (10).

Not to put too fine a point on the matter, and ignoring only time scale, even those of us who espouse the cause of organisational learning and human choice are usually as stuck with our patterns and traditions as are generations of parrots with their genetic inheritance. Standing for a

different perspective on the world, a different outcome in a company, or even, dare one say, a different presentation in a conference with an established tradition, risks the fate of a "green" in a colony of "*Norwegian Blues*", or of a member of any community who does not conform to the mores and paradigms of membership. Companies reject, conferences have no listening for, and individuals do not hear, that which does not conform to their expectations on the world. It is safer then to talk about learning within an established pattern rather than risk membership. Yet, for a pattern to be changed the risk may be needed. We, none of us, learned to walk without risking abandoning the comfortable pattern of crawling.

## WHAT SOURCE THEN THESE PATTERNS?

For the parrot, and every other living species, the source of the pre-ordained order, the root, at one and the same time of the ability to maintain ***improbable complexity*** and the stuckness in the tradition of a particular species, **is** very simple. It is the genetic code, the genome of the species or the genotype of the individual. All species, including our own, are lumbering robots created, specified by their genes, to act, and interact with others, in ways that maximise the chances of those genes leaving copies of themselves in the world.

As many are now pointing out [see postscript] the parallels of organisation and organism extend beyond the maintenance of improbable complexity. Both are complex systems that evolve and adapt, that self-organise in webs of complex relationship. Both depend for their survival on strategies, explicit or otherwise that anticipate the future: usually as a continuation of the present. Both depend, for their ability to maintain order on what Murray Gell-Mann (14) has called schemata or Richard Dawkins (8) replicators: codes bent on their own reproduction through the systems they create.

DNA as a genetic replicator provides a biochemical language through which are encoded and transmitted the instructions that enable and specify every complex, self-maintaining, living system (22)(23)(36). The individual gene is a strip, a phrase, in this language that survives because it manages to convey utility to its host in the competition for survival in the organic economy. To do so the gene must act in the context created by the thousands of other genes in a given genome; as words and phrases, to convey meaning, act within the context of the language and tradition in which they are used.

Our starting proposition is simple. It is to take literally Dawkins suggestion that there exists in the cultural world a second class of replicators, - ***memes*** - transmitted from brain to brain through human language and cultural artefacts.

The distinction meme (10) (27) (28) may not yet be one that has a context in the language and intellectual tradition of many readers. It does not yet exist as a separate perceptual and linguistic entity through which many people view the world. So what is it? In brief:

***a unit of cultural transmission that propagates from mind to mind conveying a meaning in the process; a meaning that may mutate each time it is transmitted but one that is captured and remembered.***

Any linguistic or other cultural artefact can serve. As Dawkins puts it:

*When you plant a fertile meme in my mind you **literally parasitize my brain**, turning it into a vehicle for the memes propagation in just the way that a virus may parasitize the genetic mechanism of a host cell.*

Tunes, fashions, advertising slogans, phrases, linguistic distinctions, gods, technological recipes, paradigms, mental models, perceptual filters and all the stuff of tradition, language, culture and cognition are memes and complexes of coexisting memes. The advertising industry runs on memes without bothering to explore the fact. Every slogan, every image it creates, is designed to infect the mind and influence the actions of the recipients. Hence its fundamental irrationality (21).

A geological metaphor provides an image of the physiology and psychology of memes at work upon the individual, or collective, mind. Imagine a landscape, eroded over time to provide streams, rivulets, and rivers interspersed between higher plateau. It provides a simple example of a self-organising, locked-in, system. Over time accumulations of rainfall carve out stream and river beds and settle into pools and lakes. Any new rainfall will no longer find its own way but will rather take, and reinforce, the already sculpted path of least resistance. Though the falling rain may be evenly distributed across the land, in its collection and flow across the landscape, it will tend towards a predetermined route, one taken by previous rainfalls.

Just as the rainfall follows established routes so perception follows established ways of seeing. Technically, even if the light sources which perturb the back of the retina, or the acoustic waves the eardrum, are identical, what will be noticed from all that could be seen or heard will depend on the perceptual lens through which we view the world. What is there is not independent of the receiver. What is there is what we have been trained [or conditioned or have learnt] to see. We may discard, indeed we can be blind to, anomalies that do not fit. The self-organised pattern which we call our thinking grants a particular perceptual blindness and rigidity to our perceptions of the world.

Exploring the analogy further we could say that an idea, a single thought, an utterance, *-a meme in fact*, is like the single raindrop. It falls upon a pre-formed perceptual memescape. Isolated thoughts gather together in a string - a pattern of co-existing memes - which we might compare to a few drops congregating together in a splash of water. With sufficient mass the splash of water starts to flow into streams and rivers which are, if we like, the connectors between the raindrops and the pools and lakes, if not the oceans, of our thoughts. The pools and lakes we may view as concept pools and theory lakes. Thus a self organising system is inherited and developed in which the flow of perception takes a certain course: it follows a certain pattern, a largely given paradigm.

Patterns in companies, habits and rules of behaviour, codes of thinking, systems of language, states of relationship coalesce in similar fashion around shared landscapes of perception (6)(16)(27)(28)(29).

Take this simple idea of memes as a reality in the world and you arrive at an explanation of a plethora of observations from many field of scholarship concerning the behaviour of human systems with self-fulfilling prophecies, self-replicating patterns, mental-models [e.g. (2) and unwritten rules [e.g.(6)(27)(31)] at their core. The **ECLO** is a meme around which a pattern called conference has emerged; a pattern similar to that of many another conference but a pattern with its own emergent tradition of rituals, conventions and unwritten rules of behaviour; a pattern that both enables and circumscribes what we may or may not learn from one another within the context of this conference.

To understand the phenomenon of emergent organisation we should ask not what benefit it confers on the participants but what benefit it confers on the meme. How does the emergent tradition of the conference assist a meme called ECLO infect more minds?

## MEME'S EYE VIEW

Dawkins main argument in *The Selfish Gene* is that to understand evolution and its products we should examine it, and them, not from the benefit a particular phenotypic behaviour confers on an individual organism or a species, but from the perspective of a genes [or a complex of genes] replication. We want to make the same point in respect of memes in organisations. From the perspective of one set of memes enthroned in an individual mind, or shared by a collective organization what matters is their replication in meme-space, not the benefit or otherwise conveyed to the host. As Dennett (10) traces, many memes in the paradigmatic landscape of the humanities triggered defensive routines to Dawkins suggestion; an example of the generic "**not invented in my head**" syndrome that inhibits learning at many levels (3).

We want now to demonstrate this switch of perspective by comparing the evolutionary history in the world of two memes; two ideas that have infected the minds of a population of organisational theorists and practitioners in recent years. One of these will, we suspect, be easier for readers of this volume to visualise as an alien virus seeking to parasitise other minds so we will start with that one.

## The Bad and the Ugly? BPR as a Meme

Think back some eight years to the late 1980ís when the Business Process Movement first became noticeable in management practice and language. Around about then, in most corporations in most sectors, the pressure to extract more value from white-collar activity became overwhelming. The competitive advantage of doing so, and the threats in not doing so, were simply too large for companies to ignore. The environment was ripe for new approaches to structuring work and organisation.

From the flux of variation and mental mutation that is the soup of business theory and practice there then arose several variations on the theme of Process Management seeking meaning space in the new environment. Most could be considered adaptations [more strictly exaptations] of prevailing manufacturing approaches. Most large American corporations and some European ones announced process as a key plank of their strategy for the 90ísí. As one director of a British company recently recalled to us "*My boss went to the USA and caught the BPR bug*" Yet it was

not then the only bug in town. A brief survey of the popular management literature in late 1990 and early 1991 would have revealed:

- Business Process Review
- Business Process Simplification
- Business Process Management
- Business Process Innovation
- Business Process Improvement
- Business Process Control
- Business Process Transformation
- Business Process Re-engineering

Each came with subtle shades of meaning and recipe depending which corporate initiative or guru you were talking to. The differences were more theoretical than significant. Now, five years later, we hear only the one term used.

That dynamic - widespread proliferation and experimentation followed by stabilisation around one, or a few designs, is common if not universal in the introduction of new technologies. Consider the history of the motor car, the aeroplane, the personal computer, the video cassette tape or the typewriter keyboard. When a new technology, or idea creates a new niche in the system of interactions that is an economy, or when a new demand opens a new opportunity, the response of the market is a radiative bloom of innovation followed by stabilisation around one, or a few, dominant designs (4)

A similar dynamic is found in new industries as in new technologies. New companies enter new markets, either as new entrants or when established players seek to diversify beyond their original sector. The market for Facilities Management services in the UK is currently undergoing just such a bloom with participants competing for share on the basis of traditions in IT, technological services, property and estates management, architecture and interior design, catering and hotel services, engineering, construction, image processing and consultancy. As the market matures a new stability will doubtless emerge (30). In the process new business relationships, and new patterns of relatedness, are emerging to challenge old ideas about what is core and non-core and how such services are purchased (6).

The dynamic is also an old one. The first instance on record occurred 570 million years ago in the so-called Cambrian Explosion when the technology of multi-celled life first expanded into the opportunity space of a sea of edible algae. Life experimented with a range of fundamentally different archetypes before settling down to the score or so of phyla which have survived to this day. The story of the Burgess Shale made famous by Stephen J Gould (17) serves as an example of a pattern since repeated at many scales of biological and cultural evolution.

But to return to fads in general and BPR in particular. The rise of fads in management is well documented (e.g.32), as is the disillusionment of many who practice the latest recipe only to find it does not yield the result they expected. BPR is not unique in this regard but - in seeking an explanation - our normal convention is to ask either *What's wrong with the recipe* or *What's*

*missing in the people who tried it, such that it did not work.* We add recipes to cure the failure of the last recipe heaping fad upon fad rather like the old lady who swallowed the fly.

What we do not do is look at the rise and fall of a business fad from the perspective of the fad itself, as a replicator bent on blindly infecting as many minds as possible. We do not ask: What is seeking its own replication here? Consider the following thought experiment.  
BPR Inc.

We invite you to stand in the meme's perspective. See the world through its eyes. Perhaps it will help to imagine yourself as the CEO of \*New Fad plc\*. Your shareholders will judge your reward by the number of new minds you 'infect'. This is not dissimilar to running an advertising campaign by the way. Launching and marketing a fad is not unlike launching and managing any other product [except the fad that succeeds promotes itself].

**Stage 1.** Your 'product' has got to have some sort of relevance out there. Launching a product that the market not only doesn't know they want, but aren't ready for yet, is a recipe for commercial disaster. So is the premature launch of a new idea, business solution or paradigm. Galileo, for example, had this problem with the prevailing paradigms of the day. Just as there was no listening for a heliocentric solar system in the prevailing orthodoxy of medieval Europe so, before about 1990, there was simply no listening - no space in the collective business meaning-scape for Process. When the need emerged many memes competed, as we have seen, to fill it.

**Stage 2.** You have to find some niche. If your fad is not relevant and useful to someone forget it. Any fad that is going to get talked about and written about - propagated in the world - must, almost by definition, first have to be useful to someone. Those who first label a new management technology are unlikely to talk enthusiastically about what failed.

**Stage 3.** If you are 'lucky' [and there are a whole lot of questions buried between those quote marks], your fad hits/ creates a rising market. When a lot of essentially similar fads offer much the same thing under a slightly different name [or brand] some will succeed and some will not. It may be pure chance as to which first gets critical mass. Replay the tape of business theory since 1990 and we might be talking about BPT or BPI rather than BPR [just as we might be playing the tape on Betamax or Video 2000 rather than VHS format machines].

**Stage 4.** Early share is critical because of the process by which a fad replicates. As it succeeds consultants with a living to make, managers with a name to make, business school professors with an image to make, journalists with copy to fill, or just plain seekers after enlightenment start to play. Positive feedback rules. Just as VHS became a more favoured format the more users it had (4) so BPR became the easier language to use the more it was used.

For the really successful fad a whole structure of articles, books, recipes, societies, conferences, internet discussion lists etc. etc. emerges. This structure has a vested interest in the replication of the meme by which it is infected, even if it is not sure what that meme means any more.

Congratulations CEO you did it if you got to this stage. Go directly to the bank and collect your \$200.

But think, while you do what has happened:

Somewhere in the process above a switch occurred. To infect other minds, to acquire believers, reputation becomes more important than performance. Whereas early sales may be on the basis of benefit conferred to a company later sales are on the basis of status conferred to an individual and even current fashion. We have the fad because everybody has it.

Memes remember have no foresight. They exist to go on replicating and preserve the structures and behaviours which enable them so to do. A virus succeeds even if it kills those it infects provided it is transmitted first. Syphilis increases libido [or so we have read]. Re-engineering may drive companies to corporate anorexia (12) in the name of being lean and mean but that is of no concern to the re-engineering meme, provided only that it can go on replicating. Its lack of foresight may ultimately be its downfall. Space is created for a new fashion in articles asking: What's wrong with BPR. New more vigorous fads invade old territory until they, too, keep people equally stuck.

## THE LEARNING ORGANISATION MEME

The concept of Learning Organisations emerged into the world of business theory at about the same time, or shortly before the BPR meme. Interestingly it seems to have evolved largely separately several times. Perhaps learning as meme already had its niche in the language of academia. What researcher after all does not like to think that they belong to an institute of learning [and how many university departments that you know would you really classify as learning organisations?].

Perhaps also the learning organisation has had a harder time infecting the listening of the practical business audience because the language [the meme] triggers more defensive routines in the memetic pattern of the practicing manager who has escaped all that stuff. Many a successful manager is after all doing, what observers might choose to call organisational, or action, learning without needing the benefit of the learning organisation meme. They may call it innovation, or performance, or even Process Re-engineering. They may even simply call it managing!!!!

We do not actually know what a Learning Organisation is. We do know that competing definitions abound and that is there is no common sense of meaning conveyed by the term, even within the bounds of this conference. John Harvey -Jones gives a, relatively pragmatic, senior businessman's definition when he refers to the Learning Organisation as *the Philosophers Stone of Business; the thing everyone is searching for that no one he knows has found*. By contrast Thurbin (34) argues it can be reached in 17 days! Some equate the LO with the capability to adapt, react and change in contrast others [especially Senge and those inspired by him] who would rather reserve the LO as something of a vision, a quest, a search for generative innovation and a new sense of community in organisations; one in touch with traditions other than the dominant paradigm of western business.

Another more operational viewpoint sees a LO as one in which ideas and lessons learnt are, openly and quickly adopted elsewhere in the company, or for some pioneers the whole supply chain. Not many companies manage such learning well despite databases of good practice, and

books of lessons learnt. Some professional services firms do it very well and a few companies have cultures where it just happens as a matter of course.

Other schools seek to capture the high ground of learning for a particular function within the corporation, usually the corporate centre hence planning as learning benchmarking as learning audit as learning and even control as learning. Most notably of course this pattern prevails in the HR function where learning often appears as a flag behind which the personnel and training specialists fight for their place in the sun and share of the corporate resource. Consider this definition of the top [sixth] level of maturity of an organisation from a recent article in the UK *is* IPD magazine:

**Training and Learning are the process through which strategy is formulated:** Are HRD people at the top table as both actors and facilitators in policy formation? Is the organisation acquiring the characteristics of a Learning Company?

Why the right hand state is a prerequisite of the left is not explained.

Yet, albeit gradually, a Learning Company is becoming perceived, even by pragmatic business people as something they want to be. The Learning Organisation is not so much a meme as an opportunity space that many other memes are seeking to occupy. Whereas various, similar, recipes and names competed for market share of Process Management space, different meanings compete for the LO label. The confusion - and righteousness - thus engendered probably hinders any of them becoming dominant.

It also enables much pleasurable discussion and debate into what a Learning Organisation really is: a debate which unfortunately falls frequent victim to the *my meme right your meme wrong syndrome*. We can be so stuck in our unconscious traditions, biases and language, so thrown to a condition of being right that we fail to enquire into what is possible.

## SO WHAT IS POSSIBLE?

The science of genetics has revealed the value of genetic diversity as a source of innovation; a richness from which new forms of order can emerge, or indeed can be deliberately sourced. The new science of complex systems is showing that order emerges from a diversity and critical density of inter-relatedness and inter-connectedness (20) (21)

New futures in the human world, new individual perspectives and aspirations, new systems of beliefs and assumptions, new relationships at the inter-personal level, new organisational innovations and inter-organisational relationships likewise emerge from new and different connections between people and ideas. New language, and new listening, can create new futures. Just as genetic diversity is a source of biological wealth so memetic diversity is the source of cultural wealth. It resides in different thinking and is exchanged and grown in different conversation.

Neither memes nor language is unique to the human species. Parrots have their own language as well as the ability to mimic ours. Parallels with human organisations may strike the reader!!

What can differentiate us is our freedom to choose. We do not have, in Dawkins words, to suffer the tyranny of the selfish replicators. We have the capacity, if we choose to exercise it, to see or at least to enquire into the rivulets and lakes of our perceptual landscape and ask how the world around us, and our interaction with it, might change if we choose to operate from a different perspective. When we do we can, perhaps, cease to be victims of our prior experience and conditioning and instead become authors of different futures: futures in which we can enroll others if we create the space for the emergence of a different meaning space; one that offers a rationale, a set of rules, and a relationship for learning.

**This is key.** We have the latent capacity to be, and perform, differently, individually and collectively

The learning organisation can foster memetic diversity if it creates a context, an environment, where thinking and acting differently in pursuit of different futures is acknowledged and enabled. Creating memetically isolated populations is an easy enabler (26)(29) but the resultant tension requires the balance of a context which encourages exchange of ideas between such populations (see 4 for the fundamental theory). The danger, in many organisations, including those that seek to perpetuate the Learning Organisation is that the debate centres on what the Learning Organisation is. Social and political processes ensue that disenfranchise or excommunicate those who do not conform to the party line. *Different people placed in the same system produce the same result* (2). The result is the slow death of learning. Hence we offer our concluding paradox:

**The learning organisation will only be a learning organisation when it is not stuck with discussing what a learning organisation is.**

If instead, in any organisation, the conversation concerns what the organisation will be and do, and how people will relate to each other in the doing much more is possible.

## **POSTSCRIPT**

A danger of departing from established conventions of academic papers is that one fails to convey and acknowledge the depth and breadth of contribution derived from the works of other authors. We include here a brief reference to some of the contributions that have had a major influence on the emergence of our own memes. In earlier contributions to the ECLO series (16) (27) and elsewhere (6) (28) (29) we have tried to provide more complete references. Our pending book “Shifting the Patterns” explores all the issues presented here in more detail.

Dennett’s recent book (10) offers a rigorous review of the emergent field of memetics. Our intellectual synthesis would grant more influence to autopoiesis (23) (24), self-organisation (1) (18) (19) (33) (36) and punctuated equilibrium (11) (15) (29). We would assert that memes explain the well documented influence of traditions on behaviour (6) (12) and the link between being, speaking and results (e.g.6). Many of Morgan’s images of organisation (25) can be explained by the same proposition as can the successes and failures of many approaches to the facilitation of change and learning. Ultimately one does not need an explanation of a fundamental process to utilise its effects. You can drive a car without understanding the

workings of the internal combustion engine. You are less likely to be able to tune it for a different level of performance.

## REFERENCES

- (1) Allen P. M. *Coherence, Chaos and Evolution in The Social Context*. Futures 1994 26:583-597
- (2) Akhlaghi F. *The Critical Role Of People In Delivering Services* ICM London conference address 1995
- (3) Argyris, C. *Reasoning, Learning and Action: Individual and Organisational* . San Francisco, Jossey-Bass 1982
- (4) Arthur, W. B. *Positive feedbacks in the economy*. McKinsey Quarterly Journal 1994 1:81-95
- (5) Axelrod, R. *The Evolution of Co-operation*,. New York, Basic Books Inc. 1984.
- (6) Batchelor, J and Shaw, R. *The Shifting Patterns of Business: Transforming Purchasing for the New Eras* Proceedings of the North American Association of Purchasing Managers Annual Academic Conference 1996, 11-25
- (7) Bateson, G. *Steps Towards an Evolution of the Mind*, London, Paladin, 1973
- (8) Dawkins, R. *The Selfish Gene*, Oxford, Oxford University Press, 1976 and 2nd Ed. 1989
- (9) Dawkins, R. *The Blind Watchmaker*. London, Longmans, 1986
- (10) Dennett, D. C. *Darwin's Dangerous Idea: Evolution and the Meanings of Life*, New York, Simon and Schuster, 1995
- (11) Eldredge, N. and Gould, S. J. *Punctuated Equilibrium: An Alternative to Phyletic Gradualism*, In T. J. M. Schopf (ed) Models in Palaeobiology, San Fransisco, Freeman Cooper and Company 1972 82-115
- (12) Hamel, G. and Prahalad, C. K. *Competing for the Future*, New York, McGraw Hill, 1995
- (13) Flores, F, and Winograd, T. *Computers and Cognition*, Ablex, 1987
- (14) Gell-Mann, M. *The Quark and the Jaguar: Adventures in the Simple and the Complex*, London, Little Brown, 1994
- (15) Gersynk, C. J. Punctuated Equilibria as a Model for Organisational Change *Academy of Management Review* 1991 16:10-23

- (16) Gibbons, B. and Shaw, R. The Learning and Performance Dynamic in T Campbell and V Duperret-Tran (eds) *Second Conference of the E.C.L.O* 1995, 121-128
- (17) Gould, S. J. *Wonderful Life : The Burgess Shale and the Nature of History*, New York, Norton,. 1989
- (18) Hull, D. *Science as a Process*,. Chicago, University of Chicago Press, 1986
- (19) Kauffman S. A. *Origins of Order: Self-Organization and Selection in Evolution* Oxford, Oxford University Press, 1992
- (20) Kauffman S. A. *At Home in the Universe: The Search for The Laws of Complexity*, New York, Oxford University Press, 1995
- (21) Kay, J. *Foundations of Corporate Success: How Business Strategies Add Value*, Oxford, Oxford University Press, 1993
- (22) Lloyd, T. *The Nice Company*. London, Bloomsbury, 1990
- (23) Maturana, H. R. and Varela, F. *Autopoiesis and Cognition: The Realization of the Living*, Dordrecht, Reidel, 1980
- (24) Mingers, J. *Self-Producing Systems: Implications and Applications of Autopoiesis*, London, Plenum Press, 1994.
- (25) Morgan, G. *Images of Organization*, London Sage, (1986)
- (26) Nonaka, I. and Takeuchi H. *The Knowledge Creating Company: How Japanese Companies Create the Dynamics of Innovation*, New York Oxford University Press, 1995
- (27) Price, I. The genetics of the learning organisation, in T Campbell and V Duperret-Tran (eds) *First Conference of the E.C.L.O* 1994, 81-96, E.C.L.O Conference Proceedings 1994
- (28) Price, I. *Organisational Memetics?: Organisational Learning as a Selection Process*, Management Learning 1995, 299-318
- (29) Price, I. and Evans, L. *Punctuated Equilibrium: An Organic Metaphor for the Learning Organisation*, European Forum for Management Development Quarterly Review, 1993, 93(1) 33-35
- (30) Price, I. and Akhlaghi, F. *Emergent Order on the New Frontier: Future evolution of the UK FM Market Place*, UFMR Special Report No 1, in prep.
- (31) Scott-Morgan, P. *The Unwritten Rules of the Game*, New York, McGraw Hill, 1994
- (32) Schapiro, E. *Fad Surfing In The Boardroom*, New York, John Wiley and Sons, 1995

- (33) Stacey, R. D. *The Science of Complexity: An Alternative Perspective for Strategic Change Processes*, Strategic Management Journal 1995 477-495
- (34) Thurbin, P. J. *Implementing the Learning Organisation: The 17-Day Programme*, London, Financial Times & Pitman Publishing, 1994
- (35) Van den Broek, H. Business trends and the learning organisation, in T Campbell and V Duperret-Tran (eds) *First Conference of the E.C.L.O* 1994, 1-16
- (36) Waldrop, M. M. *Complexity: The Emerging Science at the Edge of Order and Chaos*, New York, Simon and Schuster, 1992

## PUNCTUATING ORGANISATIONAL EQUILIBRIUM

### Shifting the patterns that limit

By  
If Price

Active Personal Learning, Guildford UK and Facilities Management Graduate Centre, Sheffield Hallam University

Tel/fax +44-114 253-4032: email ifprice@aol.com & I.price@shu.ac.uk

Text of an article published in the Critical Linkages Newsletter special issue of September 1997

It was the end of a typical one-day management conference in New York. I happened to be in town and dropped in to post conference cocktails to meet up with a friend who had been speaking. As the crowd thinned I fell into conversation with a guy who, at first sight, looked like another stereotyped, cautious executive but who began to speak with passion and enthusiasm of the new mood that was sweeping the company he worked for [a large food manufacturer]. They had, he said with pride, just completed their ninth record quarter in a row; results driven not by lay-offs but by people thinking and behaving differently.

How? I said, having just spent two years living through what was shaping up to be a similar, though not so painless transformation.

I guess said my companion it started with our new CEO. He is Australian. They are different.

Too right sport said I - giving free rein to a long subdued Sydney drawl. What did he do?

Well I guess it started when he blew up the headquarters.

What? You mean literally.

That's right. Ours was a company proud of our traditions. The Head Office was in the original building designed by..... *[the founder]* when he built the first plant. Bruce *[I've forgotten the man's real name]* had it dynamited. He videoed the occasion and sent everyone in the company a copy. They got the message.

At the time I was working for the world's third largest oil company leading what the CEO had called the Process Review project. Part of our role was to be one of the catalysts for a transformation that he had kicked off two years earlier. Part was to examine what practice and theory existed in other organisations and in the consulting firms and business schools from whom they were drawing inspiration. We were coming to see organisational change and learning as dynamic and individualised process. And we were beginning to find metaphors from biology and geology particularly appropriate to capture the essence of the change process. Perhaps it helped that I was once a geologist.

In a nutshell we had concluded that the process of organic evolution in a dynamic ecosystem provides the paradigm with which to understand organisational adaptation, both in general and in the specific case of our own experience. Contrary to popular misconception evolution is not a

process of gradual steady state change but is instead the organic response to punctuation of a normal equilibrium. Similar dynamics exist in organisations. The process of lasting change, however it is achieved is one of interrupting whatever it creates and sustains an organisation's equilibrium. Big bangs in the HQ are one way to do it, literally as in the case above, or metaphorically. They are not the only solution for, contrary to the one best recipe schools of both managers and their advisers, there is no one best way. There may however be one basic process.

## OUR EXPERIENCE

The big bang in the company I then worked for, BP Exploration, was instigated by a series of management conferences and special projects designed quite specifically to raise the level of creative tension in the organisation and to weaken the influence [as well as reduce the size] of various central functions. After a year of raised tension each operating centre in the company was charged with making, over the ensuing three years, a breakthrough in its operating margin. The details do not matter. Suffice it to say that the figure demanded exceeded reasonable expectations from industry benchmarks and was generally held to be impossible. Every operation achieved it in under two years, each in its own way, with the help of whatever advisers it chose to employ.

Our Process Review team lurked in the background as a stick which might be used by the CEO if he had to but without the formal power to influence how any change was achieved. In the event the stick remained unused, at least where the formal operations of the organisation were concerned. Every local management team discovered its own solutions, each of which involved challenging the patterns of thinking and behaviour that had previously prevailed; its own assumptions about its business.

In this respect the company was like few others that we visited during the course of the project. In the others, in case after case, we saw equivalent head-office groups and their chosen advisors involved in designing and mandating, or facilitating, the process of change. In very few such cases were lasting results achieved even when the change processes gathered a momentum and life of their own.

For a year and a half the process was distinctly uncomfortable for those of us in the team. The lack of formal authority to act within a conventional structure was hard to take for a group whose managerial careers had, until then, been developed within conventional frames of reference concerning authority and conventional political structures to get things done and advance one's own career. Only gradually did we realise that a perceived source of weakness when we started was a source of strength which triggered our own innovations in learning from the experience of business units undergoing change and in discovering means to energise that process. As we did so our paradigms shifted from a primary focus on the formal business process of how things are done to an understanding of the need to integrate business and human dynamics if lasting differences are to be made. An organisation is an amalgam not only of hardware and software [physical resources and operating systems] but also of liveware; people and the thinking processes, belief systems, behavioural codes, and conversational habits that shape their day to day actions.

## LIVING COMPANIES?

Since then of course our experience of the failure of many centralised exchange initiatives has been validated many times by research reported in business literature. Yet still the search for the one best way to tune the machine that is the organisation continues. Gradually however, with advances in theory and practice, with insights from the newer business gurus and with progress in applying the emergent science of complex adaptive systems to organisational settings, a new paradigm of organisations is indeed emerging. The world of management is coming to realise that organisations, quite literally, have a life of their own. Like organic species they evolve and adapt, or else go extinct. They exist in networks of interacting entities called economies or ecologies; systems in which individuals occupy particular niches and systems that have, on various time scales, the property of maintaining a quasi-stability for long periods of time between intervening short moments of change.

To the extent that evolution penetrated classical management thinking it was as a metaphor for gradual and even managed change. *Evolution not revolution* is the plea, not least, of managers seeking to control and perhaps limit the pace and impact of change in their particular company. Yet the evidence of geological history is that evolution does not work like that: its tempo, relative to the time available, is stasis or near stasis interspersed with crisis. Evolution is either revolution or it is nothing. The new view of geological evolution is itself only some 20 years old. It started with reappraisals of what geologists call the stratigraphic record, with the evidence from oil exploration wells and particularly with the seismic technology developed in the oil industry since the early 1970s. Without repeating here details that have been amply documented elsewhere suffice it to say that since then we have come to realise that the history of both physical and biological environments on the planet has been, as an early pioneer of the new thinking put *it like the life of a soldier; long periods of boredom and short moments of terror*. Why? And more importantly why [on a different timescale] does organisational evolution show the same dynamic. What is the source of an organization's equilibrium, its stability and even its tendency to seek to preserve that stability even as the world around it changes?

For an organic species the answer is its genetic code; information chemically coded in molecules of DNA which effectively specifies the form of a species, enabling it, or more precisely individual organisms within the species to be better survivors than their rivals, but at the same time limiting the individual to be what is specified by the genes; no more and no less. From the gene's eye view of the world this is a suitable state of affairs for the gene is a replicator, interested without conscious forethought in producing copies of itself. The inherent tendency of genes to replicate is the driving force of evolution. To do so they must build survival machines, organic bodies, that will reproduce in a world populated by competing survival machines encoded by rival gene complexes. These complex systems of interacting genes and the bodies they specify seek the stability of an Evolutionary Stable System or ESS; one in which, within the limits of fluctuating populations, a stable genetic mix is preserved. Most variants are swamped in the larger gene pool and bursts of innovation [evolution's moments of change] happen when either the biologic equilibrium is physically punctuated or when small populations evolve in genetically isolated communities. So do organisations have their equivalent of the genetic code?

We believe they do. Examine any organisation and you will find a series of unwritten but widely followed rules of the game; rules that specify sensible behaviour for those who wish to remain in

and prosper in the organisation. You will find mental models or paradigms, codes of behaviour and patterns of values and beliefs that govern who belongs and who does not. You will find particular assumptions and meanings conveyed and created in the nuances of language and jargon that an organisation employs. You will find traditions that are accepted as part of the way things are around here.

My colleagues and I have stopped trying to make precise distinctions between say a belief system, a set of shared values or theories in use, or paradigms or mental models. It is not clear where the unwritten rules of a particular company become its culture or the way things are round here. We prefer to say that every organisation has a pattern or a mix of shared thinking and language which enables it to operate. Until some such shared pattern emerges, whether it is a strategy, sense of business purpose and objectives, a set of religious beliefs, a shared language, or a prevailing set of theories underpinning an academic discipline no organisation exists. There is a chaos of random interactions upon which some would be leaders seek to impose order by force of personality or rewards. The collective pattern defines the organisation. In the language of complexity theory it is the attractor around which the organisation stabilises.

Conventional views have it that we humans create organisations and perhaps that they in turn create patterns but consider the organisation from the pattern's point of view. By enabling, and in effect specifying a certain form of organisation a pattern seeks, again without conscious forethought, its own replication in the world. It works in precisely the same way as the genotype of an individual organism or the genome of an organic species. It is in fact a replicator, enabling the appearance in the world of the survival machine that will ensure its own replication. In the language of biological ideas now finding their way into philosophy, psychology and organisational theory the pattern is the memetic code of the species.

Just as fundamental change in the organic systems, whether designed or evolved, requires a shift in the prevailing genetic codes so we argue in organisational systems a change requires a shift in the prevailing memetic or mental pattern; the collective meaning set that is transmitted by an organisation's unwritten rules of behaviour, by its language and indeed by its cultural artefacts. Just as religious buildings, whatever their other functions, help preserve the belief set embodied in a particular religion so a corporate edifice, like the one destroyed in the example with which I started the article, helps preserve the underlying pattern of an organisation. To create a change that pattern must be shifted, whether by design or accident. Those who design change programmes that do not shift an existing pattern should not be surprised when the pattern strikes back; for when the chips are down the pattern is not there for the good of the organisation. It is, like The Selfish Gene, there for itself. It operates on the minds of people who carry it to ensure its own survival, as surely as a flu virus operates on their bodies, inducing behaviours selected to enhance a replicator's chances of replicating. Change the pattern, either by changing the context in which it operates [dynamite can do the trick<sup>1</sup>] or by changing the conversation and rules by which it is transmitted and you change the organisation.

## SOURCES AND ACKNOWLEDGMENTS

In reaching the theories put forward here my colleagues and I have drawn upon emerging ideas from many fields of practice and theory. With a colleague Ray Shaw co-author of a forthcoming

book “*Shifting the Patterns: Transforming the codes of personal and organisational performance.*” I try to explore these ideas in more detail and acknowledge the many sources from whom we have drawn inspiration.

Readers wishing an academic summary may find it helpful to refer to my paper *Organisational Memetics: Organisational Learning as a Selection Process*: Management Learning 1995 Vol 26/3 pp299-318. A less academic approach, with updated references is Price. I and Shaw R. *Parrots Patterns and Performance: Emergence of a New Management Replicator* delivered in May 1996 to the annual meeting of ECLO; [The European Consortium for the Learning Organisation] and available from them, and over the www from September this year.

During the Process Review experience I had the unique opportunity of working with Peter Scott-Morgan whose approach to analysing The Unwritten Rules of the Game, was subsequently published by McGraw Hill in 1994. Shifting those rules remains for me one of the essences of pattern shifting.

*Ifrynn [If] Price is a consultant/ educator/ researcher and writer/speaker on exchange. He originally trained and worked as a geologist in many parts of the world. Before creating the Process Review Team described here he was manager of BP's world-wide Exploration and Production Research Division. In April 1993 he left corporate life to co-found The Harrow Partnership Limited; a group devoted to helping people and organisations release the potential trapped in their limiting patterns. If is also a Visiting Fellow of Sheffield Hallam University where he teaches benchmarking and organisation transformation and facilitates several organisational learning forums. With his partner Hazel he also directs Active Personal Learning a company dedicated to helping individuals and smaller business grow through the use of accrediting acquired competencies.*

## THE AUTHORS' REPLY TO PAUL MARSDEN'S BOOK REVIEW: MEMETICS AND THE EDGE OF CHAOS

By

If Price and Ray Shaw

Facilities Management Graduate Centre, Sheffield Hallam University  
and The Extraordinary Performance Company Ltd, Reigate.  
I.Price@shu.ac.uk RayShaw1@compuserve.com

In reviewing Shifting the Patterns (STP) Marsden identifies us as taking memetics to the edge of chaos; a comment we welcome if only because our intellectual underpinning for the book draws on that other emergent science - complexity - as much as it does on memetics and standard neo-Darwinian evolution. Unlike say Dennett (1995) who sees in complexity and Kauffman's (1993) ideas a search for a skyhook, we see complexity and a general science of selfish replicators as necessary complements to one another in the search for the understanding of the origin and self-maintenance of complex order (Price, forthcoming). Complexity offers the 'edge of chaos', as the setting for the emergence of new order and we make no apology for also offering it as the metaphor for the release of individual and organisational possibility.

Marsden is quite correct that we attempt to use what he has aptly termed "*the memetic stance*" (Marsden, 1998) to prompt solutions to real-world issues rather than write a book on memetics. STP emerged as our manifesto for the release of human potential. Its conception involved the dilemma of our desire to both produce new scholastic ideas and yet to write outside the standard patterns of 'management books'; patterns which tend to lump the market into the categories of 'scholastic' or 'how-to'. The memetic stance seems to us to be indeed an "*ontologically minimalist heuristic*" (*ibid.*); one capable of unifying a wealth of broadly evolutionary explanations of human phenomena at all scales from the individual to the 'inter-organisational'. Though we wrote mainly in the context of companies we see the message as pertaining to all forms of organisation and are delighted that, within a month of STP reaching the world we had favourable feedback from readers concerned with such patterns as their marriages, drug addiction and better support for children at risk as well as from a more conventional management audience.

However when Marsden argues that "*rather than take the conventional view that individuals are the hosts for memes which are transmitted through individual behaviour, STP argues that companies should be understood as the proper hosts to memes*" we must take issue and urge a more inclusive perspective. It seems to us questionable to label that the conventional view (see for example Hull 1988, Lloyd 1990) but more importantly we would argue that forms of 'company' more complex than an individual are an emergent property of shared memes, transmitted through language and cultural artefacts. They do then indeed become vehicles for the replication of those memes.

Marsden goes on to cite the problems of our:

*"Unorthodox understanding of the relationship between meme and organisation ... particularly with respect to the organism/company (dis)analogy. Organisations are not physically bonded,*

*rather relations between their constituent parts are informational and not physical and thus not confined to predetermined paths. Organisations themselves do not reproduce whereas organisms do. Thirdly unlike organisms, organisations have phylogenies. Fourthly, and perhaps most critically, companies do not have any needs themselves, unlike individuals, so it is difficult to see how evolution can operate at the level of the organisation."*

To deal, briefly, with each point, first extended phenotypes (Dawkins, 1982) are not physically bonded neither are, say, symbiotic organisms co-evolving, yet replicator theory explains the resultant complex order. Second organisations, or at least organisational strategies and processes, are very definitely reproduced through imitation (Lloyd, 1990). Third McCarthy et. al. (1997) have demonstrated organisational phylogeny and introduced the prospect of the twin sciences of organisational memetics (Price, 1994; 1995) and organisational cladistics. Finally, while companies may start as vehicles to deliver individual needs we would argue they become institutions with, in a sense at least, their own - or their memes' - needs. Who in an organisation has not heard the expression 'what this organisation needs is .....'

We are then accused of exacerbating problems in the partial gene/meme analogy by conflating 'pattern' and 'meme'. True perhaps but we in good company. Dawkins (1976) makes very clear his use of 'The Selfish Gene' as a metaphor for the interaction of several poorly defined bits of a chromosome interacting in complex chromosome system. We tried to make the same point clear (p159/160). The science of genetics has had a century's start on memetics and manages to embrace both generalist and specific concepts of gene. We would agree the need for a memetics to evolve its own specialist terminology but suggest the generalist conflation of meme, like gene, benefits the wider appreciation of the memetic stance. Research memeticists will undoubtedly evolve their own special memes if only because replicators abhor a vacuum in replication space. The process will, inevitably, be a selection process between variants on the meme of meme and its derivatives. We may for example find ourselves requiring at least two distinct meanings of 'host', at the level of mind and of organisation. But then where is the host of the gene? Is it the species, the body, the cell, the nucleus?

We also admit to freely switching between gene and virus analogies (we challenge their looseness). 'Double-helix' replicator theory suffers a similar confusion. DNA (gene) based organisms are also hosts to RNA based viruses. If memetics is restricted to the viral level then science needs another word for the wider organisational replicator. The alternative is to keep 'meme' general and derive more specific terminology where it is needed. "*Confusion over the unit and level of evolution*" is not unique to memetics. "*It is not at all clear whether rules evolve, organisations evolve or more amorphously, patterns (of rules behaviours and organisations evolve).*" Memes, we argue, evolve through the emergent web of institutional rules, language and artefacts that they enable.

Marsden continues:

*"If memes are taken to be socially acquired cultural instructions then the cogency of STP largely holds together, and the framework presented is somewhat similar to Cloak's cultural ethology of i-culture and m-culture (Cloak 1975). Indeed the authors might have usefully exploited the clarity of this framework, but instead no reference is made to Cloak. On the other hand, by*

*adopting Dawkins' meme concept uncritically, the authors also add the ambiguity inherent in Dawkins various conceptualisations (see Gatherer 1998) to their own framework."*

We agree and were, at the time of writing unaware of Cloak's distinctions. At the level of scientific objectivity those distinctions are indeed important. In STP we were more concerned with utility and the 'tyranny of selfish replicators'.

We were likewise unaware, at the time of writing of "*Allison's (1992, 1993) work on the spread of non utility-maximising rules through populations* or Burns and Dietz's institution theory (1992a, 1992b, 1996, 1997a, 1997b). We are pleased to agree their relevance and welcome yet another body of organisational observation amenable to the memetic stance or the selfish meme heuristic.

We do though see the 'long periods of stability - short moments of terror' (Ager 1973) dynamic of the evolution of the complex geological system on which we all live as a fundamental property of replicator driven systems, a conclusion we sought to substantiate by reference to observations at various scales of existence from the geological to the corporate (see also Price and Kennie, 1997). "*Reports of the death of gradualism have been greatly exaggerated, and Gould is*" - indeed - "*not the only evolutionist around*". Unfortunately some of the others seem to us so preoccupied with how evolution ought to operate that they are less concerned with the evidence of how it has operated (patterns / memes again!). The geological - biological debate over places at the high table of evolution may yet mirror itself in memetics.

Then we reach "*the most controversial aspect of STP for memeticists is the idea that the self, and importantly consciousness, are not memetic, and that a non-memetic independent freewill somehow allows individuals in organisations to act with true intentionality. In other words, the dualist homunculus is left untouched by the authors' conceptualisation of memetics*". We stayed out of that one because of concern for utility rather than 'truth'. Too far down that line of reasoning, even if it is neurophysiologically 'true' lies the danger of abrogation of personal or social moral and ethical responsibility. We do not however see a need for a dualist homunculus to evoke a mind capable of inquiring into its own existence and own choices. There is a fully self contained explanation. Memetic patterns which permit an inquiry into other memetic patterns residing in the same brain, grant themselves a certain trick in replication space. Some memes have always granted the minds that carried them some power of inquiry into underlying assumptions. Sometimes they got their host excommunicated / burnt at the stake / made redundant or otherwise disposed of but they nonetheless replicated and also, on balance, enabled the progress of cultural evolution. The memes of inquiry live on because they also enable their hosts (at all levels of organisation) to achieve greater results and are an antidote to other, more parasitic, memes. We see much power in Blackmore's (1996) conception of meme-eating memes. In deploying them, people can at least influence the evolution of companies, through both selective transmission and also the creation of new memes derived from a new understanding of the patterns that code for performance. While the school of meme theory, "*concerned with deconstructing pervasive homuncularism understanding human behaviour*" has undoubtedly brought new insights and explanations of human behaviour we would return to Dawkins' original conception of the meme as an example of the power of replicators to enable and maintain complex order. We hold therefore that there is a science of organisational memetics; one that has the potential to finally provide an explanation for the existence of

organisations, by recasting in one simple heuristic numerous other observational stances and perhaps even one capable of giving people the dualist homunculus that they have not previously had. Now there's a thought!

## References

- Ager, D.V. (1973) *The Nature of the Stratigraphical Record*, London: Macmillan.
- Allison, P. (1992) Cultural relatedness under oblique and horizontal transmission rules. *Ethology and Sociobiology*, **13**:153-169.
- Allison, P. (1993) The Cultural Evolution of Beneficent Norms. *Social Forces*, **71**:279-301.
- Blackmore, S. (1996) The Psychology of Awakening, International Conference on Buddhism, Science and Psychotherapy, Dartington. <http://www.memes.org.uk/meme-lab/DART96.htm>
- Burns, T.R. and Dietz, T. (1992a) Cultural Evolution: Social Rule Systems, Selection, and Human Agency. *International Sociology*, **7**: 259-283.
- Burns, T.R. and Dietz, T. (1992b) Human Agency and the Evolutionary Dynamics of Culture. *Acta Sociologica*, **35**:187-200.
- Burns, T.R. and Dietz, T. (1996). Emerging Order and Disorder: Modernization and Globalization in an Evolutionary Perspective. Report, Uppsala Theory Circle, Uppsala University, Uppsala, Sweden.
- Burns, T.R. and Dietz, T. (1997a). Evolutionary Sociology: Selective Environments, Human Agency and Institutional Dynamics. Conference on Sociological Theory and the Environment, Zeist, The Netherlands, March 20-23, 1997 (Research Committee 24, Environment and Society, the International Sociological Association).
- Burns, T.R. and Dietz, T. (1997b). Human Agency and Evolutionary Processes: Institutional Dynamics and Social Revolution. In: B. Wittrock (ed.), *Agency in Social Theory*. London: Sage Publications.
- Cloak, F.T. (1975). Is a cultural ethology possible? *Human Ecology*, **3**:161-182.
- Dawkins, R. (1976,). *The Selfish Gene*. Oxford: Oxford University Press.
- Dawkins, R. (1982).*The Extended Phenotype*. Oxford: Oxford University Press.
- Dennett, D.C. (1995). *Darwin's Dangerous Idea: Evolution and the Meanings of Life*, New York: Simon and Shuster.
- Hull, D. (1988). *Science as a Process*. Chicago: University of Chicago Press.

Kauffman, S.A. (1993). *The Origins of Order: Self Organization and Selection in Evolution*, New York: Oxford University Press.

Lloyd, T. (1990). *The Nice Company*, London: Bloomsbury

Price, I. (1994). The genetics of the learning organisation. In P. DePotter and T.L. Campbell (eds.), *Proceedings of the First Conference of the European Consortium for the Learning Organisation*, London: Gower.

Price, I. (1995). Organisational memetics?: organisational learning as a selection process. *Management Learning*, 299-318. <http://members.aol.com/ifprice/ifresch.orgmem.html>

Price, I. (forthcoming). Images or Reality? Metaphors, Memes and Management. In M. Lissack and H. Gunz (eds.) *Managing the Complex*, Westport: Quorum Press.

Price, I. and Kennie, T.R.M. (1997). Punctuated Strategic Equilibrium and Some Leadership Challenges for University 2000. 2<sup>nd</sup> International Dynamics of Strategy Conference, SEMS, Guildford. <http://members.aol.com/ifprice/pkacstrat.html>

Price, I. and Shaw, R. (1998). *Shifting the Patterns: breaching the memetic codes of corporate performance*. Chalford: Management Books 2000. <http://members.aol.com/ifprice/stp.html>

Marsden, P. (1998). Memetics and social contagion: Two sides of the same coin? *Journal of Memetics - Evolutionary Models of Information Transmission*, 2. [http://jom-emit.cfpm.org/1998/vol2/marsden\\_p.html](http://jom-emit.cfpm.org/1998/vol2/marsden_p.html).

McCarthy, I.P., Lessure, M., Ridgeway, K. and Fielder. M. (1997). Building a manufacturing cladogram, *International Journal of Technology Management*, 1:269

# STEPS TOWARD THE MEMETIC SELF

## A commentary on Rose's paper: Controversies in Meme Theory

By  
If Price

Facilities Management Graduate Centre  
Sheffield Hallam University  
I.Price@shu.ac.uk

### 1 Introduction

As the editor of JoM-EMIT points out, some authors in memetics have suggested (explicitly or implicitly) that 'consciousness' or a 'Self' beyond the constructs of memes and genes can choose or design memes (most famously Dawkins' (1976) call to 'overthrow the tyranny' of the replicators). The alternative position eloquently argued by Rose holds that attributing variation and selection of memes to some form of conscious self beyond the reach of the meme "*not only undermines the value of meme theory as an evolutionary process, but is also a kind of giving up. If we intentionally design and choose memes 'why do we need an evolutionary theory of culture at all?*".

This essay is written as an invited contribution to the debate, one I am glad to make if only because Ray Shaw and I (1998) have written what one reviewer (Marsden, 1998b) termed a 'manifesto for a velvet revolution' against the tyranny of selfish replicators and dropped easily into the language of individual, goal-directed, choice without addressing the question of whether such a self is, in point of scientific fact, a logical impossibility. Here I may redress the balance. The effort comes in three sections. First I seek to clarify the meaning of meme being used. I then ask whether it is necessary to invoke some conscious self beyond memes, some residual dualist homunculus, if individuals, and groups, are to free themselves of the tyranny of the selfish replicators. In these two sections I will seek to hold to the conventional scientific interpretation of truth. Are we arriving at a better explanation for the working of the natural world and human social systems within that world? In the third section, and the third section only I will address the limits of that particular meme of truth.

### 2 What meme of meme?

Price and Shaw (1996) used memes drawn from organisational theory and practice to seek an illustration of what may be a typical life cycle of memes in other spheres. To create some replication niche for themselves memes, when they first evolve, must enable some form of result, confer some benefit for those who are their first host. Whether this is a universal need may be open to debate but enabling is a useful strategy for replicating. At some point however 'successful' memes, where success is judged purely in terms of replication, need not confer any benefit, provided only that they spread. Business fads provide numerous examples.

From the meme's perspective it is not necessary that replication has a high fidelity. Many business fads (and other memes) become labels whose carriers try and attach different meanings to them. The 'meme' meme is undoubtedly currently enjoying a considerable radiation, one accompanied by a wide, and growing, divergence of meanings attached to it (Edmonds, 1998).

Such divergence is probably another inherent facet of memetic replication. Compared to genes, memes have traded replicative fidelity for replicative speed. From the selfish meme's perspective, so what? Success, for the 'meme' meme is purely the number of individual human minds who now store some form of meaning associated with the word 'meme'. If it causes confusion then so be it. At least the 'meme' meme is replicated.

The diversity of meaning does though limit carriers of the 'meme' meme when they try to inquire into the power of memes and the limits, or otherwise, of memetics as a theory of either the individual mind or of various forms of social organisation involving more than one mind as members. Rose (op. cit.) is again lucid on the confusion which has arisen from Dawkins' original use of cultural artefacts, Cloak's (1975) 'm-culture'; rather than the cultural instructions inside the brain which produced these artefacts Cloak's i-culture to exemplify memes. In trying to arrive at a theory of organisational memetics (Price 1995) I have taken the same interpretation as Rose and conceived of memes as ideas or information, in essence the i-culture. Memes, under this interpretation, may be replicated not only via cultural artefacts but also through language (c.f. Cavalli-sforza and Feldman, 1994) and through the implicit unwritten rules found in any form of social organisation. My own memes in this area derive from Hull's (1988) interpretation of the unwritten rules of the scientific process as vehicles for maintaining particular paradigms and Scott-Morgan's (1994) work on unwritten rules.

The confusion should not be allowed to detract from Dawkins' original point, the power of replicators to enable the emergence, and maintenance over time, of forms of order which are out of thermodynamic equilibrium with their environment. Darwin's 'Dangerous Idea' (Dennett 1995) seems to offer not only the one truly sustainable 'scientific' explanation we possess for the why the biosphere is the biosphere but also the potential for an equivalent explanation of all forms of organisation. The combination of memetics and complexity (Gabora, 1997; Price, 1995; 1999) offers rich potential as we examine not individual memes passed on through imitation, but rich and densely interconnected memetic patterns. An ironical property of complex adaptive systems (CAS) is their tendency to seek and maintain evolutionary stable states. We need to consider, as forms of memetic phenotype, not just individual 'viruses of the mind' but also all forms of social organisation (social CAS). Any individual mind is a product not only of particular individual memes but also of the wider social memes too which it has been exposed. Hence for example I view my mind as carrying a memetic pattern imprinted (among other things) by the English language and tradition (with overtones of the Australian sub-species), by early training as a geologist, and by a career spent within managed commercial organisations. One individual mind can be a member of many different memetic organisations. Can that mind also have any form of free-will, any ability to choose between the various memetic strains that have made it what it is?

### 3 The memetic self

A more generalised theory of replicators applied to the complex system that is the human brain / mind would hold any one brain / mind combination to be the product of some mix of our genetic and memetic inheritance. Where the boundary is crossed, if indeed it is, is yet another of the unanswered questions of memetics. As I understand it one school of cognitive evolutionary psychology (e.g. Pinker 1997) seeks to argue a genetically specified capability for perception,

image coding, memory, language and information processing as capable of yielding a conscious mind without recourse to memes. I prefer Dennett's view of a genetically coded brain converted to a mind when it becomes the resting place for memes. Price and Shaw (1998) *exapted* a De Bono metaphor in the image of a landscape, eroded over time to provide streams, rivulets, and rivers interspersed between higher plateau as a simple example of a self-organising, locked-in, system. As rain falls on a virgin landscape so it tends to find the paths of least resistance: the soft rocks and minor depressions of the undulating territory. Over time accumulations of rainfall carve out stream and river beds and settle into pools and lakes. Any new rainfall will no longer find its own way but will rather take, and reinforce, the routes which already exist.

Just as the rainfall follows established routes, so perception follows established ways of *seeing*. Technically, even if the light sources which perturb the back of the retina are identical, what will be noticed from all that could be seen, will depend on the perceptual lens through which someone views the world. The optimist's half-full glass is the pessimist's half-empty one! What is 'there' is not independent of the viewer. What is there is what we have been *trained* (or conditioned or have learnt) to see. Our training in terms of our mental models means that we will not see certain other things which do not fit with the memes we carry. We may discard, indeed we can be blind to, anomalies that do not fit. The self-organised pattern which we call our thinking grants a particular perceptual blindness and rigidity to our perceptions of the world - the very foundation of such things as stereotypes and prejudices - common to all human experience and found, for example in the way one department in a company may view another.

What holds for light waves perturbing the retina, holds equally for acoustic perturbances of the eardrum. Exploring the analogy further we could say that an idea, a single thought, an utterance, a meme in fact, is like the single raindrop. It falls with others upon a pre-formed perceptual landscape. Isolated thoughts gather together in a string - a pattern of co-existing memes - which we might compare to a few drops congregating together in a splash of water. With sufficient mass the splash of water starts to flow into streams and rivers which are, if we like, the connectors between the raindrops and the pools and lakes, if not the oceans, of our thoughts. The pools and lakes we may view as concept pools and theory lakes. Thus a self-organising system is inherited and developed in which the flow of perception takes a certain course; it follows a certain pattern, a largely given paradigm.

Patterns in the brain influence *seeing* (or more accurately perceiving). Patterns, and seeing influence behaviour so that behaviour follows certain patterns. It may be argued that we see the world less as it is and more as we are, and that we act perfectly consistent with how we see the world. There is a certain alignment with our thinking, our perceptions and our actions-in-the-world. Thinking, seeing, and behaving tend to follow pre-existing patterns.

Yet that pattern carries some ability to inquire into, and make sense of, the environment around it. An inquisitive ability, albeit a limited one, seems to be one of the evolutionary tricks of our species biological endowment. Genetic selection hit (for reasons which may be fortuitous - e.g. see Morgan 1997) on a combination of a large brain, an enhanced vocal capacity, an opposed thumb and an upright posture. Whether the experiment would have succeeded if it had not provided the environment for accelerated, memetic, evolution is a moot point. Our hominid ancestors had to live in a world where many faster and physically stronger predators must have

regarded them as an alternative lunch. Tool making, communication, and a faster rate of learning from each other and their environment proved advantageous evolutionary adaptations. Equally a tendency to accept, without question, the mores, beliefs and behaviours which granted any individual membership of a particular tribe, or other social group conveyed a different form of selective advantage. There was safety in numbers. One's own replicators were less likely to end up as some other replicators' lunch.

The inquiring (within limits) mind and the socially acquiescing (within limits) mind is thus perfectly explicable as a natural evolutionary development; a product of Darwinian selection. That mind enabled the species which carried it to outgrow the limits of their genetic inheritance, and periodically to challenge the limits of their memetic inheritance. Such challenges did not always come easily. The history of the development of ideas is replete with examples of the painful ends which awaited those who challenged prevailing epistemological orthodoxies. While Dawkins phraseology, 'escaping the tyranny of the selfish replicators' may be accused of a certain literary excess (a useful memetic trick) it is not necessary to invoke a self out with the meme-gene combination to explain self-awareness or questioning of a particular status quo. It may not even be the combination! An intriguing possibility is that curiosity, the inquiring brain, is an inherent part of our genetic inheritance whilst conformity is memetically induced. It is probably memes more than genes that limit our ability to question a prevailing reality.

Memes for common technologies, common languages and common belief systems sustained and enabled various forms of social organisation from, say, early foraging / hunting groups to any of the world's major nations or religions. If one takes this memetic stance to social organisation it becomes logically difficult to interpret history as other than the record of memetic evolution; an observation which, in turn, offers a perspective of the extent of the tyranny of the selfish replicators. Hobbes' famous characterisation of the life of man as "solitary, poor, nasty, brutish and short" reflects the fate of most members of the memetic species, living subsistence existences with any surplus devoted to the perpetuation of the prevailing memes and the power / caste structure which carried them. In the process most members of the human species became, in large measure, victims of the meme set they found themselves members of. Too great a challenge to an existing status quo has always carried the risk of retaliation by the 'powers that be', that is the agents of the 'memes that be'. The inquiring mind risked the loss of anything from membership of a particular group to life.

Memetics seems to me to offer the prospect of an explanation of the inherent danger of runaway ideological tyrannies, and in and of itself it is thus an antidote to them. Science is itself a memetic selection process (Hull, 1988), and one not free of its own tendency to seek the largest share it can of available resources for its own replication. It is certainly subject to parasitism by malignant and socially repulsive memes like eugenics. It has though enabled a position where we have, in a way that none of our ancestors did, a more accurate understanding of the world around us and our place in it. Perhaps it is only through the memes / genes for the inquiring mind and an appreciation of the power of selfish replicators that we may at last escape at least some of the tyranny.

## **4 Beyond the science meme**

I have tried to write, in the previous two sections, from a perspective of scientific truth. I am inculcated by the science meme and would normally defend the utility of that meme in the pursuit of that form of truth, at least in most contexts.

Switch the context though and even the memes of scientific truth can become limiting. The unchallenged assumption that there must be a single 'truth' or 'right way' becomes a means of defending a particular world view or of creating a particular, self-fulfilling, reality. Assertions with the generic format of declaring something to be true can hold those who make them stuck in a certain reality, parasitised by a certain mind set, victims of a particular memetic pattern. The phenomenon works at scales from the individual to the societal. More can sometimes be gained by asking not 'is a particular perspective true', but 'is that perspective valid', and 'are there alternatives which would enable me to achieve a different result?' To a believer in memes it is difficult to see any particular religion for example as true, yet one can perhaps see many religions as valid, if only because, by providing a rationale for desirable codes of social behaviour they enable certain common value systems. Achieving even a semblance of commonly accepted ethical standards may be harder in a world without religion. This is not to say that there are not many circumstances in which religions limit the minds of those infected by them, but merely to allow an inquiry into the possibility of contexts, in which religion might enable, not limit.

The assertion of truth contains also a dark side. An ever present danger in the logic of saying we are all creatures of our memes is the denial of individual responsibility for ones actions, so any action, or even 'any action not strictly prohibited' one individual takes to exploit others is excused. Whether greed and ownership are memetic or genetic is a moot point but truth can be harnessed in the service of the unbridled growth of either.

Because memes may be scientifically true - as the explanation for the complex system that is the human mind - it does not seem, to me at least, to deny that mind the property of inquiring into itself. Indeed, finally abandoning any self outside memes and genes seems to demand more, not less, self awareness. The memetic self either abrogates responsibility for its behaviour 'It wasn't me - I'm just a vehicle' or accepts that the last chance to escape responsibility for ones actions in the world has past. Memetics meets morality.

## **Notes**

1. It is not strictly relevant to the discussion but it may be that this process should be modelled in terms of 'hitchhiker' memes - different subsets of meaning that try and get a free ride in replication space by attaching themselves to an existing terminology. Our 1996 paper used the example of the 'Learning Organisation'.

## **References**

Dawkins, R., (1976). *The Selfish Gene*. Oxford: Oxford University Press. (2nd Edition, 1989).

- Dennett, D. C., (1991). *Consciousness Explained*. New York: Little, Brown & Co.
- Dennett, D. C., (1995). *Darwin's Dangerous Idea*. London: Penguin.
- Edmonds, B., (1998). On Modelling in Memetics. *Journal of Memetics - Evolutionary Models of Information Transmission*, **2**. [http://jom-emit.cfpm.org/1998/vol2/edmonds\\_b.html](http://jom-emit.cfpm.org/1998/vol2/edmonds_b.html)
- Gabora, L., (1997). The origin and evolution of culture and creativity. *Journal of Memetics - Evolutionary Models of Information Transmission*, **1**, [http://jom-emit.cfpm.org/1997/vol1/gabora\\_1.html](http://jom-emit.cfpm.org/1997/vol1/gabora_1.html)
- Hull, D. L., (1988). *Science as a Process*. Chicago: Chicago University Press
- Marsden, P. (1998a). Memetics and Social Contagion: Two Sides of the Same Coin? *Journal of Memetics - Evolutionary Models of Information Transmission*, **2**. [http://jom-emit.cfpm.org/1998/vol2/marsden\\_p.html](http://jom-emit.cfpm.org/1998/vol2/marsden_p.html)
- Marsden, P. (1998b). A Review of: Shifting the Patterns: Breaching the Memetic Codes of Corporate Performance by If Price and Ray Shaw. *Journal of Memetics - Evolutionary Models of Information Transmission*, **2**. [http://jom-emit.cfpm.org/1998/vol2/marsden\\_p2.html](http://jom-emit.cfpm.org/1998/vol2/marsden_p2.html)
- Morgan, E., (1997). *The Aquatic Ape Hypothesis*. London: Souvenir Press
- Pinker, S., (1997). *How the Mind Works*. New York: W. W. Norton
- Price, I., (1995). Organisational memetics: organisational learning as a selection process. *Management Learning*, 299-318.
- Price, I. and Shaw, R., (1996). Parrots, patterns and performance [The learning organisation meme: emergence of a new management replicator]. In T.L. Campbell (ed.) Proceedings of the Third Conference of the European Consortium for the Learning Organisation, Copenhagen.
- Price, I. and Shaw, R., (1998). *Shifting The Patterns: Breaching the memetic codes of corporate performance*. Chalford: Management Books 2000.
- Price, I., (forthcoming). Images or Reality? Metaphors, Memes and Management. In M. Lissack and H. Gunz Eds. *Managing the Complex*, Westport: Quorum Press,
- Scott-Morgan, P., (1994). *The Unwritten Rules of the Game*. New York: McGraw Hill.

# **RECONSTRUCTION OF ORGANIZATIONAL PHYLOGENY FROM MEMETIC SIMILARITY ANALYSIS: PROOF OF FEASIBILITY**

**By**

**If Price and Andrew Lord**

Facilities Management Graduate Centre,

Sheffield Hallam University, Unit 7 Science Park, Sheffield, S1 1WB, UK.

## **Abstract**

A successful phylogenetic reconstruction of the known pattern of descent of the main post-reformation Christian Churches has been achieved from a computerised analysis of aspects of their present day memetic pattern. The result confirms the feasibility of a new approach to organisational memetics and conducts a first empirical test of the hypothesis that, if organisations are construed as evolving, those with a common ancestor should show greater replicator similarity than more distant relatives.

## **1 Introduction**

Memetics has debated at length the internalist versus externalist positions and the question of whether the individual or organisation should be considered as the memetic 'phenotype'. We approach the topic from an organisational stance (c.f. Price, 1995; Gell-Mann, 1996; Price and Shaw, 1998; Williams, 2000). In brief the stance holds that organisations evolve (Aldrich, 1999) and are constructed or enabled by shared patterns (Schien, 1985; Price and Shaw, 1998), paradigms (Hull, 1988) or schemata (Gell-Mann, 1996). Whether conceived as pattern, paradigm or schemata the proposition would be that they are replicated in the organisations concerned. As noted by Marsden (2000), the stance could be argued to recapitulate an older evolutionary tradition in social sciences, and indeed economics, however the organisational replicator, whatever it is, can be termed meme by reason of the specific origin of the term in the context of a general theory of complex organisation enabled by replicators.

Wilkins recently reminded the memetics discussion list of his earlier (Wilkins, 1998) suggestion that evolutionary cladograms provide a potential means of contrasting the memetic codes of related organisations [note1]:

One such technique that could be used is the Wagner Similarity method (cf *Wiley and others 1991*). This method establishes an instance matrix of characters (presence and absence, but it could be any value  $0 > x > 1$ ) and calculates the sum of the modulus of differences to give the Wagner distance between taxa (or theories). From this a net diagram of relationships, or an unrooted tree, can be drawn up to give a sharp notion of the overall similarity of taxa or theories. The same data can be recast as a rooted cladogram if a sister-group can be selected. Software, both commercial and public domain, is available to perform these analyses and others using well-established algorithms. While this will not definitively establish ancestry of a conceptual lineage, it can be used to test hypotheses of ancestry, and to overcome the Whiggish tendency of

historians to read preferred modern views back into a historical subject, such as Darwin's reliance or not on now-discredited views such as embryological recapitulation.

This research note [note 2] is written to report what we believe to be the first successful explicit test of such an approach. Rather than a Wagner similarity we have used Hamming distance [note 3] (a metric of digital error) and the UPGMA (Unweighted Pair Group Method with Arithmetic Mean) clustering strategy (Sneath & Sokal, 1973; Lapointe and Legendre, 1994) on putative meme strips [note 4] rendered as binary data; i.e. a 'taxon' has or does not have a particular character. We have previously explained (Price and Lord, 2000) the principle behind the method, an examination of further isomorphism between biotic and abiotic complex systems, the reasons for the choice of a test group; and the evidence of an early instance of a successful iteration between inferred and historical pattern of descent and an analysis of memetic similarity. In brief, religious bodies, particularly post-reformation Christian Groups were selected as a test population because they displayed:

- Wide ranging diversity within the group
- Distinguishable characteristics
- Prominent belief systems
- Sufficient historical records
- Periodic evolutionary blooms [note 5]

Comparison with biological evolution leads to the hypothesis that a phylogenogram constructed from an analysis of memetic similarity should reconstruct the known pattern of historical descent. The attempt raises issues of terminology. Hull (1988) has used the schism in systematics between cladistics and phenetics as almost his type example of science as a process of selection between competing paradigms (or meme complexes) each of which replicates via the forms of organisation which it enables. Organisational memetics is arguably no more than a generalisation of Hull's proposition (e.g. Price, 1995; Price and Shaw, 1998). Cladists have criticised pheneticists as lacking objective rigour or the logic of seeking to unravel a pattern of evolutionary descent. Some at least acknowledge that the advent of molecular phenetics has, at best, blurred the distinction and at least lead to a resurgence in phenetics. Here we are borrowing from systematics but seeking to test memetics rather than reconstruct descent.

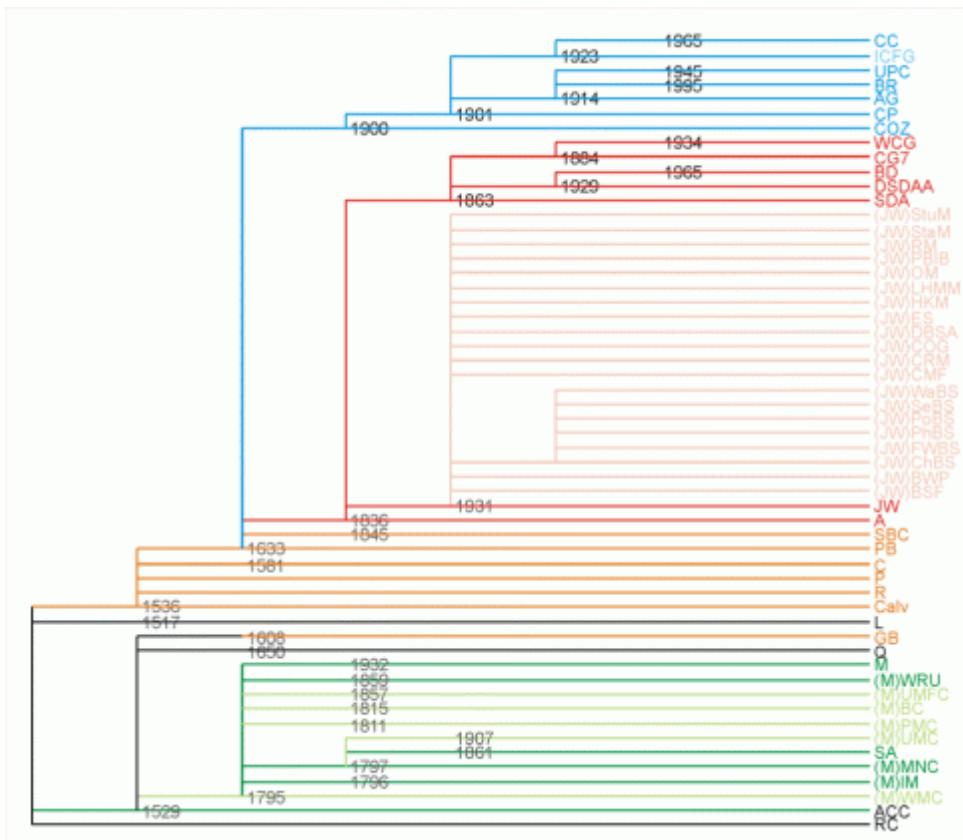
In strict terms we are attempting a phenetic analysis of memetic data however, as with molecular phenetics, we are suggesting that the phylogenogram will reveal a history of descent: i.e. will produce the same output as a cladogram. In the domain under investigation the true history of descent is, within certain bounds (see below), established. The objective is not to classify with a view to unravelling history, it is to test the proposition that the organisations examined might be memetically encoded by comparing what is in effect an inferred cladogram (albeit one derived from phenetic methods) with one known from historical sources. The test, if successful, provides evidence compatible with the inferred presence of a cultural replicator in evolving organisations. If memetics becomes (as we believe it should [note 6]) to studies of evolution in the social sciences broadly what genetics has become to studies of biological evolution, then it may develop its own language and terminology. To those whose perspective on evolution is more geological than biological the selection competition between phenetics and cladistics seems to generate more heat than light. The distinctions may prove unnecessary, or even unhelpful for memetics. Such debates lie in the future. Our concern is to see whether a particular analytical

approach is doable, and if so, does it hold out a promise of utility in the field of organisational memetics.

## 2 Method

### 2.1 Data

The meme strip shown in Table 1 (appended) was gathered from Internet and other written sources [note 7]. We will leave to others the behaviouralist versus internalist debate and see Table 1 as a list of beliefs and practices that between them characterise and underpin different Christian Sects: beliefs and practices that seem to be replicated in the minds of adherents to a particular denomination and that, in being replicated sustain that denomination. We have applied them to the taxa set [note 8] listed in Table 2 (appended), and, where possible, to individual sub-sects, though here the differences became hard to distinguish. The historical derivation of these branches is illustrated in Figure 1. Taxa in lighter colours are not subsequently analysed. They represent schisms within particular religions, for example a variety of sub groups of Jehovah's witnesses, where we have not yet been able to confirm apparently minor differences in belief and practice [note 9].



**Figure 1:** Historically derived phylogeny

## **2.2 Software**

A software routine [note 10] has been written to analyse the progressive pairwise similarity of strips of binary data using, as noted above, the UPGMA on Hamming distances. The taxa character matrix is searched for pairs with the least distance between them. The average of such a pair then becomes a new 'parent taxa' which is subjected to future iterations. Fuller details of the validation will again be described in a longer paper but in brief it successfully reconstructs hypothetical, randomly generated data sets, and also, somewhat to our surprise approximates a biological phylogenetic tree from a set of claimed therapeutic properties of essential oils (Lawless, 1992) [note 11].

## **2.3 Results**

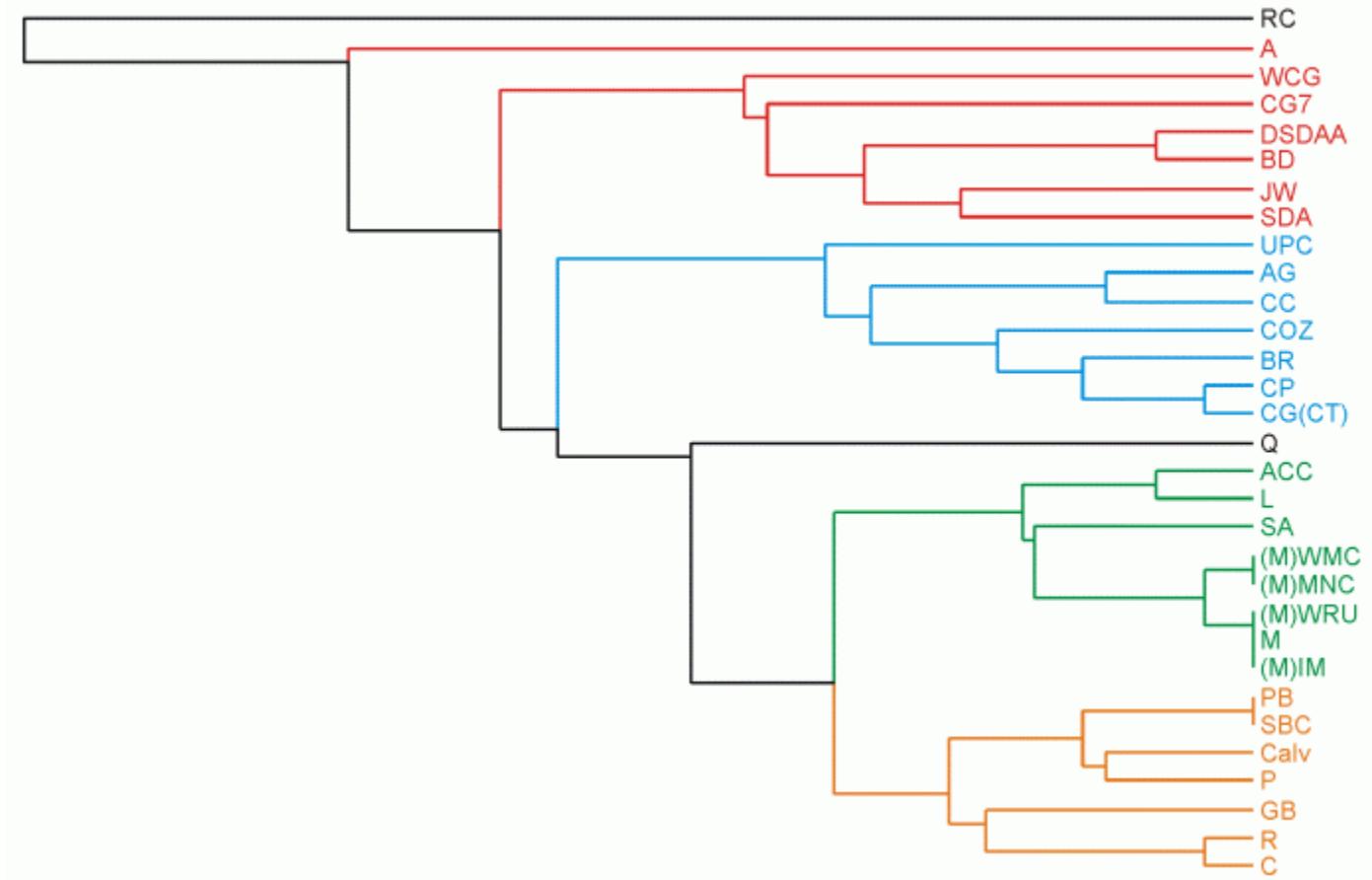
Initial results of the tests on our hypothetical religious clades [note 12] (Figure 2) were encouraging, in that clusters of obvious common origin were revealed. However higher order relationships failed to reconstruct accurately: the derived tree did not match the one drawn from historical sources. In particular clusters of Calvinist denominations showed greater dissimilarity between themselves than one such cluster did to major Episcopalian branches. This gave rise to the suspicion that 'some memes may be more powerful than others'; i.e. that the presence of a certain meme in a particular denomination or group of denominations might inhibit its further splitting [note 13]. In particular the non-episcopalian denominations showed greater variety than the did the episcopalians. In essence the 'Episcopal meme' and the resulting presence of a hierarchical bishopric arguably restricted the possibility of further evolution.

The initial test had given the presence of such authority the same weight (one unit of a binary string) as any other characteristic; a feature, which it was hypothesised, might undervalue the power of an established hierarchy as an antidote to too much change. In order to further test this hypothesis the reconstruction was then tested by assigning greater weight to the bishop meme. Weighting was simulated by inserting extra bishops into the meme strip [note 14]. In successive tests the switch from 7 to 8 bishops produced a tree sufficiently comparable to known patterns of historical descent (Figure 3) to claim a largely successful reconstruction. The switch is illustrated in Figure 4.

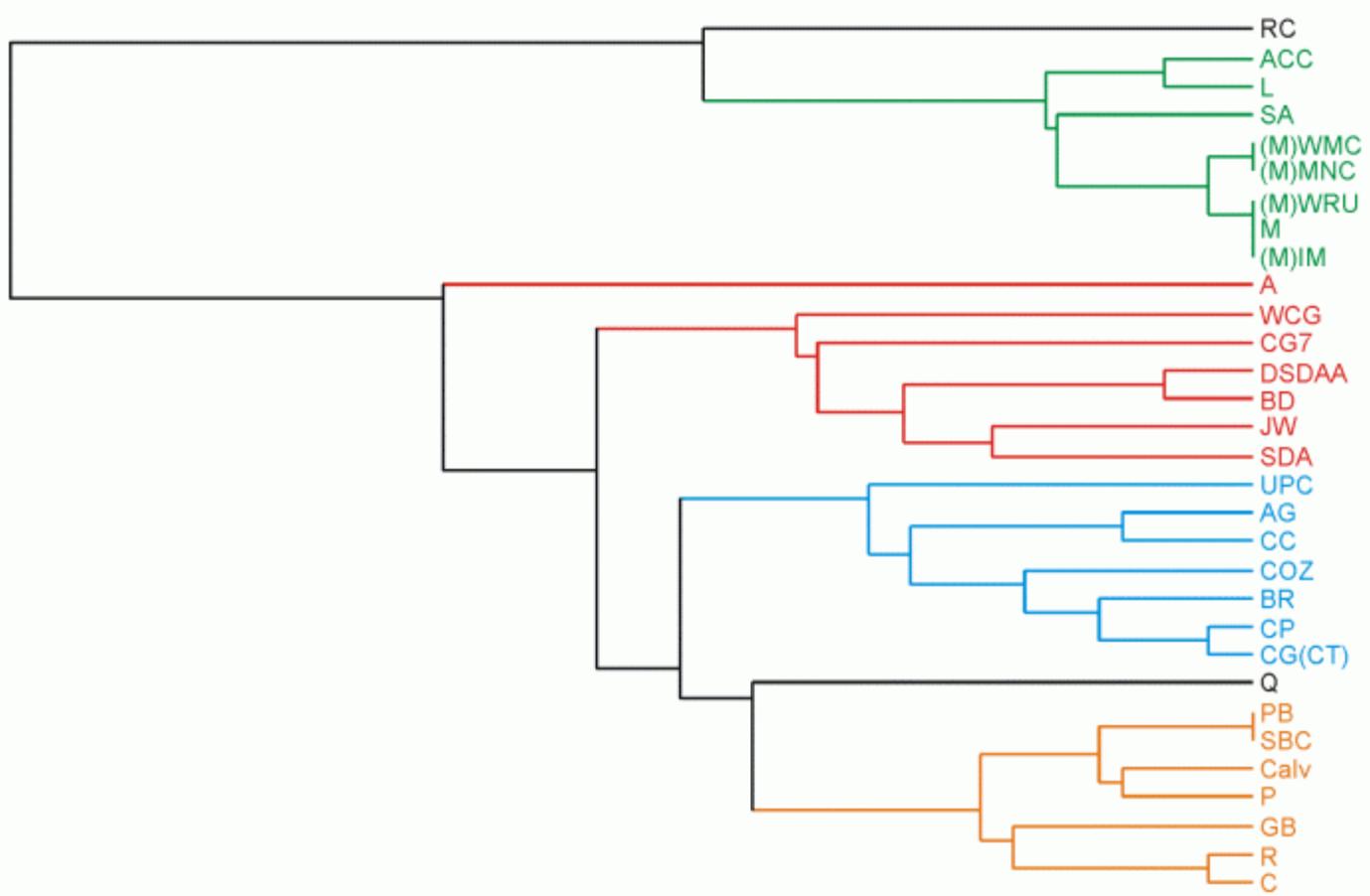
## **3 Discussion**

Figure 3 also reveals interesting questions for further analysis. The ancestral Roman Catholic Church does reconstruct as the isolated clade, without significant further splitting: a feature, which would suggest the meme of the apostolic succession to have even greater influence than the bishop meme. Remove one and some evolution is possible; remove two and a much greater radiation occurs. The Lutheran and Anglo-Catholic (i.e.: Church of England) clades reconstruct in the most plausible phylogenetic scenario, as co-descendants from some common ancestor. A conventional historical description would treat this as inaccurate. After all Henry VII founded the Church of England in a separate historical event and saw himself as Defender of the Faith against the Lutheran schism. In ideological terms however one could say that post Luther there were two clades, a Roman Catholic one and an unincorporated non Roman Catholic one, which subsequently precipitated and evolved separate branches. The issue mirrors the taxonomic debate

as to whether one biological species can evolve from another or whether two evolve from a common ancestor.



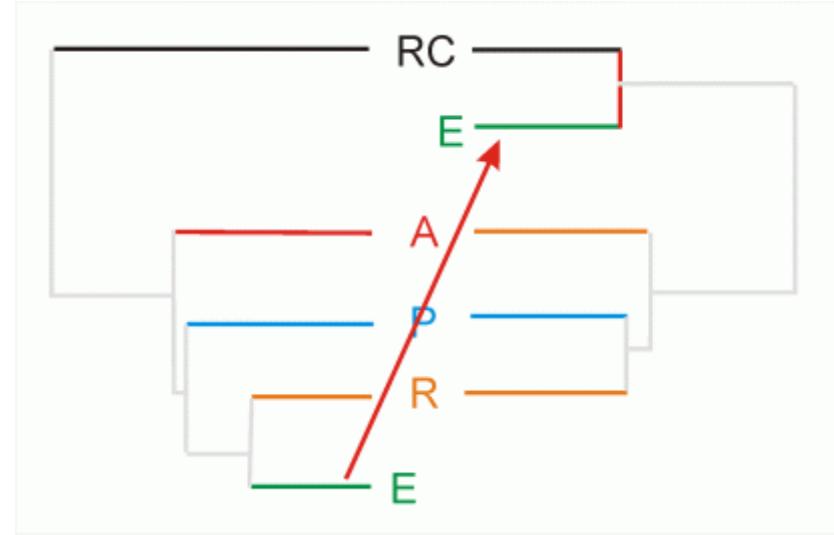
**Figure 2:** Phenogram derived from character states when the "Episcopal" meme weighs 1



**Figure 3.** Phenogram with the "Episcopalian" meme given a weighting of 8

There are two significant errors in Figure 3, The Salvation Army and the Quakers. Neither is necessarily surprising. The Salvation Army have replaced an Episcopalian polity with one modelled on a quasi-military structure and the unique implications remain to be investigated. Having confirmed the need to weight power structures in one form of hierarchy it is not unreasonable to suspect its influence in another. The Quakers are in some respects the complete opposite in having gone further than any other 'church' to question authority of position and indeed to blur the boundary between religion and ontology. Both the failures of the reconstruction can then be considered to be exceptions. The approach is demonstrated to be worth further investigation and the results are consistent with a view of organisations as evolving entities with recognisable patterns or memes at their core. The Episcopalian clades in particular are interpretable as organisations devoted to replicating of the 'bishop' meme. There are wider issues raised for understanding the history of Christianity however discussing them takes us beyond the scope of what is possible in a note of this kind. There are likewise issues for methods and terminology of operational research into memetics. Both merit further debate and consideration. At this stage our concern is to demonstrate that the hypothetical approach suggested by Wilkins can apparently be operationalized in practice. A more comprehensive paper will be submitted in due course. Testing of the software variants for other applications continues [note 15]. Meanwhile the evidence that a cladogram derived from a memetic code can

successfully reconstruct patterns of descent seems encouraging news for the emerging science of organisational memetics and a tool with considerable potential in evolutionary analyses of social organisations.



**Figure 4:** The effect of the changing reconstruction of the phenogram when the Episcopal meme is replicated 8 times in the character string (NB not a computer generated diagram)

## Notes

Our conception would see the 'pattern' (sensu Price and Shaw, 1998) or 'memeplex' of an organisation as an interconnected system of individual memes, much as Dawkins uses, and is clear about his use (1989 p. 271), the metaphor of the selfish gene as shorthand for a complicated system of interconnected and interdependent genes.

We emphasise that this contribution is intended as a note for rapid publication. A much fuller description of both the method and the data set will be submitted in due course.

And we acknowledge an anonymous reviewers pointing out that one is in fact derived from the other.

We cannot claim that the character set assembled for a given denomination necessarily represents a full memeplex. We are listing a series of characteristics postulated to be parts of that memeplex and informally describing them as meme strips.

The fossil record, and modern ecosystems, reveals periods of evolutionary radiation as new niches open up, The Burgess Shale and the Galapagos finches are well known examples. We expected tracking of postulated memetic divergence would be easier through equivalent events such as the religious diversity developed by European settlers in North America.

The existence of other evolutionary traditions in social science and other aspirations of scholars in the emerging field is acknowledged. The approach we take remains valid whether or not the replicators are considered as memes.

The primary source, where available, is the home page of the denomination concerned. We have also sought pages classifying and comparing religions and standard reference works. A fully annotated and hyperlinked character matrix of our 'taxa' is in preparation and will be submitted for comparative analysis as soon as the necessary referencing can be completed. Scholars interested in evaluating our results are welcome meanwhile to contact the authors.

A reviewer has commented on the potential bias towards 'sects', particularly of the millennialist variety, at the expense of variation within other traditions. It arises from our search for a sufficiently large population; many variants on the theme. There are of course variations within the apparent boundaries of larger religions and the same reviewer highlights historical diversity within Roman Catholicism. Comparative tracking of 'microevolution' within an apparent 'species' was judged beyond the research at this stage.

Borrowing again from biotic terminology in the absence of an alternative it may be that these groups have more in common with geographic 'varieties' or 'sub-species' than they do with 'species'. In Aldrich's (1999) terms they have not developed sufficiently rigid boundaries.

Provisionally entitled the Memetic Evaluation of Numerically Derived Evolutionary Lineages. A detailed description will be separately submitted. Meanwhile researchers who may want to submit this approach to critical scrutiny are welcome to contact the authors.

It is not suggested that these properties are memetic, or necessarily scientific. The data set was used as an independent test of the software. It seems reasonable however to assume that genetically related species would produce similar active principles thereby exhibit comparable toxicological/therapeutic properties.

The denominations which history shows have separated since the reformation in a process that we would regard as evolutionary, i.e. random within denomination variation was, on occasions, sufficient to generate new forms that secured their own niche in the total population of Christian religions. History records how the selection process and the attempts of established churches, or rulers adhering to them, to eradicate potential rivals was frequently a survival matter for those involved. We use the term 'clade' for groupings of denominations in the absence of an established hierarchical terminology in 'memetic' taxonomy.

Compare the observation, whose attribution escapes us, that paradigms do not change in the face of evidence; they change when the old guard finally dies or retires.

Working within the constraints of binary encoding of denominational characteristics this was the simplest method we could conceive to test the proposition.

In particular its application to market strategies of different firms appears promising.

## References

Aldrich, H. (1999) *Organizations Evolving*, New York, Sage.

Dawkins, R. (1989) The Selfish Gene 2<sup>nd</sup> Ed. London Penguin

Gell-Mann, M. (1996) Address to the US National Defence University " In the case of societal evolution, the schemata consist of laws, customs, myths, traditions, and so forth. The pieces of such a schema are often called "memes," <http://www.dodccrp.org/comch01.html>

Hull, D. (1988), Science as a Process. Chicago, University of Chicago Press.

Lapointe, F-J. & Legendre, P. (1994) A Classification of Pure Malt Scotch Whiskies. Applied. Statistics.**43**(1):237-257

Lawless, J. (1992) The Encyclopaedia of Essential Oils, Shaftsbury: Element Books.

Marsden, P. (2000). Forefathers of Memetics: Gabriel Tarde and the Laws of Imitation. Journal of Memetics - Evolutionary Models of Information Transmission, **4**. [http://jom-emit.cfpm.org/2000/vol4/marsden\\_p.html](http://jom-emit.cfpm.org/2000/vol4/marsden_p.html)

Price I. (1995) Organisational memetics?: organisational learning as a selection process Management Learning **26**(3):299-318

Price, I. and Shaw, R. (1998), Shifting the Patterns: Transforming the performance of people and companies Chalfont, Management Books 2000, UK.

Price I and Lord, A.S. (2000) Isomorphism in biotic and abiotic Complex Adaptive Systems International Conference of Complexity and Complex Systems in Industry, Warwick, UK.

Schein, E. H., (1985) Organisational Culture and Leadership: A Dynamic View San Francisco, Jossey Bass

Sneath, P. H. A. and Sokal, R. R. (1973) Numerical Taxonomy, San Francisco: Freeman. Wiley, E.O., D. Siegel-Causey, D.R. Brooks, and V.A. Funk (1991), The Compleat Cladist: A Primer of Phylogenetic Procedures, The University of Kansas, Museum of Natural History, Special Publication No. 19.

Wilkins, J. S. (1998). The evolutionary structure of scientific theories. Biology and Philosophy **13**(4):479-504.

Williams, R. (2000) The business of memes: memetic possibilities for marketing and management'. Management Decision, **38**(4):272-279

## Appendix: Character Set and Taxa

**Table 1:** Selected Characters

Class	Character
Status	Extant, Incorporated, Movement, Theology
Ecclesiology/Polity	Hierarchical, Autonomous, Episcopalian, Presbyterian, Congregational, Charismatic Leadership, Papal Authority, Prophet Founder, Women Ministers
Theology	Trinitarian
Christology	Son, Arian, Monophysite, Modal Monarchianism
Soteriology	Sola Fide, General Atonement, Particular Atonement, Predestination, Faith plus works
Eschatology	Amillennial, Premillennial, Postmillennial, Date Setting
Hermeneutics	Fundamentalist
Sabbath:	Sunday, Saturday
Sacramentology	Baptism, Infant, Belief, Total Immersion, Jesus Only Formula, Spirit Baptism, Holy Eucharist, Confirmation, Penance, Orders, Matrimony, Extreme Unction
Lifestyle	Excommunication, Lifestyle, Proselytizing, Exclusivist
Additional:	Soul Sleep, Investigative judgement, Annihilationist, No Hell

**Table 2:** Shortlist of Terminal Taxa

Vernacular Name	Key	Parent
Roman Catholic	RC	<root>
Anglican Catholic Church (CofE)	ACC	RC
Lutheran	L	RC
Calvinist	Calv	RC
Reformed	R	Calv
Presbyterian	P	Calv
The Wesleyan Methodist Connexion	(M)WMC	ACC
Independent Methodists	(M)IM	(M)WMC
The Methodist New Connexion	(M)MNC	(M)WMC
Salvation Army	SA	(M)MNC
Wesleyan Reform Union	(M)WRU	(M)WMC
The Methodist Church	M	(M)WMC
The Society of Friends (Quakers)	Q	ACC
Congregationalist	C	ACC
General Baptist	GB	ACC
Particular Baptist	PB	Calv
Southern Baptist Convention	SBC	PB
Adventist	A	PB
Jehovah's Witnesses	JW	A
Seventh Day Adventist	SDA	A
Davidian Seventh Day Adventist Association	DSDAA	SDA
Branch Davidians	BD	DSDAA
Church of God Seventh Day	CG7	SDA
Worldwide Church of God	WCG	CG7
Church of God (Cleveland, Tennessee.)	CG(CT)	PB
City of Zion	COZ	PB
Classical Pentecostalism	CP	COZ
Assemblies of God	AG	CP
Brownsville Revival	BR	AG
United Pentecostal Church	UPC	AG
Calvary Chapel	CC	CP

## THE SELFISH TEXT: THE BIBLE AND MEMETICS

By

Hugh S. Pyper

Dept of Theology and Religious Studies University of Leeds  
The Bible into Culture Colloquium, Sheffield 9-12 April 1997

Simply in terms of the number of copies currently in existence, the Bible represents one of the most successful texts ever produced. Whereas other great texts of the ancient world have either been lost or else exist only in a relatively small number of copies, the Bible is ubiquitous. It exists in over two thousand different languages and in many of those languages it exists in multiple translations. Something identifiable with the Bible in its present form has existed for nearly two millennia. If 'survival of the fittest' has any validity as a slogan, then the Bible seems a fair candidate for the accolade of the fittest of texts.

According to the collective authors of *The Postmodern Bible*, it is a 'truism' that the bible has exerted more influence on Western culture than any other book (The Bible and Culture Collective 1995: 1). In art, literature, politics and religion, biblical thought-forms, narratives and quotations are all-pervasive. As Western culture becomes globalised, so too does the bible. It is said that between a quarter and a third of all Japanese households possess a bible, in a country where only one or two percent of the population have any Christian adherence. This is because it is regarded as essential background for a proper understanding of Western culture. One effect of the spread of western culture through trade, conquest as well as missionary activity has been the spread of a collection of ancient Hebrew and Greek texts to every corner of the globe. Where Western culture goes, the bible goes too.

The purpose of this paper is to explore more fully how the biblical texts have achieved this remarkable success. It will be obvious that the model by which I intend to do this is a Darwinian one. Indeed, I propose to turn to one of the fiercest contemporary critics of the biblical world view, Richard Dawkins. His book *The Selfish Gene* (1976/1989), itself a runaway best-seller, has popularised the admittedly controversial idea that human beings, indeed all living organisms, can be construed as the 'survival vehicles' for their genetic material.

This claim is a variant on Samuel Butler's well-known description of a hen as 'an egg's way of making another egg'. An organism is a gene's way of making another gene. More pertinent to our purposes is Dawkins' further claim that there is a strict analogy between the processes of biological evolution and the development of human culture. This idea has been taken up by Daniel Dennett who adapts Butler's epigram to read 'A scholar is just a library's way of making more libraries.' (1991: 202). It is the following further adaptation of this slogan that forms the proposition that this paper will discuss: *western culture is the bible's way of making more bibles*. In an attempt to see if and how far this rather bold assertion can be defended, we will analyse in more depth the nature of the analogy that can be drawn between biological and culture evolution and in particular the usefulness of Dawkins' concept of the 'meme' in this context.

## **Genes, memes and texts**

In *The Selfish Gene* (1976), Dawkins contends that just as biological evolution can be studied at various levels—the gene, the genome, the individual, the gene-pool or the species—so cultural evolution can be looked at in multiple levels, from the spread of the simplest catch-phrase to the rise and dominance of the great civilisations of China or the Islamic world. In terms of evolutionary biology, the main point he argues in the book is that the clearest way to think about evolution is to work from the point of view of the smallest replicating entities, in the case of genetics, the gene. By analogy, studies of cultural evolution in Darwinian terms will proceed best by examining the smallest replicating units in culture. It is these that he designates as 'memes'.

He illustrates the concept as follows:

Examples of memes are tunes, ideas, catch-phrases, clothes-fashions, ways of making pots or building arches. Just as genes propagate themselves in the gene pool by leaping from body to body via sperm or eggs, so memes propagate themselves in the meme pool by leaping from brain to brain via a process which in the broad sense, be called imitation (1976:206).

Later, however, Dawkins becomes concerned to distinguish between a meme as a unit of information lodged in a brain and the phenotypic effects of that meme, such manifestations as the tune or the idea (1982:109).

Definitions are difficult, to say the least, in these areas. What exactly constitutes a meme or a culture defies classification and the recent literature on memes is bedevilled by shifting definitions and unsupportable generalisations and comparisons. Extravagant claims for the explanatory power of this concept have been made, including claims that the key to human self-understanding is in the new 'science' of memetics.

As the originator of the term, Dawkins has been far more modest in his assessment of the value of the concept. Dennett, who as a philosopher has written with more rigour on the subject than most, elaborates the concept of the meme in Chapter 12 of his *Darwin's Dangerous Idea* where he defines memes as 'self-replicating complex ideas which form *distinct memorable units*'. Even so, the concept remains notoriously fluid and therefore liable to abuse. How could it be applied to the bible? The bible if anything seems more like a repository of memes than a meme itself or even a 'meme complex'. In the ensuing argument, we will be using the term 'meme' for its convenience and its heuristic power in applying Darwinian insights to cultural developments but with a weather eye on its slippery nature.

## **The Bible as replicator**

At this point, however, I want to turn to a more fruitful line of enquiry provided by Dawkins' later, more rigorous discussion of the concept of the *replicator*. This he defines as follows: 'A replicator is anything in the universe of which copies are made'. (Dawkins 1982:83). At a banal level, that is a claim that can undeniably be made for the bible and so it may be of interest to explore the ramifications of this analogy.

Note, first of all, that Dawkins carefully does not say that the replicator must be *self*-replicating. There is a fundamental point here which has often been missed. DNA is sometimes described as a 'self-replicating' molecule. In one sense this is true. Given the right environment, a molecule of DNA is capable of acting as a template so that an exact copy of itself is produced. The important point is that it needs the right environment. DNA on its own cannot reproduce itself; it needs a complex of enzymes which will guide and manage the process. In fact, it is only at the level of the cell that we find a replicating structure which can produce copies of all its parts from raw materials in a simple environment.

In that sense, the bible is no different from DNA. Shut a bible, or even two, in a cupboard and you will certainly not find more bibles when you come back. Leave a jar of DNA on a shelf, however, and it will not increase either. Only in the context of a cell, of a 'survival machine', will we find DNA reproducing. Likewise, Bibles can only reproduce through the agency of a human reader who then takes steps to ensure that more copies are produced.

The crucial point which lifts all this from the level of a truism is the way in which Dawkins then refines the concept of the replicator. He distinguishes between an active replicator, the nature of which has some influence over the probability of its being copied, and a passive replicator, the nature of which is immaterial. DNA, the replicating molecule which encodes genetic information in cells is an example of the first, in that it exerts phenotypic effects on the environment through the protein synthesis it enables, which in turn influence whether it will be copied.

Active replicators modify their environment in such a way as to enhance their own reproductive capacity. Dawkins makes a particular study of the interactions between parasites and hosts. For example at a simple level, a gall wasp larva will carry genes that encode for the synthesis of chemicals that mimic the growth hormones of an oak tree, inducing the tree to grow an unusual structure which serves to protect and feed the wasp larva. Here the wasp genes are acting on the phenotype of the oak tree, not the wasp, in a way that enhances the wasp's reproduction but which may have a deleterious effect on the tree. The relevance of this parasite/host model for the consideration of the bible will become clearer as our discussion develops.

Unfortunately for our purposes, Dawkins' example of a passive regulator is a sheet of paper which is Xeroxed. On the face of it, this undercuts any analogy between the genetic material and the bible. He goes on, however, to concede that some pieces of paper are much more likely to be copied than others because of what is written on them. They then become active replicators as they convey information which acts on the reader and her environment in such a way as to induce her to copy the text. The argument we will pursue is that in this sense, the bible is indeed an active replicator, one which alters its environment so as to increase its chances of being copied. The intriguing questions then become how the bible alters its environment to increase the likelihood of its being replicated and why it has been so conspicuously successful in this.

Dawkins goes on to discuss other aspects of active replicators. It is a fundamental point in his argument that no process of replication is infallible. Strikingly, a favoured metaphor to explain this in popular genetic texts is the variability of the biblical text in different translations or through processes of copying. Robert Pollack in his *Signs of Life* for instance, sets out six English versions of James 4:5 to illustrate the phenomenon of alleles, the existence of variant

forms of the same gene within a population or indeed an individual genome (1994:38). Dawkins himself uses the 'mistranslation' of the Hebrew for 'young woman' as 'virgin' in the Septuagint version of Isaac 53 as an example of the potentially enormous phenotypic effects of a small change in DNA. He also provides a footnote explaining the Hebrew and Greek texts complete with citations in Hebrew, remarking that 'I suppose the scholars of the Septuagint could at least be said to have started something big' by this (1989: 16). This infiltration of biblical examples into the texts of popular genetics is an intriguing phenomenon to which we will return.

The crucial consequence of this variation is that when it occurs, some replicators may turn out to be less efficient than others at replication and so will tend to be replaced by the more successful replicators. For active replicators, whose nature affects their success in achieving replication, such variation may have a remarkable effect on their reproductive ability. Those which replicate most efficiently will, if all else is equal, come to predominate in the population.

Yet variability in itself is not enough; it must be coupled with stability. If 'successful' variants are to survive and out-compete the others, they must be conserved over time. Dawkins sets out the conditions for a successful replicator in the following slogan: longevity, fecundity and fidelity. The replicator must last long enough to reproduce, it must be capable of producing a sufficient number of copies, and these copies must be accurate. To ensure accuracy, the genetic material has a whole complex of 'editor enzymes' which repair and correct copying errors in DNA. So too the biblical text has become sacrosanct with a premium put on its accurate reproduction. The great complexes of the Masoretic apparatus and the libraries of biblical criticism which have sought to preserve the text in its 'original' form are the evidences of this.

The stability of a particular text or a particular DNA sequence, what Dawkins calls its 'longevity' is an important factor. The replicator must maintain its identity over time. Equally important, however, is its capacity to throw up variants which, when conditions change, may confer an advantage on the organisms which bear them. It is this balance between the ability to reproduce faithfully a particular variant but also to be able to produce variation if the circumstances favour it that confers reproductive success on any replicating system.

## **The cell and the community**

The bible then operates as a replicator in a way analogous to DNA. Like DNA, it stores information which can be read and translated and which contributes to its own reproduction. This, however requires the action of another level of agency. In the case of DNA, this agent is the cell where the information contained in DNA is translated into proteins which both structure and control the host of chemical processes which are necessary to sustain the life of the cell, and therefore the reproduction of DNA. In the case of the bible, the agency is a human community which will recopy and disseminate the text.

The crucial question then becomes how the active effect of the cellular DNA on the constitution of the cell or organism which is its 'survival vehicle' is paralleled in the case of the bible. That a case can be made is evident from the fact that the analogy has been pursued in the opposite direction, notably by Robert Pollack. In his book *Signs of Life* he explicitly embraces the analogy of DNA as text.

... I have organised this book around the notion of DNA as a work of literature, a great historical text. But the metaphor of the chemical text is more than a vision: DNA is a long skinny assembly of atoms similar in function, if not form, to the letters of a book, strung out in one long line. The cells of our bodies do extract a multiplicity of meanings from the DNA text inside them, and we have indeed begun to read a cell's DNA in ways even more subtle than a cell can do (1994: 5) Molecular biology is shot through with metaphors of reading: translation, transcription, reading enzymes and the like. Pollack extends this metaphor by suggesting that the genome is like an encyclopaedia, where the volumes are represented by the chromosomes, the articles by the sets of genes encoding for a particular character, the sentences by the genes themselves. Words are domains and letters are base pairs (1994: 21).

Nor is he alone in this. Dawkins himself speaks of the tempting analogy of seeing DNA as a 'family bible' (1995: 39), a record of our ancestry, slightly different for each one of us although he quickly goes on to point out flaws in this metaphor. Dennett makes the point that the strict analogy between genes and memes can be maintained on the ground that they are both 'semantic entities', by which he means that they constitute information which can be variously encoded. A gene is not simply a piece of DNA, although to be effective it has to be expressed as such. It could equally be said to be encoded in a sequence of letters on a page, just as a meme may be contained in the pages of a book.

But as texts, both DNA and the bible have to be read. In the case of DNA this is a matter of the synthesis of RNA and through it of particular amino acid sequences in cellular proteins. In the case of a text such as the bible, the analogous process, in Dennett's view, is that its memes influence a human mind and so influence a common meme-pool as to ensure the physical survival of the text. Dennett expresses this as follows: '...memes still depend at least indirectly on one or more of their vehicles spending at least a brief pupa stage in a remarkable sort of meme nest,; the human mind.' (1995: 349)

Mere reproduction of a text as a physical artefact is not enough to ensure its continued survival as Dennett makes explicit. Copies of books will only endure so long and the relative youth of even the earliest complete manuscripts of the bible bears this out. He quotes an analogy from Manfred Eigen who points out that a Mozart symphony cannot be said to survive as a living cultural entity unless it is played and replayed and checked for continuing value against other compositions. The bible must be read and must make itself read if there is to be reproduced. Its success in achieving this is what makes it an example of a highly adaptive active replicator. In this view the biblical reader, then, acts as the site of transfer of the information contained in the text to the meme-pool in which he or she operates. The book itself encodes memes which once active in the mind lead the human agents of that meme-pool to produce more examples of the text. But like all memes, in Dennett's view, they encounter competition. People have a lot of other things to do with their time and energy besides copying bibles, indeed a lot of other texts to read. What has lead to the particular success of the bible in this competition for mental space?

## The viral bible

Controversially, this can best be looked at under the rubric of its 'infectivity'. In a paper entitled 'Viruses of the Mind', Dawkins gives an account of the propagating power of what he calls a mind virus (Dawkins 1993). By this term he means a piece of information which ensures its own duplication without regard to the survival of the system it exploits. Viruses are propagated differently from the genes of their hosts. For instance, influenza viruses spread by coughs and sneezes rather than by being incorporated into a viable embryo for the next generation. This means that, unlike host cells whose genes will only be propagated by the reproduction of the organism of which they are part, viral genes and viruses have no vested interest in the reproductive success of their host.

So, how would a successful 'mind-virus' operate? The problem for a virus is that it must be incorporated into the replicative machinery of its host. What is the parallel mechanism among viruses of the mind? Such a meme will have to instill in the host a mechanism of conserving the meme, and a mechanism of propagating it. It would ideally act like the gall wasp to divert the host's energies to its own reproduction. It would also, however, be well advised to have a mechanism of conserving its variability so that any changes in the environment, including the intrusion of other foreign memes, and in particular any developments in the host's own immune system can be either countermanded or else outflanked. My tentative suggestion is that the bible instills a meme in its readers which aligns its own survival with that of the reader and his or her community. 'Your survival depends on mine' is the message that the bible gives. If the primary evolutionary drive is for survival, then a virus or a meme that 'persuades' its host that it is necessary to the host's own survival and therefore conveys a reproductive advantage will have an instant welcome into the replicatory machine. The virus becomes a symbiont, an organism which co-operates to mutual benefit with its host, rather than a parasite.

Of course, it is only in hindsight that the nature of the association can be known with certainty. It is quite possible that an organism will live quietly as a symbiont and then suddenly turn on its host at a later stage. Images of John Hurt and the parasitic alien bursting out through his body wall are only too apt in this connection, but the process may be a much quieter one. It may be that a false offer of reward to the host is never cashed out. From the invader's point of view, this matters little as long as it achieves its goal of its own reproduction.

## Strategies of survival

What the bible has to offer the communities it needs to reproduce is the unique variety of powerful strategies of survival it enshrines. Dawkins and other writers on memetics frequently cite the example of the 'God-meme' as a meme which has a powerful record of propagation across time and space. From a theological point of view, of course, the reduction of the complexity of human accounts of God to a single meme is a gross oversimplification. What seems to underlie their reading is rather a meme which predicates human survival on something other than purely biological grounds, which offers a space, not only for bodily survival but for memetic survival.

This has resonances with the account that Zygmunt Bauman offers of the whole enterprise of human culture. Culture, he claims, is a human construct designed to fend off the threat of death. It is a survival mechanism, which finds a way of promising a form of survival in the face of the inevitability of individual death. For Bauman, the Jewish tradition is the clearest case of the subsuming of individual death in communal survival. The individual may die, but his or her genes and memes will carry on. The duty of the individual then, in the sense of his or her best survival strategy becomes one of ensuring the survival of the group, not simply his or her own prolonged life. Christianity has adopted the alternative strategy of a promise of immortality, in that the believer's death is caught up in the context of the resurrection of Jesus. Both genetically and memetically, the afterlife of the believer is strictly irrelevant except in so far as belief in personal immortality act to sustain the continuity of the meme pool.

Both of these strategies are offered to the reader of the biblical text, together with stern warnings of the likely outcome of failing to abide by the word of the text. This is also aligned to a particular set of strategies which reinforce the integrity of the biblical text. In order to maintain continuity and identity, any organism or any gene-pool has to be able to filter out undesirable interlopers.

Again, a cunning usurper will both penetrate these defences but also quickly turn them to its own advantage, keeping out competitive genes and installing itself as the object of the host's attention. On an organismic scale, the efforts of a baby cuckoo to throw the host species' own eggs and chicks out of the nest but also its success in subverting the host's nurturing instincts to its own development are a classic case. The cuckoo succeeds in side-tracking the mechanisms of rearing the young which have evolved for the vital task of reproducing the host species.

The bible contains powerful instructions as to its own unique worth and the limits to be placed on the infiltration of foreign information or texts into the communities which propagate it. The whole process of canonisation, for instance, reveals a complex interaction between text and community which serves, for example, to oust the fledgling apocrypha and turn the community's attention to the ever-growing task of copying and commentating on the biblical text with an increased sense of its importance and of the need for its conservation.

The propagation of the text and the founding of new communities are also linked to the survival of the reader and his or her community, or meme-pool. The Hebrew bible is full of admonitions about the duty to hand down its teaching, and by implication the text, to the next generation. Secondly the text contains a strong message of evangelisation. The survival of the reader's community depends on the production of new texts and new communities. This complex of memes and of strategies forms a powerful ensemble to ensure the accurate transmission of the text.

## Biblical variation

What then of biblical variability? Superficially, of course, bibles show variation. The physical appearance of the bibles on our shelves is very different from that of the scrolls found at Qumran, an evolution that has something to do with ease of reading, portability and changes in the mechanics of reproduction. In another obvious sense, the bible has evolved out of its

component parts which themselves have undergone a long process of development. It now exists in a number of forms: the Tanakh, and the various canons of the Christian churches. Despite this variation it might be argued that within each community it has on the whole developed a fixed form.

However, that form itself preserves a great variety of strategies of survival. There is an analogy here perhaps in the way in which variant genes can be maintained in a population even if they have no particular advantage, or are perhaps deleterious. In many organisms, chromosomes and the genes they contain are carried in pairs. This means that an individual may carry two different variants of any gene, such variants being termed alleles. In most cases, one of alleles is dominant, so that in an individual who carries two copies of the gene, only the dominant characteristics are expressed. The consequence of this is that the individual may carry without any disadvantage another allele which could if expressed have a deleterious effect, but might also, in changed circumstances, turn out to confer an advantage. The sexual process leads to the constant reassortment of alleles which means that the population will be able both to express the alleles of the gene but also to carry them under the cloak of the dominant phenotype.

It is tempting to speculate whether some of the redundancies and doublings in the biblical text may have a similar function in that they can preserve maverick readings. These can be ignored by the mainstream interpreters, especially if an interpretative parallel to the dominance mechanism is in play, where by that verse or passage can be 'corrected' by appeal to other verses in scripture or the perceived overall theological thrust of the material. One day, however, the alternative reading may prove of interest or use to a particular interpretative community which then propagates the bible on the basis of that alternative reading.

Furthermore, even in its canonical form the bible can still generate variety. The information contained between the covers of any given edition the bible varies and develops, especially in terms of marginalia and commentary, which may, at times, have outweighed the biblical text in terms of importance. It is only necessary to count the number of editions of the bible currently available to realise how in adapting to the needs of different communities, cultures and age groups, the contents of the physical entity called the bible can vary widely. These variations serve to widen its appeal, or in other words to enable it to gain entry to and propagate itself in a whole variety of new environments.

In this connection, one of the most obvious sources of variation and adaptive strategies that the bible is its translatability. Translation is a good trick for increasing the number of bibles—I certainly possess at least 17 versions of the scriptures and I suspect that most regular biblical readers, let alone scholars, possess more than one version, something which would be unlikely for many other books in their libraries. From the blindly functional point of view, there is paradoxical advantage for a text in being written in a dead language once it has achieved a cultural dominance in another language group. It can potentially always be re-translated because the precision of the match between the words and the meanings cannot be guaranteed in the way that the text itself insists is important. It is more open to revision than a text inscribed in the native language of a culture, except in rare instances where a text preserves an older form of the evolving tongue. It is possible, for instance to find modern language paraphrases of Shakespeare, but there is a great resistance to producing new English versions of his works, whereas once the

interesting conservatism over the Authorised Version was broken, the floodgates of new biblical translations opened. Additionally, there is implicit permission for the text to be translated into the vernacular language of any community which uses the book, again increasing both its diversity and adaptability but also the sheer numbers of copies in existence.

This is by no means a one way process. The success of the bible in reforming the communities into which it moves through translation is also striking. As Carroll and Prickett observe in their introduction to the World Classics edition of the Authorised Version

What we *can* observe is that it was not just the Bible that was transformed in the course of successive reinterpretations. The Vulgate, a single, authoritative, monolingual instrument for the entire Western Church, was the instrument of the new imperial power of the Roman Church. Luther's Reformation translation of the Bible was to change the German language for ever; his commentary on Romans to set the agenda of theological debate for centuries. Tyndale's translation of the New Testament, on which the Authorised Version was to be so closely modelled, did the same if not more for English. (Carroll and Prickett 1997: xiv-xv)

Communities based on the bible may have a strong interest in conserving it unchanged. From the point of view of the bible, however, its ability to adapt to new communities is an essential part of its success. The fact that much human ingenuity has been expended on ensuring that the bible does *not* change and that such mutations have at times been physically rooted out merely goes to show the strong pressure on the text to mutate and its potential for evolution. As we noted before, the interests of the text and those of its nurturing community may not coincide.

This leads to the conclusion that when another champion of scientific realism, the astronomer Carl Sagan, lays into the Bible in his attack on what he conceives of as superstition in his recent book *The Demon-haunted World* he in fact reveals one source of its reproductive success. Contrasting the love of one's enemy enjoined in the Gospels with the celebration of holy war of Joshua, he writes 'The Bible is full of so many stories of contradictory moral purpose that every generation can find scriptural justification for nearly any action it proposes, from incest, slavery and mass murder to the most refined love, courage and self-sacrifice' (1996: 275). Indeed, and this has surely contributed to its survival. A book to which both the Apartheid regime in South Africa and its most fervent opponents can turn to justify their position may not offer simple moral precepts, but does ensure that both sides will own their own copies. For the survival of the *book*, its amazing capacity to sustain opposing camps is a very successful strategy. Part of its success is its very diversity, but also the fact that diversity can be differently enacted. As is well known, every cell in the human body contains the same genome, but this is differently expressed in different tissues to give cells which vary radically in form and function. The difference is the particular portions of the genome which are read. So too the bible contains more information than any one community can readily assimilate, especially as it may seem mutually contradictory or impossible to apply in a given situation. What then happens so often is the formation of a canon within the canon, where the community opts to read and follow a particular smaller set of instructions, read with a particular interpretative slant. This may change over time, giving a flexibility and yet continuity to the community. Biblical communities themselves show a capacity for survival which consists in a knack of maintaining continuity through change.

## Biblical advantage

It was allegedly the physician to Frederick the Great who when asked by that monarch to give a proof of the existence of God replied 'The continued existence of the Jews'; an existence bound up with the identity, adaptability and continuity that the bible confers. In a more theoretical vein Sir Peter Medawar attributes the biological success of human beings as a whole to a new form of inheritance; exogenetic or exosomal heredity. 'In this form of heredity, information is transmitted from one generation to the next through non-genetic channels—by word of mouth, for example, and by other forms of indoctrination; in general by the entire apparatus of culture' (1977: 14; quoted in Dennett 1995:342). Henry Plotkin in his *Darwin Machines and the Nature of Knowledge* cites the Bible and the Koran as just such devices of exosomal storage. He speculates on the selective advantage of such texts to the cultures which retain them by drawing on Bartlett's work on the degradation of oral narratives which implies that over time any group will retell a tale in such a way as to bring it into line with accepted norms. Plotkin argues that the 'exosomal storage of memes' in the bible may have preserved 'richness subtlety and beauty' in cultures which possess the book (1994: 220).

That continuity is bound up with the continued existence of the bible. The community of readers sees it as its duty to ensure the survival of the book. More than this, it sees the book as the guarantor of its own continuity and survival. The book itself contains a whole plethora of strategies for survival, and in particular, is the record of an amazing feat of cultural continuity as the diaspora communities of Jews manage to retain a sense of themselves as Israel, as members of one 'meme-pool' of cultural exchange, protected by firm filters from external memetic contamination. The fact that historians might take leave to differ over the actual continuity of the community and its immunity to outside infection is surely a proof of the power of the meme complex in question. Despite the available evidence of all that might have led to its dissolution and disappearance, the community is maintained, and the text is preserved.

Even more amazing is the development of communities of those supposedly excluded by the text, the Gentiles, who find ways of identifying themselves as Israel and arrogating to themselves both the promises and the duties imposed by the text, chief among which is the duty to ensure the connivance and dispersal of the text. Here the 'gene-pool' of Judaism, with its claim of descent from Abraham, is replaced by a meme-pool, a claim of descent from Abraham's faith, a line of argument already presaged in the Old Testament itself.

This is an astonishing success and one of crucial importance to the propagation of the text. The consequence of its incorporation into the canon of the Christian bible is an exponential leap in the number of copies produced. However, it may also be true that the text turns against the communities that have sustained it if that is to its advantage.

The horrid record of Christian anti-Semitism shows the consequences of the reappropriation of the filter mechanisms for memetic purity being turned against the original host community as the bible takes on a new existence as Christian Scripture. A prime exemplar of the selfishness of the text might be seen in the reformation where the text operates to cause a major breach and disruption in the community which sustained it in order to take advantage of the new technology of printing through the propagation of a meme that removed the authority of interpretation from

the institution to the individual and to the possibilities of reproduction within vernacular language communities. The peril of too close an association with the host community may be that the text will fall with the community that guards it. The success of the bible has been predicated on its ability to 'jump ship' when necessary.

Again, this is a slightly self-parodic example of the prevalence of the intentionalist fallacy in the discussion of Darwinian replicants. It is patently a fallacy to argue that the bible provoked the Reformation in order to increase its own population, but the facts remain. Whatever damage the Reformation did to the church and to the victims of the religious wars that accompanied it, it was certainly good for the bible.

## The survival of the bible

What then of the bible in late twentieth century society where the traditional communities of interest in the bible may be thought to be in danger of collapse? Selection is a cruel business as many species and their DNA find every day. Surviving intact for a hundred million years is no defence when your habitat is suddenly filled with industrial pollutants. The best that can be said for any replicator is that it has survived *so far*. Tomorrow is another day and will perhaps bring an insurmountable challenge.

Even if the worst comes to the worst in terms of a diminishing community of biblical readers, the important thing for the text in a memetic view is not that it be read but that it be copied. By achieving an iconic status within a culture, the text can relieve itself of the pressure of seducing readers. The baptismal bible or wedding bible may be a gift that is never read, and no-one is likely to open the pages of the court room bible on which oaths are sworn, but they must be complete as the community is well endowed with reverence for the canon that says 'all or nothing': an incomplete bible is not a bible. This may work more powerfully for the New Testament than the Old, as the New circulates independently, but it nonetheless still allows for the reproduction of redundant, unread material within the New Testament. This, in turn, will allow for the introduction for the sake of iconic completeness of the even less appealing Old Testament [1]. So the bible can survive in some sense on the vestiges of a culture which valued it.

However, the situation is by no means as gloomy as that. As I was finishing this paper, by coincidence two documents arrived together in my pigeonhole. One was a copy of a review of several modern popular debunking biographies of Jesus, including that by A.N. Wilson, which finished with the comment from an eminent critic, 'At least he sends us scurrying back to the gospels'. The second was a leaflet from the National Bible Society urging donations for the dispatch of bibles to the displaced and starving population of Zaire.

Whatever one thinks of the anti-biblical polemics of an A.N. Wilson and the response of the Bible Society to the disasters of war, there is a common feature to these documents; both seem to serve to increase the sale and distribution of the bible. The Bible Society can still launch an appeal to increase the number of bibles in the belief that the bible will contribute uniquely to the survival of the people of Zaire. A.N. Wilson, who seeks to debunk it and would no doubt pour scorn on the Bible Society's work finds himself both propagating biblical memes in his own texts

and sending his readers back to consult the original text. The biblical text is not affected by the fact that the person who reads it is only doing so to refute it as long as there is a sufficient cultural community or meme- pool to maintain the argument and therefore sustain the need for the text.

Yet the bible has always shown an astonishing facility in generating communities that will see it as worth transmitting. What memetic pull is it that brings together a group of scholars of diverse backgrounds and beliefs to Sheffield to discuss the bible and culture, for instance? Has the bible 'succeeded' in making a bid beyond its native environment of the religious community, one which may be severely threatened in a new memetic ecology. Is it now able to persuade communities of readers to consider it as a cultural artefact, using the same memetic appeal as Homer and Plato? If so, what a tribute to the extraordinary staying power of the particular combination of memes which the text and the communities it builds around itself enshrine. Having formed communities about itself for two thousand years, often by co-opting its enemies, is the bible proving able to do this again by infiltrating not religious but cultural discussion?

A telling example of this ability of the biblical text to infiltrate the most unlikely communities is the very genre of popular genetic writing of which Dawkins is the most celebrated practitioner. For someone who evinces such a suspicion of the influence of the bible, he makes a surprising number of references to it. His main rival, both as a best-selling writer of popular genetics and as advocate of what to Dawkins seems a 'heretical' view of Darwinism is Stephen Jay Gould, whose books are shot through with biblical allusions. Another populist geneticist, Stephen Jones', recent book, *In the Blood: God, Genes and Destiny*, a companion to a television series, is an interesting case in point. In it he covers such topics as the Lost Tribes of Israel and the concept of Armageddon, in the process alluding to a large number of biblical texts and outlining many biblical stories. It would be ironic, would it not, if we were to conclude that Dawkins himself has become a 'survival machine' for the bible, a 'meme nest' for its dispersed memes which may induce readers who would otherwise leave their bibles unread to go back to the text.

Dawkins, however, is merely one articulate representative of a much wider conversation in a global gene pool which could loosely be designated as 'western culture'. Insofar as we have seen that the survival of the bible seems to be predicated on the persistence of its peculiarly effective set of memes which induce reading communities to propagate it, it is Dawkins inescapable cultural environment which is in evidence here.

But are we then simply the victims of the bible? Dawkins ends *The Selfish Gene* with a rallying cry 'We are built as gene machines and cultured as meme machines, but we have the power to turn against our creators. We, alone on earth, can rebel against the tyranny of the selfish replicators' (1976; 215). Is he actually a witness that we may think we can rebel against biblical memes but that such replicators have an uncanny power to survive all our efforts? The telling passage in his *River out of Eden* (the title itself needs no comment) on the 'ravishing' poetry of the Authorised Version of the Song of Songs, and the 'lifetime's repetition' which has given it its own haunting appeal despite the possible inaccuracy in translation (1995:40) argues that aspects at least of the bible have succeeded in inserting themselves into the 'meme nest' of Dawkin's mind in such a way that they are transmitted, if not replicated.

Dennett comments on Dawkin's challenge to the power of the memes, 'This "we" that transcends not only its genetic creators but also its memetic creators is ... a myth.' (1995:366). Dennett himself ends his 1995 book with an ambivalent plea for the preservation of meme complexes such as religions for their cultural enrichment. 'I love the King James Bible', he declares. 'My own spirit recoils from a God Who is He or She in the same way in which my heart shrinks when I see a lion pacing neurotically back and forth in a small zoo cage. I know, I know, the lion is beautiful but dangerous; if you let the lion roam free it would kill me; safety demands that it be put in a cage' (1995: 515). This is a rather extraordinary paragraph which is somewhat baffling in its implications. What is it that distresses Dennett so much? It seems to be the use of gendered language of God, not the idea of God. Is the King James Bible the cage, in which case what is the love he bears it, or does it need to be in the cage with the lion because of its dangers?

My own view, and no doubt I here manifest symptoms of my own freight of memetic viruses, is that the bible has so firmly entrenched a place in our culture that it is ineradicable. It is not a parasite but a constituent part of the great complex of meme complexes that can be designated 'western culture', part of the exosomatic genome of that culture's members. More than that, I see it in Plotkin's terms as an indispensable source of what might be called 'memetic diversity.' In agricultural genetics, one of the most worrying trends has been the loss of diversity from the appellations of food plants and animals. There are obvious superficial gains, not least to seed companies and fertiliser manufacturers in growing vast tracts of pure stands of the 'best' varieties, the judgment of 'best' depending on the particular values of the grower or the market. Ease of marketing may well win out against nutrition. However, there is a potential disaster looming if the super variety is suddenly attacked by a pathogen or if there is a major climatic shift. A variety may be fit for the purpose and the conditions of the moment, but what if conditions change?

Here it becomes vitally important to maintain a 'gene pool' of wild relatives of the crop plants which may themselves have all sorts of drawbacks from the point of view of the technology of farming, but which have shown themselves able to fend for themselves in this competitive world over time. Such wild populations contain a huge diversity of genetic material maintained over time and a vast potential for diversity and for change. Can we view the Bible as a sort of cultural 'memetic reserve'? Parts of it may seem irrelevant, redundant, even detrimental to our survival, but it has kept going. As Medawar and Plotkin indicate, it may serve to maintain a memetic richness and complexity, a inexhaustible source of variety which may contain the unexpected counter to forces that threaten to impoverish our cultural lives.

But Dennett's rather inarticulate declaration of love for the bible suggests other possibilities. This paper has, of course taken a slightly wry look at a provocative re-reading of the dynamics of cultural development. Nothing in the theory of memetics can help us to establish the truth or falsity of a meme; it can only deal with its frequency and prevalence. Questions of reference are not raised. Indeed, the practitioners of memetics have erected some pretty formidable filters to debar any such questions. Methodologically, this may be necessary to prevent muddled thinking, but methodology is not truth despite its strong tendency to become so. The easiest way to filter out a proposition is to declare it to be either meaningless or false. A very different account could have been given on the premise that the bible reflects the encounter of God with the complex web of human culture and individuality, a premise which *methodologically* Darwinism cannot

entertain. The attempt, however, to follow through such a methodology is a discipline which I hope has brought to light intriguing connections which any account of the relations of the bible and culture might need to take on board.

More radically, however, is survival, as Darwinism must have it, really the primary value, or is it in turn a methodology which has become a truth? The lion may be dangerous, and human culture as Bauman argued may well have been a device to keep it caged and to ensure survival, just as the wild lion has been confined practically to game reserves. The beauty of the cage, then, is in some sense engendered by the lion. Letting the lion out would have disturbing implications for culture as well as for scientific method. But is there no other mode of co-existence between human and beast, between human beings, the bible, and the God which its cages and displays? Here I want simply to recall that Isak Dinesen has a wonderful account of the mutual respect between the hunter and the lion which grows from the fact that each knows that the other is hunter as well as prey. Somewhere in this may be a dynamic that can lift us beyond the mechanisms of the meme and into true encounter.

## **Footnote**

The Bible carries a great deal of seemingly redundant information: the detailed instructions for the construction of the tabernacle, for instance. Intriguingly one of the more unexpected findings of molecular biology is that a large proportion of the DNA in any genome is seemingly also redundant, consisting of simple repeating sequences which do not encode any gene. This redundant DNA is, however, copied and transmitted to the next generation as faithfully as any other. Opinion is still divided as to the function of this material. It may have structural or geometrical implications.

## **References**

- Bauman Z. (1992) *Mortality, Immortality and Other Life Strategies* (Cambridge, Polity Press)
- Carroll R.P. and Prickett S. (eds) (1997) *The Bible: Authorised King James Version with Apocrypha* (Oxford, Oxford University Press)
- Dawkins R. (1976) *The Selfish Gene* (Oxford, Oxford University Press)
- (1982) *The Extended Phenotype* (Oxford, Oxford University Press).
- (1989) *The Selfish Gene* New Edition (Oxford, Oxford University Press).
- (1993) 'Viruses of the Mind' in Dalhousie, B. (ed.) *Dennett and His Critics: Demystifying Mind* (Cambridge MA, Blackwell)
- (1995) *River out of Eden: A Darwinian View of Life* (London, Weidenfeld and Nicolson).
- Dennett D.C. (1991) *Consciousness Explained* (London, Allen Lane).

- (1995) *Darwin's Dangerous Idea* (London, Allen Lane).
- Jones S. (1996) *In the Blood: God, Genes and Destiny* (London, HarperCollinsPublishers).
- Medawar (1977) 'Unnatural Science' *The New York Review of Books* Feb 3: 13-18.
- Plotkin H. (1995) *Darwin Machines and the Nature of Knowledge: Concerning Adaptations, Instinct and the Evolution of Intelligence* (London, Penguin)
- Pollack R. (1994) *Signs of Life: The Language and Meanings of DNA* (London, Penguin).
- Sagan C. (1996) *The Demon-haunted World: Science as a Candle in the Dark* (London, Headline Book Publishing).
- The Bible and Culture Collective (1995) *The Postmodern Bible* (New Haven, Yale University Press).

# DO ANIMALS HAVE MEMES?

By

**Simon M. Reader & Kevin N. Laland**

Sub-Department of Animal Behaviour, Department of Zoology

University of Cambridge

Madingley, Cambridge, CB3 8AA, UK.

Tel: +44 (0) 1954 210301, fax: +44 (0) 1954 210247

smr1001@cam.ac.uk, knl1001@hermes.cam.ac.uk.

## Abstract

Imitation has been put forward as a defining feature of memetic transmission. Since there is currently poor evidence for imitation in non-human animals, such definitions have been interpreted as restricting meme theory to the study of human behaviour patterns and birdsong. We believe this is a mistake. Human capacities for imitation, teaching and language may well account for the extraordinary diversity of human culture compared with animal proto-cultures, but imitation is simply one mechanism of transmitting acquired information between individuals. As long as information is transmitted with sufficient fidelity to be replicated in the brain of the receiver, any social learning process will do. Non-human animals may be poor imitators, but many are excellent social learners. We argue that the meme concept can, and should, be applied to animal cultural transmission.

**Keywords:** social learning, cultural transmission, acquired information, animals, imitation, replicator, meme.

## 1 Introduction

In her recent book *The Meme Machine* Susan Blackmore (1999) suggests that what makes humans special in comparison with other animals is our extraordinary capacity for imitation. According to Blackmore, imitation is the linchpin of meme transmission, and as evidence for imitation in animals is weak, humans alone can be regarded as possessing and transmitting memes.

In this article we present a counter position to Blackmore's, arguing that animals should be allowed the right to carry memes. We present our arguments, not as a criticism of Blackmore's position, but rather in an attempt to open up an informed debate over this interesting issue. We agree that evidence for non-human animals routinely imitating is weak, but argue that imitation need not, and should not, be the defining feature of a meme. Transmission fidelity, not the psychological process underlying transmission, is a determining feature of whether a meme can spread and replicate.

## 2 Imitation and the meme

The definition of the term 'meme' has been the subject of some controversy (e.g. Rose 1998). However, transmission of acquired information by imitation is a common element to many

definitions. For example, a meme has been defined as "*an element of a culture that may be considered to be passed on by non-genetic means, esp. imitation*" (Oxford English Dictionary), as "*a unit of imitation*" (Dawkins 1989, p.192), and as "*whatever it is that is passed on by imitation*" (Blackmore 1999, p. 43). In these definitions the use of the term 'imitation' is frequently qualified by the phrase "*in the broad sense*" (Dawkins 1989, p.194; Blackmore 1999, p.7 and p.43). Indeed, Dawkins's first definition of the term appears to equate imitation with all kinds of cultural transmission processes ( "*[a meme is] a unit of cultural transmission, or a unit of imitation*", Dawkins 1982, p.192, reversed original italics), and includes examples of cultural transmission such as the spread of milk bottle opening in British tits (Fisher & Hinde 1949; Dawkins 1982, p.109). Other authors do not specifically address this issue, but use the term 'imitation' to refer to social learning processes in general (Bonner 1980, p.18 & p.165; Goodenough 1995) and agree that animals possess memes (Bonner 1980, p.188). Blackmore however explicitly excludes all non-imitative forms of social learning as capable of transmitting memes, and so, with the exception of birdsong, restricts memes to humans (1999, p.49).

There is little doubt that animals regularly acquire learned information from one another. The term 'social learning' refers to learning that is influenced by observation of, or interaction with, another animal or its products (Box 1984). Experimental investigations have revealed that imitation is just one of several processes that can result in social learning (Galef 1988). There exist numerous reports of novel behaviour patterns spreading through animal populations through social learning processes (for reviews see Lefebvre & Palameta 1988 or Galef 1988). Famous cases include termiting in chimpanzees (Goodall 1964), food washing in Japanese macaques (Kawai 1965), the opening of milk bottles in British birds (Fisher & Hinde 1949), dietary preferences in rats (Galef 1996), birdsong (Jenkins 1978), and fear of snakes in rhesus monkeys (Mineka & Cook 1988). In the majority of cases where the transmission process is investigated, behaviour patterns are not transmitted by imitation, but instead result from other, simpler processes such as local enhancement, where an animal's attention is drawn to an object by the actions of another, in a manner that results in learning (Galef 1988). Yet animals clearly have behavioural traditions based on acquired information transmission. Moreover, many of these animal traditions appear to change over time (e.g. Gibbs 1990) in a manner perhaps consistent with the predictions of memetic evolution.

So why do researchers place such a premium on imitation? It would seem that, for Blackmore, the key to memetic transmission is that something is learned about the form of a behaviour rather than about the environment (1999, p.49). For example, Blackmore argues that milk bottle opening in British birds is not a meme since "*the tits already knew how to peck; they only learned what to peck*" (Blackmore 1999, p.49, reversed original italics).

However, while in the past definitions of imitation have stressed the learning of a new behaviour through observation (e.g. Galef 1988), it is now widely recognised that imitation is not solely, and may rarely be, concerned with learning a new motor pattern: imitation is also concerned with learning the context and consequences of an established motor pattern (Heyes 1993; Heyes 1996, p.381). For illustration, consider the case of someone learning to play tennis through imitation. Here the apprentice is not learning to run, hold racket like objects, hit things or wave their arms around - all of the required motor patterns are already part of their repertoire. They are learning to move their body in a particular way (for instance, to move the racket towards the ball), at a

particular time (when the ball is approaching), to generate a certain result (to hit the ball in the desired manner). The tennis playing meme is not exclusively concerned with motor patterns: it is a complex of information that also concerns the location appropriate for the behaviour, the objects with which one interacts, and the consequences of the behaviour. To take another example, this time the 'making pumpkin soup' meme described by Blackmore (1999, p.61), we have a case where the recipient of the new recipe is not preparing food, cooking, or even making soup for the first time. The cook is simply carrying out an existing motor pattern (that of making soup) with an ingredient novel to the soup-making context but familiar in other contexts.

Exactly the same logic applies to the milk-bottle-top opening birds. They are not learning to peck any more than the tennis apprentice is learning to run around or hold rackets: that motor pattern is already part of their repertoire. They are learning to peck a particular object (the milk-bottle), found in a particular location (on a doorstep), to generate a particular consequence (the cream reward). Similarly, Japanese macaques are not learning to move their hands in water; they are learning that if they move particular objects (the sand-covered sweet potatoes) in water they can generate a desirable reward (that is, sand-free food). Imitation is not a criterion upon which the meme-carrying of animals and humans can be distinguished.

### 3 Reconstructed memes

A second, but related, objection to allocating meme status to animals is that in cases of animal social learning most or all of the meme is not transmitted, but rather reconstructed. To quote Blackmore once more, "*other forms of social learning do not support a replication system with true heredity, because the behaviour is not really copied*" (1999, p.50). To remain with the milk-bottle-opening example, current experimental and theoretical analyses (Lefebvre 1995; Sherry & Galef 1984) concur with Fisher & Hinde's original conjecture that the majority of birds have learned the behaviour pattern in some way from comparatively few individual innovators. There is no evidence that the birds transmit any more information than a tip-off that milk bottles contain cream (Hinde & Fisher 1951), but that is enough for each observer to reconstruct the behaviour of milk-bottle opening. Here it is the meme for milk-bottle opening that is transmitted, not the specific opening technique, which is quite variable even within individuals (Fisher & Hinde 1949). This reconstruction process is very different from the direct copying of DNA, and leads Blackmore to question whether it is legitimate to describe such a process as replication. Yet while this particular meme may not be directly copied, it none the less appears to be replicated with unerring reliability. The first report of milk-bottle opening was in Swathling near Southampton in 1921 (Fisher & Hinde 1949) and since then, it has spread throughout Britain, into several other European countries, across to tens of other species, and is apparently still prevalent some 75 years later. If fidelity in the form of reliable reproduction is a criterion upon which memes are judged, milk-bottle-top opening will surely qualify. It is highly likely that cream-drinking is just one element of these birds' socially learned foraging repertoire, implying that other feeding memes may have exhibited a corresponding decrease in frequency. There is no reason to think that the birds could not institute a mutant variation to exploit some other food source.

This raises a policy decision for memetics. Should reconstructed memes count as genuine memes? We argue that they should for two reasons. First, virtually all memes, including those of

humans, involve an element of reconstruction (Sperber 1996). Humans when they imitate rarely do so perfectly, and they are typically forced to re-evaluate and adjust their behaviour in the light of sensory feedback (Piaget 1962; Yando et al. 1978; Custance et al. 1995). To argue that largely reconstructed memes are not memes would require an arbitrary and unenforceable rule to be employed regarding just how much reconstruction is allowed before acquired information qualifies as a meme. Reconstruction cannot be quantified in any meaningful sense. Second, reconstructed memes fit the Darwinian model as effectively as perfectly transmitted memes. They too can replicate and evolve, and to eliminate them on arbitrary grounds at this early stage in the science of memetics risks eliminating a large number, maybe even the majority, of interesting cases of social transmission that may benefit from memetic analysis. What counts is not how transmission occurs, but whether the product is similar in the transmitter and receiver - in other words, what counts (amongst other things) is fidelity (Dawkins 1989).

## 4 Information versus phenotype

A further criticism of memetic definitions based upon imitation is that emphasis is shifted from the transmission of information to the transmission of that which is imitated, i.e. behaviour patterns, or the memetic phenotype. Delius (1991) was explicit on the fundamental nature of this distinction, stating, "*memes stress the transmittance of coded information rather than of behaviour itself*". Delius defines memes as "*synaptic patterns that code cultural traits*" (where cultural traits are "*behavioural items acquired through social learning*"). Boyd and Richerson made a similar point, arguing, "*the essence of culture is encoded information rather than the behaviours that result from this information*" (1985, p.43). This discrimination is important, since it seems entirely feasible that the same meme could produce very dissimilar behaviour patterns in different individuals or in different circumstances (see e.g. Boyd & Richerson 1985, p.41-43), and that identical information could be learned from different behavioural performances (Sperber 1999). For example, the meme for making pumpkin soup could be transmitted by a perfect demonstration of the technique, or by a demonstration riddled with obvious mistakes. Prior knowledge of cooking would ensure that the meme for soup making was reconstructed accurately, without the inclusion of such mistakes.

## 5 Birdsong: a special case?

Although animals in general are frequently regarded as not having memes, researchers commonly make an exception of birdsong. However, learning a song does not involve learning a novel motor pattern, and would appear on the surface to be inconsistent with Blackmore's criteria for memes. Yet perhaps we are being pedantic, since the acquisition of birdsong involves the social learning of novel behavioural elements, if vocalizations can be regarded as such. Moreover, no doubt, most birdsong meme enthusiasts regard song learning as a form of imitation. Yet social learning theorists have long made a distinction between vocal and motor imitation (Galef 1998; Heyes 1994). The former is regarded as a less challenging form of social learning, since there is a greater correspondence between the sensory feedback from the learner's own vocalizations and those of its tutors than in motor imitation, for which the sensory experiences of *doing* and *seeing others do* are typically very different. What is more, birdsong is a highly restricted form of social learning, different from the general capacity for social learning found in humans. The processes underlying song acquisition allow birds to learn song and

nothing else. That is not to say that songbirds are incapable of other forms of social learning, but rather that these other forms probably rely on alternative mechanisms to song learning. Whatever the processes underlying birdsong, there is no evidence that they are more similar to the processes underlying human culture than other forms of animal social learning.

So why is an exception made of birdsong? We suggest two reasons, both unconnected to the psychological processes that underlie learning. First, Dawkins mentions birdsong in *The Selfish Gene* (1989, p.189). Dawkins's authority carries a great deal of weight in memetics, and if he says birds have memes that is enough for many. Second, birdsong has been subject to some of the most elegant empirical work on memes (e.g. Burnell 1998; Lynch 1996; Lynch et al. 1989). It turns out that birdsong lends itself to the meme concept beautifully, perhaps better than almost anything else, even human culture. We would welcome similar memetic analyses of other forms of animal social learning.

## 6 Memes in non-human animals

We end with three illustrative examples of cultural transmission in non-human animals to which meme theory may profitably be applied. First, an example from the apes. In a recent synthesis of long-term field studies of chimpanzees across Africa, Whiten et al. (1999) documented "*significant cultural variation*" after ecological explanations for this variation were discounted, with 39 different behaviour patterns identified that are traditional in some communities but absent in others. An earlier paper by McGrew (1994) reached similar conclusions. Whiten et al. (1999) argue that these behaviour patterns are most likely acquired by a mix of imitation and other processes that can result in social learning, but Whiten (personal communication) agrees that these different behaviour patterns can be classed as memes, as defined by the Oxford English Dictionary, especially in the sense that Boyd and Richerson (1985) emphasized. For example, the essential 'idea' of ant-dipping by the two-handed method used at the Gombe National Park has been observed in several successive generations, despite minor variations in its manifestations (Whiten, personal communication). Cultural variation has previously been identified only for single behaviour patterns in non-human animals, such as the local dialects of song-birds, and it may be possible to examine the dynamic relationships between the different memes which make up chimpanzee proto-cultures.

The meme concept need not be restricted to animals phylogenetically close to humans. In an experimental study, Curio et al. (1978) demonstrated that observer birds could learn to mob a non-raptorial bird, and even a plastic bottle, as a consequence of witnessing another bird mob at its presentation. This behaviour was transmitted along chain of six individuals, with each individual acting as a demonstrator for the next bird in the chain. Here a meme for the context in which mobbing is an appropriate response is transmitted. However, there were limitations to what could be transmitted. For example, birds did not learn to mob an empty cage, and learned to mob the plastic bottle stimulus less strongly than the bird stimulus. There may be similar restrictions on the successful transmission of human memes.

Third, an example from fish. Mating sites of the bluehead wrasse (a coral reef fish) have remained in daily use over 12 years (four generations) without changing location (Warner 1988). There was no obvious correlation of sites with resources, and many more potential mating sites

available. Entire populations were experimentally replaced, and the transplanted populations chose new sites; but where females were transplanted to locations containing native females, the newcomers only mated at established sites, implying that tradition rather than resources is important in determining mating site location. After further experiments Warner proposed that these arbitrary traditions were probably transmitted by females following one another (Warner 1990), a very simple social learning process. Here we have an example of a long-lived, arbitrary tradition transmitted without imitation. Again, we can consider the mating site preference that each fish learns as a discrete, replicating unit of information, and examine the diffusion dynamics of these memes.

## 7 Conclusions

We argue that a defining feature of a meme is that of a culturally transmitted replicator. The proposed memes in our examples show the three characteristic qualities of replicators detailed by Dawkins: copying-fidelity, fecundity, and longevity (1989, p.194; see also Dennett 1995, p.343). As such, these proposed memes seem as likely as any human meme to undergo evolutionary processes through heredity, variation, and selection. Hence we believe that memetics should not prematurely rule out these animals' memes. Indeed, animal social learning may be a useful testing ground for the meme concept. After all, it is entirely plausible that in future years animal social learning may eventually be judged to fit the meme model better than human culture.

## Acknowledgments

This work was supported by a BBSRC studentship to Simon Reader and a Royal Society University Research Fellowship to Kevin Laland. Jeremy Kendal made helpful comments on the manuscript.

## References

- Blackmore, S. (1999). *The Meme Machine*. Oxford, Oxford University Press.
- Bonner, J. T. (1980). *The Evolution of Culture in Animals*. Princeton, NJ, Princeton University Press.
- Box, H. O. (1984). *Primate behaviour and social ecology*. London, Chapman and Hall.
- Boyd, R. & P. J. Richerson (1985). *Culture and the evolutionary process*. Chicago, University of Chicago.
- Burnell, K. (1998). Cultural variation in savannah sparrow, *Passerculus sandwichensis*, songs: an analysis using the meme concept. *Animal Behaviour*, **56**:995-1003.
- Curio, E., U. Ernst & W. Vieth (1978). The adaptive significance of avian mobbing II. Cultural transmission of enemy recognition in blackbirds: Effectiveness and some constraints. *Zietschrift für Tierpsychologie*, **42**:184-202.

- Custance, D., A. Whiten & K. Bard (1995). Can young chimpanzees (*Pan troglodytes*) imitate arbitrary actions? Hayes and Hayes (1952) revisited. *Behaviour*, **132**:837-859.
- Dawkins, R. (1982). *The Extended Phenotype*. Oxford, Oxford University Press.
- Dawkins, R. (1989). *The Selfish Gene*, 2nd edition. Oxford, Oxford University Press.
- Delius, J. D. (1991). The nature of culture. In *The Tinbergen Legacy*, M. S. Dawkins, T. R. Halliday & R. Dawkins (eds.), 75-99. London, Chapman & Hall.
- Dennet, D. C. (1995). *Darwin's Dangerous Idea: Evolution and the Meanings of Life*. New York, Touchstone.
- Fisher, J. & R. A. Hinde (1949). The opening of milk bottles by birds. *British Birds*, **42**:347-357.
- Galef, B. G. (1996). Social enhancement of food preferences in Norway rats: a brief review. In *Social Learning in Animals: the Roots of Culture*, C. M. Heyes & B. G. Galef (eds.), 49-64. London, Academic Press.
- Galef, B. G. J. (1988). Imitation in animals: history, definition, and interpretation of data from the psychological laboratory. In *Social Learning: Psychological and Biological Perspectives*, T. Zentall & B. G. Galef (eds.), 3-25. Hillsdale NY, Erlbaum.
- Gibbs, H. (1990). Cultural evolution of male song types in Darwins medium ground finches, *Geospiza fortis*. *Animal Behaviour*, **39**:253-263.
- Goodall, J. (1964). Tool-using and aimed throwing in a community of free-living chimpanzees. *Nature*, **201**:1264-1266.
- Goodenough, O. R. (1995). Mind viruses: culture, evolution and the puzzle of altruism. *Social Science Information*, **34**:287-320.
- Heyes, C. (1993). Imitation, culture and cognition. *Animal Behaviour*, **46**:999-1010.
- Heyes, C. M. (1994). Social learning in animals: categories and mechanisms. *Biological Reviews*, **69**:207-231.
- Heyes, C. M. (1996). Genuine imitation? In *Social Learning in Animals: the Roots of Culture*, C. M. Heyes & B. G. Galef (eds.), 371-390. London, Academic Press.
- Hinde, R. A. & J. Fisher (1951). Further observations on the opening of milk bottles by birds. *British Birds*, **44**:393-396.
- Jenkins, P. F. (1978). Cultural transmission of song patterns and dialect development in a free-living bird population. *Animal Behaviour*, **25**:50-78.

- Kawai, M. (1965). Newly-acquired pre-cultural behavior of the natural troop of Japanese monkeys on Koshima Islet. *Primates*, **6**:1-30.
- Lefebvre, L. (1995). The opening of milk-bottles by birds: evidence for accelerating learning rates, but against the wave-of-advance model of cultural transmission. *Behavioural Processes*, **34**:43-53.
- Lefebvre, L. & B. Palameta (1988). Mechanisms, ecology and population diffusion of socially learned, food-finding behaviour in feral pigeons. In *Social learning: Psychological and Biological Perspectives*, T. Zentall & B. G. Galef (eds.), 141-163, Erlbaum.
- Lynch, A. (1996). The population memetics of birdsong. In *Ecology and Evolution of Acoustic Communication in Birds*, D. E. Kroodsma & E. H. Miller (eds.), 181-197. Ithaca, New York, Cornell University Press.
- Lynch, A., G. M. Plunkett, A. J. Baker, et al. (1989). A model of cultural evolution of chaffinch song derived with the meme concept. *American Naturalist*, **133**:634-653.
- McGrew, W. C. (1994). Tools compared: the material of culture. In *Chimpanzee cultures*, R. W. Wrangham, W. C. McGrew, B. M. de Waal & P. Heltne (eds.), 25-39, Harvard University Press.
- Mineka, S. & M. Cook (1988). Social learning and the acquisition of snake fear in monkeys. In *Social Learning: Psychological and Biological Perspectives*, T. Zentall & B. G. Galef (eds.), 51-73. Hillsdale NY, Erlbaum.
- Piaget, J. (1962). *Play, Dreams and Imitation in Childhood*. London, Routledge & Kegan Paul.
- Rose, N. (1998). Controversies in meme theory. *Journal of Memetics - Evolutionary Models of Information Transmission*, **2**. <[http://jom-emit.cfpm.org/1998/vol2/rose\\_n.html](http://jom-emit.cfpm.org/1998/vol2/rose_n.html)>
- Sherry, D. F. & B. G. Galef (1984). Cultural transmission without imitation: milk bottle opening by birds. *Animal behaviour*, **32**:937-938.
- Sperber, D. (1996). *Explaining Culture: a Naturalistic Approach*. Oxford, Blackwell.
- Sperber, D. (1999). The future of 'meme': retirement or redefinition? Paper presented at King's College Research Centre Forum, Do Memes Account for Culture?, University of Cambridge, UK.
- Warner, R. (1988). Traditionality of mating-site preferences in a coral reef fish. *Nature*, **335**:719-721.
- Warner, R. (1990). Male versus female influences on mating-site determination in a coral-reef fish. *Animal Behaviour*, **39**:540-548.

Whiten, A., J. Goodall, W. C. McGrew, et al. (1999). Cultures in chimpanzees. *Nature*, **399**:682-685.

Yando, R., V. Seitz & E. Zigler (1978). *Imitation: A Developmental Perspective*. Hillsdale, NJ, Erlbaum.

# **TRANSMISSION OF CULTURAL TRAITS BY EMULATION: AN AGENT-BASED MODEL OF GROUP FORAGING BEHAVIOR**

**By**

**Robert G. Reynolds, Robert Whallon, and Steven Goodhall**

Dept. of Computer Science, Wayne State University

Detroit, Michigan 48202

Museum of Anthropology, University of Michigan

Ann Arbor, Michigan 48104

&

Senior Project Manager

Compuware, Farmington Hills

## **Abstract**

The goal of this research is to assess the impact of Culture on decision-making behavior. Specifically we are concerned whether the emergence of human culture provided humans with an adaptive advantage over non-human primate counterparts in terms of hunter-gathering capabilities. Reynolds has proposed several mathematical models of hunter-gatherer and primate decision making based upon differences in human and primate cultural traits [1]; these were labeled the cultural algorithm and vector voting models respectively.

In this paper an agent-based implementation of the vector voting model using Swarm is presented. Learning takes place in the model by emulation. The performance of this model in a variety of patchwork landscapes is compared with that of a random walk model. The impact that the distribution of food resources has on the outcome is also tested using several different strategies such as equal sharing, round robin, satisfied first, neediest first, and fixed rank order. The results suggest that the vector voting model performs best when certain resource sharing strategies are used (fixed rank order, satisfied first) as opposed to others such as neediest first, and equal sharing. In the latter case it behaves equivalently to a random walk. Thus, when knowledge is not shared equally, it is not adaptive to share resources equally either.

## **1 Introduction**

The goal of this research is to assess the impact of Culture on decision-making behavior. Specifically we are concerned with the question of whether the emergence of human culture provided humans with an adaptive advantage over non-human primate counterparts in terms of hunter-gathering capabilities. Reynolds has proposed several mathematical models of hunter-gatherer and primate decision making based upon differences in primate and human cultural traits [1], these were labeled the vector voting and cultural algorithm models respectively. In the vector voting model each individual's vote was based upon their own knowledge and knowledge was not shared between individuals. The decision made by the group was a consensus based upon the weights and opinions of the members and was modeled after the patterns of interaction seen among primate groups. In the cultural algorithm model the individuals' knowledge was pooled and used by a central decision maker to produce a decision.

The basic context in which these two decision-theoretic models were compared was a two-dimensional cellular space divided into R discrete sub-regions or cells each of unit area. The task facing the model groups was to compute the answers to various spatial predicates or queries about the region based upon the agents' current knowledge. The models were analyzed theoretically and it was shown that the ability to form a collective intelligence through the pooling of knowledge with Cultural Algorithms had some distinct advantages over the vector voting model.

In particular, predicates such as the best direction within a region in which to forage, the direction containing the most resources, was limited by the maximum area over which each individual had knowledge. On the other hand, pooling of that knowledge theoretically allowed a group to make these decisions over the entire region.

However, even in a social system where knowledge is not directly pooled, learning can take place. For example, Tomasello and Call [2] state that there are many similarities between how humans and primates understand their social worlds. Each has its own cultural system. In an extensive survey of primate cognition studies they conclude that, "*all primates live in basically that same type of social world, in which they individually recognize conspecifics and appreciate both the vertical (dominance) and horizontal (affiliative) relationships that hold between group members. They also have the ability to predict the behavior of conspecifics in many situations based upon combinations of cues and insights, and in some cases to affect the behavior of group mates via various social and communicative strategies.*"

They state that the basic difference between primate and human cultures is that in the latter the "*inter-subjectivity of human linguistic symbols- and their perspective nature as one offshoot of this inter-subjectivity- means that linguistic symbols do not represent the world directly, in the manner of perceptual or sensory-motor representations, but rather they are used by people to induce or construe certain perceptual and conceptual situations in order to attend to them in one way rather than another*".

Thus, learning in a primate social system relies heavily on **emulative learning**. An individual watches another perform an action and observes the state changes that result. Learning in this context is directly associated with sensory motor activities relative to objects in the environment. While humans can acquire knowledge in this way as well, they are able to support the **imitative learning** of concepts. With imitative learning "*an individual understands others as intentional agents, like the self, that have a perspective on the world that can be followed into, directed, and shared.*"

The idea is that even when a group makes a decision based upon the knowledge of each individual without pooling, the physical results of that decision can be observed by everyone and learning can take place in an emulative fashion. The question of interest in this paper is what additional behaviors emerge from a group that uses the vector voting approach along with an emulative learning process. Specifically, to what extent do behaviors typically observed within a primate foraging group emerge from the use of the vector voting model? In future studies we will examine the impact that adding communication and abstraction capabilities in the form of Cultural Algorithms has on these patterns of behavior.

Here we develop an agent-based implementation of the vector voting model using the Swarm Simulation environment. In section 2 we describe the theory behind the vector voting model and the emulative learning method that serves as the basis for the primate cultural transmission here. In section 3 the Swarm implementation of the vector voting model with emulative learning is presented in detail. Next, in section 4 we describe the results of running this model in a variety of environments with a variety of social configurations. Specifically, we use our system two sets of experiments. Our first set of experiments is used to determine whether the vector voting model outperforms a baseline model which performs a simple random walk through the landscape. We also observe any generalized behavior patterns that emerge from the model group. The second set of experiments is designed to observe how differences in social interaction effect the performance of the vector voting model. Here, we propose several different ways of distributing the collected resources including approaches based upon dominance and equal sharing. We then attempt to identify which form of distribution works best in conjunction with the vector voting model. Section 5 provides our conclusions.

## 2 Implementing Cultural Transmission in the Vector Voting Model

The vector voting model represents how primate groups might make consensus based decisions. Human hunter-gatherers can also employ this technique as well and one assumes that human cultural systems are able to build upon the capabilities of the vector voting approach.

The motivation for the vector model can be seen in the following quote from Kummer [3, p. 66], *"In choosing the direction of their departure, hamadryas baboons have to rely on information gathered on previous trips. We do not know how or by whom these sites are explored and remembered, but we know that on different mornings different males of the troop strive in different directions. The decision is made by a long process in which most of the males in the troop participate while the rest seem unconcerned. I have already described how a hamadryas troop prepares for departure during its morning rest. The troop performs slow on-the spot movement, changing its shape like an undecided ameba. Here and there, males move a few yards away from the troop and sit down, facing in a particular direction away from the center."*

*Pseudopods are generally formed by the younger males and their groups, until one of the older males in the center rises and struts toward one of the pseudopods. At this, the entire troop is alerted and begins to depart in the indicated direction. Detailed observation reveals that two male roles are involved in leading the hamadryas troop: the younger initiators who "propose" certain directions and the deciders who choose among the proposed directions."*

One of the basic decisions that a foraging group can make is the direction in which to forage during the day. It was shown in [4] that it was not possible for a vector voting model to always select the direction with the maximal amount of resources within a region R. However, theoretical limits aside how does the vector voting model actually perform when coupled with emulative learning, the type of learning that is frequently observed in primate cultures.

In emulative learning the observer makes a cognitive connection between what action is performed and the state changes it produces. For example, a primate can observe another rolling over a log and exposing a number of insects. That action can be viewed as producing a state

change which can be stored in memory. In our model, the result of a directional decision produces a trajectory through the landscape. As a result of that path each individual has an opportunity to acquire food. In each cell that the group enters the resources there are divided among group members by various strategies such as priority or fixed order access or equal sharing of resources.

At the end of the day each individual can store a memory, not of the decision, but of the result. An individual does this by associating a visible landmark with the degree to which they were fed (satisfaction scale) the day they saw that landmark. A memory can more than one landmark attached to it and different individuals in the group can associate different landmarks with their memory. The memory made by an individual is a function of the food resources that they received that day. Each individual has a maximum number of memories that it can store and a memory is forgotten after a certain number of days (memdepth) unless it is used again.

Emulative learning using memories associated with the icons encountered relative to a group's previous decision can now be used to impact future decisions. Here, our region  $R$  has a number of landmarks ( $\text{numLandmarks}$ ) which are distributed through the space. The cellular space is of size  $N$  by  $M$  and a cell can have at most one landmark assigned to it. Each day after the group decides on a direction to move based upon the memories associated with the landmarks currently visible from their location (visibility).

Each individual effectively pools the satisfaction scores for the memories associated with the visible landmarks in each of the eight directions. Here the scores are represented by a preference scale from -5 to +5 where the + direction represents satisfaction and - dissatisfaction. The direction with the highest score is the direction of choice for an individual. That choice is then weighted by their status in the group. Each member then moves in the direction of their choice and is observed by the others. The group moves in the direction which achieves the highest consensus as observed by a mediator.

Group size can change based upon the extent to which individuals are fed. Individuals who have not gotten sufficient resources over a given period die and a group is removed when all of each individual have died. On the other hand if the group has been able to feed all of its members over a given period it can add a new individual up to a maximum group size. When it reaches that size it can fission into a second group.

### 3 Model Description

The **Band model** is an agent based model that simulates the foraging behavior of groups of hunter-gatherers using the vector voting method described above. The bands move in a rectangular space, looking for food. The size of the band can increase or decrease over time based on its success in acquiring resources. Fission or fusion of a band takes place when a maximum or minimum number of individuals is produced.

The model supports a variety of food distributions, agent characteristics, and foraging behaviors. Food distributions include completely random distributions, distributions based on mathematical

functions, and distributions that simulate a variable distribution based upon local conditions such as terrain, water supply, etc.

Each band member can form memories of their individual successes and failures that they can use to guide their future search for food. At the beginning of each model day, each individual evaluates these memories and selects a preferred direction of travel or vector. The members of a band are polled and the direction preferred by a majority of the members is selected according to the vector voting model given in the previous section.

The band then moves in the selected direction, collecting food as they go. At the end of the day the collected resources are distributed according to the consumption strategy used, each member evaluates how well it was fed, and associates that memory with one or more visible landmarks on the terrain. Several consumption strategies are supported here. They are fixed rank order, round-robin, equal sharing, neediest first, and satisfied first. Each will be described in more detail later.

During the execution of the model, detailed information is written to files for subsequent analysis. This information includes:

- Band and individual survival rates
- Individual satisfaction rates
- Range of movement of the band
- Distance between bands
- Longevity of bands
- Band size

These data are then analyzed to develop the graphs that accompany this report. The project is implemented using Swarm, a multi-agent modeling package developed at the Santa Fe Institute ([www.sfi.edu](http://www.sfi.edu))

### **3.1 The Basic Agent Model**

The system models the activities of hunter-gatherers searching for food in a defined landscape, these constitute our agents here. The hunter-gatherers are organized into bands that travel together in search of food. Their search is guided by fixed landmarks that are distributed across the landscape. Individual members of the bands form memories of their past successes and failures in finding food. Each memory is associated with one or more landmarks that are nearby when the memory is formed. Each memory records the degree to which the individual was successful in obtaining enough food on the day that the memory is formed.

Each day, every band selects a direction of travel based on the memories of the individual members of the band. Bands can move basically in a straight line in any of 8 compass directions (N, S, E, W, NE, NW, SE, SW). The distance traveled is limited to a maximum number of squares in the landscape grid. A limit of 4 squares has been used for all trials used in this paper. Bands move in pseudo-parallel, one square at a time, in rotation among the bands. Each band continues to move in that direction until sufficient resources are acquired to feed all band

members, an object is encountered, or the maximum distance limitation (4 squares) is met. A band will alter its trajectory in order to avoid an obstacle. Here the obstacles are either static (the region boundaries) or dynamic (other bands). It is assumed here that two bands cannot occupy the same cellular space at the same time. If a band encounters an obstacle a new direction is selected, based on the same vector voting mechanism used at the beginning of the day, in order to avoid the obstacle.

At the beginning of the day, each band member reviews the memories associated with any landmarks that are currently visible and forms a weighted sum of memory scores for each possible direction of travel. Weighting is based on the interval that has elapsed since the memory was formed or reinforced and the distance to the landmark. Old memories eventually expire. The direction with the highest score becomes that individual's preferred direction of travel. The direction in which the band will travel is selected by summing up the preferred direction for each voting member. The preferred directions of all adult members are scored and the band will travel in the direction that receives the most votes. One constraint that is placed on the selected direction is that no band can reverse its direction of travel from the previous day. If the reverse direction is selected, the second choice is used. The band then travels together in the selected direction with its members collecting food as they go.

Several algorithms are implemented in order to regulate the distribution of the collected food among the members of the band as they collect food in each new cell. These distribution algorithms include round robin, equal sharing, fixed rank order, neediest first, and a satisfied first strategy. Movement for the day stops when the band has traversed the maximum number of squares or when all members of the band have consumed as much food as they can for the day. Each member of the band has a Minimum Daily Requirement for food. A provision is made in the model for an individual to consume more than one day's requirement. This permits an individual that has not been fully fed in the past to "catch up". It also permits individuals to "get ahead" so that they will survive subsequent food shortages. An individual's failure to obtain the MDR on a given day produces a cumulative deficit in nutrition. If the deficit exceeds a selected limit, the individual dies.

At the end of each day, each member of the band forms memories based on how much food they obtained during the day. These memories are associated with one or more landmarks that are visible from the location at the end of the day. Both the visibility of landmarks and the visual abilities of individuals are controlled by model parameters. Each individual selects a random direction and records the landmarks that are visible in that direction. An alternative strategy of using the nearest landmarks for all individuals failed to produce enough variability of experience to drive the voting model.

### **3.2 The Food Distribution Patterns**

The pattern of food distributions are an important part of the agent's environment here. Initially we tried various mathematical distributions that were intended to produce gradients of food availability. For reference purposes, we also ran trials against a completely random distribution of food on a cell by cell basis. In a completely random distribution each cell is assigned a resource value between 0 and the maximum quantity allowed in the cell selected from a uniform

random distribution. Recently, we have settled on a "patchwork" distribution that is intended to simulate a semi-arid environment in which food supply varies over the entire space but is uniform within localities. Our implementation of this distribution ties each locality to a landmark located at the center of the local area. Regardless of the distribution, as a group moves through a cell it is able to remove resources from the cell. This amount is generated using a uniform random distribution between the current cell value and 0. After the group has moved on, the cell can "regrow" its resources in a linear fashion over time. The parameters for regrowth are provided at run time.

The patchwork distribution represents or attempts to emulate the attributes of a semi-arid environment. This distribution is generated by creating uniform areas of food availability around each landmark on the landscape, perhaps analogous to desert oases. This distribution is created according to the following procedure;

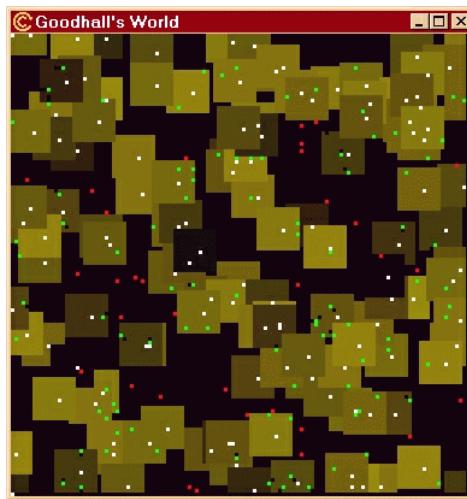
Landmarks are distributed randomly on the landscape

Each landmark is assigned a random food value between zero and the maximum permitted quantity of food in a cell.

An area of fixed dimensions around the landmark is populated with food at that value.

If the areas around two landmarks overlap, the food level will be set at the value of the last landmark processed.

The following image shows a patchwork distribution:



**Figure 1.** An example patchwork distribution of resources in the model

### 3.3 The Swarm Model Implementation

The simulation was written in objective C and runs within the Swarm simulation environment. The Swarm simulation environment provides a collection of software objects and associated

methods or algorithms (routines that manipulate these objects) to support the execution of our band model. One of the advantages of this approach is that the basic utilities that are in common to many agent based simulations are pre-written and we only need to supply the object-oriented code for the model. The fundamental swarm object within which our simulation runs is called the **observer space object**. This object controls the graphics interfaces, scheduling of agent activity, and the performance of basic model control functions such experiment length, and run initialization. It provides a framework in which the behavior of the simulation can be controlled and observed, hence the name observer space object. Table 1 gives the attributes or variables associated with the observer object along with the algorithms that it uses to control the execution of our simulation model here.

Object attributes	
Attribute	Description and usage
Display Interval	Controls how often the GUI display is updated. Set to 1 for all trials.
Experiment Duration	Number of days that the model will run before terminating.
Schedule Objects	Swarm scheduling control objects
Display Objects	Swarm display control objects

Object algorithms	
Attribute	Description and usage
Create model	Create the model objects
Run model	Execute the model

**Table 1.** The Observer Space is a swarm object that supports the graphical interfaces, scheduling, and control functions for the model.

Running within the observer object is our model. The key object in our model is the model space object. The **Model Space** is a rectangular grid depicting the agent's environment onto which other model objects are mapped. Its attributes and algorithms are described in table 2. Individual bands are placed into cells in the model space as they move about the landscape. Other objects, such as Landmarks, and Food Cells are overlaid on the Model Space and displayed on the common interface provided by the observer object. The dimensions of the grid are set by a pair of parameters. For most trials, a 128 by 128 model space has been used. Practical considerations of the data analysis software preclude large dimensions at this time.

The band model object attributes contain statistics that describe overall model behavior such as the number of bands that are dead, the life span of a band, the number of elapsed days in the model among others. In addition there are attributes that point to a lists of the current bands, the individual agents that comprise them, the current landmarks, and resource cells respectively. These constitute the basic model objects. In subsequent sections we will discuss the attributes and algorithms associated with each. There are also pointers to files that will be used to retrieve model data and archive results. Finally, the model parameter pointer references the list of model parameters that can be set upon model initialization and govern the interaction of the model objects.

The algorithms for the model space object are in charge of getting the model parameters, initializing the model according to those parameters, updating the model statistics after each model event, and logging diagnostic information about model performance that can be useful in debugging the model.

Object attributes	
Attribute	Description and usage
Day Number	Counter of the current day in the model.
Number of Dead Bands	Current count of the number of dead bands
Number of bands that have died.	Count of the total number of bands that have died since the beginning of the model.
Sum of band age at death	Sum of the age of all dead bands at the time of their death. Used to calculate average mortality.
Average age of band at death	Average age of bands that have died or are still living.
List Food Spaces	Pointers to the Food Space Objects
List of Landmarks	Pointer to the list of landmarks.
File object pointer	Pointer to the file control object
Maximum distance	The calculate diagonal distance across the landscape
Model parameters	Shared elements that can be set at the start of the simulation run to control the operation of the model.

Model Parameters	
Parameter Name	Parameter Description
BandSize	maximum size of each band (number of members)
BandSplit	size at which band can split

BandStart	size of each band at the beginning of the run (number of members)
BaseMDReq	base minimum daily food requirement for an adult
BounceMode	Behavior mode when a band reaches the edge of the space or collides with another band 0 – reflex bounce 1 – vote on a new direction
CAAcceptanceInterval	Cultural Algorithm Acceptance Interval (days) this is the interval at which the CA component accesses the data in the model space
CAInfluenceInterval	Cultural Algorithm Influence Interval (days) this is the interval at which the CA component provides feedback to the model space
CAMode	Selects the CA algorithm to be used,
caption	Trial Caption
consumeMode	mode of consumption during foraging
daysAhead	Number of days of consumption that an individual can stockpile to recover from shortfalls or to get ahead
daysToDie	number of days without MDR before a guy dies
decayDays	number of days for a dead band to decay and disappear from the display until this interval expires, the square occupied by the band is blocked to other bands
diagBand	A band number to trigger diagnostic outputs that are limited to a selected band
diagnosticDest	destination for diagnostic outputs 1 => console 2 => log, 3 => both
diagnosticInterval	diagnostic interval in days, interval at which diagnostic outputs will be written
diagnosticMask	diagnostic mask, a bitmap selecting various classes of diagnostic outputs
flipMode	select a yes no parameter to be flipped for each trial
foodDistFactor	First computational factor for food distribution. Used differently by each mode. Not always used.
foodDistFactor2	Second computational factor for food distribution. Used differently by each mode. Not always used.
foodDistMode	food distribution function
foodDistRange	food distribution range in squares usage varies by algorithm currently used only for mode 6, patchwork around Lankmarks to control the size of the patches
food.MaxValue	The maximum quantity of food that can exist in a single cell
forageMax	number of squares to forage in Mode 1

forageMode	forage mode 0=en mass while traveling 1=individual after traveling (unimplemented)
lmDistMode	landmark distribution mode
lmPerMemory	controls how many landmarks are attached to a memory when it is formed
logDataInterval	interval to write to data log (days)
memDepth	memory depth in days, number of days after a memory is formed before it is forgotten
memMode	Memory formation mode
numBands	number of bands to generate
numLandmarks	number of landmarks to generate
quitAfter	Terminate the mode after this many steps. Results in program failure. Superseded by batch mode operation.
randomDays	number of days of random movement before using memory (during this interval, bands do not starve)
regenRate	rate of food regeneration per day (<= 1.0)
reincarnateDays	number of days before a dead band is reincarnated
reproduceDays	Number of days where the band is fully fed before it can add a member.
searchMode	search mode
seedProb	probability of food in a cell (random distribution) also controls the minimum amount of food not currently used in other modes
starveMode	Mode for determining starvation
stepsPerDay	maximum number of steps to move in one day
stopEvery	interval (days) at which the model will pause
trialName	file name prefix for the trial (yyyymmdd_hhmmss) this is not specified externally, but is generated and exported at run time
visibility	default landmark visibility in squares, for how many squares is a landmark visible
vision	default vision in squares, how many squares can a Guy see a landmark
worldXSize	X dimension of the model space
worldYSize	Y dimension of the model space
ZoomFactor	swarm zoom factor, 0 = automatic 1-

Object algorithms	
Attribute	Description and usage
Get methods	Obtain model parameters and control variables
Set methods	Set model parameters and control variables
Diagnostic methods	Log diagnostic information about model execution
Tracking methods	Maintain global information on model execution, such as number of living bands, etc.

**Table 2.** The Band Model space object attributes and associated algorithms

Now each of the model objects that interact within the grid world will be described. The basic objects are: **the food space cells**, **the landmarks**, and **the bands**. The bands are composed of individual **member agents** and each member agent has a list of **memories**. We will now describe each of these objects in turn starting with the top level model objects.

The **Food Space** is a rectangular grid that overlays the Model Space. Each cell in the food space contains a quantity of food that can be consumed. Two supplemental food spaces are defined to support variation in the food supply. The first of these contains the original values calculated when the Food Space was populated. This is used to control the re-growth of food after it has been consumed. The second is used to support seasonal variation of the food supply.

Object attributes	
Attribute	Description and usage
Current food values	Array containing the amount of food available in each cell of the landscape.
Maximum food values	Array containing the permitted maximum amount of food for each cell of the landscape.

Object algorithms	
Algorithm	Description and usage
Seed food space	A group of methods that seed the food space according to the selected distribution function
Regenerate	Grow more food to make up for the amount of food consumed. Executed daily.

Vary food	Change the available food according to a seasonal cycle.
Get methods	Obtain model parameters and control variables
Set methods	Set model parameters and control variables
Diagnostic methods	Log diagnostic information about model execution

**Table 3.** The object attributes and the algorithms for the food space object. Its attributes include the current and maximum food values for all grid cells.

Just as the food cells are overlaid onto the model grid, so are the landmarks. **Landmarks** are randomly distributed throughout the landscape. They are commonly observed landscape features that can be used as reference points in the decision making of the bands. While landmarks can be randomly placed within the landscape, certain landmarks were placed in certain grid positions to ensure that the entire region was potentially exploited by the bands. For example, a Landmark is placed in each corner square of the landscape. This choice was based on early observation that bands could not navigate into the corners of the space unless there was a Landmark towards which they could move.

Landmarks are used to support memory formation and band navigation. Each memory references one or more landmarks that are visible from the location where the memory was formed. All navigational choices made by individuals are made on the basis of moving towards or away from a landmark.

Each landmark is an object whose state is described in terms of its position on the grid, the maximum distance from which it can be viewed, along with a list of memory scores that have been produced by individuals within view of the landmark. The basic algorithms for landmarks are their distribution (randomization), and the collection of statistics about their use in decision making. Currently landmarks are viewed as static objects so degradation or improvement in their visualization appears over time.

Object attributes	
Attribute	Description and usage
Position	X and Y coordinates of the landmark
Visibility	Distance at which the landmark is visible
Landmark scores	Information about the quality of memories associated with the landmark.

Object algorithms	
Attribute	Description and usage
Distribute landmarks	Distribute the landmarks more or less randomly on the space.
Landmark statistics	A series of methods that maintain current and prior history on landmark performance in memory searches.

**Table 4.** The object attributes and algorithms for the Landmark objects.

While the positions of both food cell and landmarks do not change during the course of the simulation, the bands are able to move about in search of food. **Bands** are groups of individual hunter-gatherer agents. The members of a band move and forage together within a cell. At the beginning of a simulation run, bands are placed at random on the landscape. Since the carrying capacity of the environment can be estimated, the number of bands used will influence the selective pressure on the decision making process here. In all of the runs described here, more individuals are initially present than can be supported by the environment. This was done in order to assess the robustness of the vector voting model under duress.

Each band starts with a specified number of members. This number can be reduced by the death of members or increased by the addition of new members. Both the addition and deletion of members results from the foraging activity. If the group fails to meet the MDR for some of its members over a period of time, deaths can result. If all members of a band die, the band is removed from the model. New members can be added to a band whenever all members of the band have received their minimum daily requirement for a specified number of days. Only one member is added, and the interval starts over. When the band size reaches a maximum, fission can take place where the new band is placed in the location of a deceased band.

Since band decision making is at the heart of the model, the band object is the most complex model object here. Table 5 gives its attributes and its algorithms. A number of band attributes relate to its current status in terms of creation, number of members and a pointer to a list of its members, current direction of movement, and number of members fed among others. Random days is an attribute that determines how many days the group will move at random before decision making via vector voting begins. The random movement simulates the discovery of the new territory by the band and allows them to record memories about their experience which serve as the fuel for the decision making process. Random days is an attribute that determines how many days they forage randomly until decision making begins. For the completely random model, the number of days is equal to that of the simulation length. The band algorithms include methods for initializing a new band, starting a day, finishing a day, updating its members daily information among others.

<b>Band object attributes</b>	
<b>Attribute</b>	<b>Description and usage</b>
Size	The maximum permitted size of the band.
Member list	Points to the guys in the band
Random days	The number of days of random motion to be performed before using the search algorithm. Zero for all recent trials.
Total Daily Requirement	The total amount of food required by the band on a single day
Band dead	An indicator that the band has died
Band decayed	An indicator that the band has been dead long enough to be removed from the display and also long enough to be reincarnated.
Day created	The model day on which the band was created or reincarnated
Direction	The current direction of movement for the band
Days fed	The number of days during which all members of the band have received their MDR.
Have eaten	Indicator of whether the members of the band received the MDR today
Number dead	The number of members in the band who are currently not alive.
Number in and	The number of members in the band who are currently alive
Consumption order controls	Controls the order in which members are fed.
Contact information	Counters of the number of times the band collided with or was next to another band
Day band died	The day on which the band died. Used to control reincarnation of bands.
Range date	Deflections in the X and Y axis during a measurement period.

<b>Band object algorithms</b>	
<b>Algorithm</b>	<b>Description and usage</b>
Get methods	Obtain model parameters and control variables

Set methods	Set model parameters and control variables
Diagnostic methods	Log diagnostic information about model execution
Maintain count of alive and dead	Methods to maintain the current living population of the band
Select direction	Method to obtain and tally votes from the band members
Step	Method to perform one step's processing for the band, including movement and consumption
Start day	Method to control beginning of day processing, including selection of direction
End day	Method to control end of day processing, including memory formation by band members
Reincarnate	Reincarnate a band that has previously died.
Forage	Consume food at the end of a day's movement (unimplemented future function.)
Maintain movement and contact statistics	Group of methods to maintain statistics on range of movement and contact with other bands.
Reproduce	Method to add a member to a band during good times.

**Table 5.** Band object attributes and methods. The band object is the most complex of the model objects. In particular there are a number of algorithms associated with its behavior.

Agents are individual members of a band. They are the second most complex model objects.

Each **member object** forms memories based on how much food it receives on a given day. A number of variables describe the state of an agent including their age and gender, the MDR for their category, where they are located, the landmarks that are visible from their current location, the amount of food they have gotten today, the memories that they have of their experiences, and how they have voted in terms of the bands decision making. The algorithms for an individual agent include making memories, consuming food, starting the day, ending the day, their vote in the foraging decision, and other algorithms to compute agent statistics. The description of the agent object attributes and algorithms are given in table 6 below.

Object attributes	
Attribute	Description and usage
Adult	Flag indicating whether the member is an adult or a child
Gender	Flag indicating whether the member is male or female

MDR	Minimum Daily Requirement of food
Memory list	List of memories
Vision	Number of square in the landscape that the individual can see a landmark. Not used in any current models.
Food consumed today	Amount of food consumed today
Days without food	Number of days during which the member has not received the MDR
Dead	Indicates whether the individual is dead
Food deficit	Cumulative shortfall of consumption against MDR
Number of Memories	Current number of memory entries for this individual. Used to gather model statistics
Last vote	Holds the last direction the member voted for. Use to calculate dissatisfaction.
Total dissatisfaction	Cumulative total of dissatisfaction. Used to determine consumption order.

Object algorithms	
Attribute	Description and usage
Get methods	Obtain model parameters and control variables
Set methods	Set model parameters and control variables
Diagnostic methods	Log diagnostic information about model execution
Vote for a direction	Decide which way the member wants to travel
Make a memory	Create a memory based on current location and today's consumption experience.
Start of day	Perform start of day processing, including clearing of consumption counters
End of day	Perform end of day processing, including calculation of satisfaction
Consume food	Consume food up to the MDR
Consume more food	Consume food beyond the MDR if it is available, up to the

	permitted multiplier
Forage	Look for food at the end of a day's movement (unimplemented future function)
Food needed	Calculate the current need for food. Used to control consumption order
Check whether fed	Compare the amount consumed with the MDR
Reproduce	Change status from dead to living, resetting all controls

**Table 6.** The attributes for the agent objects and their algorithms.

Just as band objects are contained in the model grid, and agent objects make up the bands, each agent has a list of **memories** that guides their decision making. At the end of each day's movement, every member forms a memory. The content of this memory is a score that indicates how well the individual was fed on that day along with the landmarks associated with the memory by the individual (good), the x and y location in which the memory was formed, and the list of landmarks associated with the memory by the individual. The basic algorithms for this object are the addition and deletion of memories. This constitutes the emulative learning component of the model. The description of the memory object attributes and algorithms are given in table 7.

Object attributes	
Attribute	Description and usage
Get methods	Obtain model parameters and control variables
Set methods	Set model parameters and control variables
Diagnostic methods	Log diagnostic information about model execution
Active indicator	Indicates whether the memory is active or expired
Directions	Direction of movement into and out of the square where the memory was formed (unimplemented future function)
Good	Score reflecting the food consumption associated with the memory –5 – obtained no food +5 – obtained as much extra food as was allowed.
Landmark list	List of landmarks associated with the memory. Used to select memories for navigation.
Position	X and Y position at which the memory is formed. This is

	maintained but is not used when interpreting a memory.
--	--

Object algorithms	
Attribute	Description and usage
Get methods	Obtain model parameters and control variables
Set methods	Set model parameters and control variables
Diagnostic methods	Log diagnostic information about model execution
Reincarnate	Clear memories as part of the band reincarnation process
Add a landmark	Associate a landmark with a memory

**Table 7.** A description of the memory object attributes and algorithms

This concludes the basic description of the objects used to support the Band simulation model in the swarm environment. In the following section we use the simulation model to ask various questions about the capabilities of the vector voting model when used in conjunction with emulative learning.

## 4 Experimental Results

The vector voting model augmented with emulative learning was applied to a variety of different environments, each with a different resource distribution patterns. The goals of these experiments are four-fold:

- To compare the model to a base line random walk through the environment.
- To observe any emergent patterns of foraging behavior that correspond to those exhibited in primate and hunter-gatherers groups.
- To observe the relative survivability of groups in terms of various cognitive and social parameters, e.g. the number of memories an individual can have.
- To observe the impact that resource sharing has on the performance of the vector voting model. That is, what forms of sharing are best able to support the vector voting decision making model are which make the model less robust.

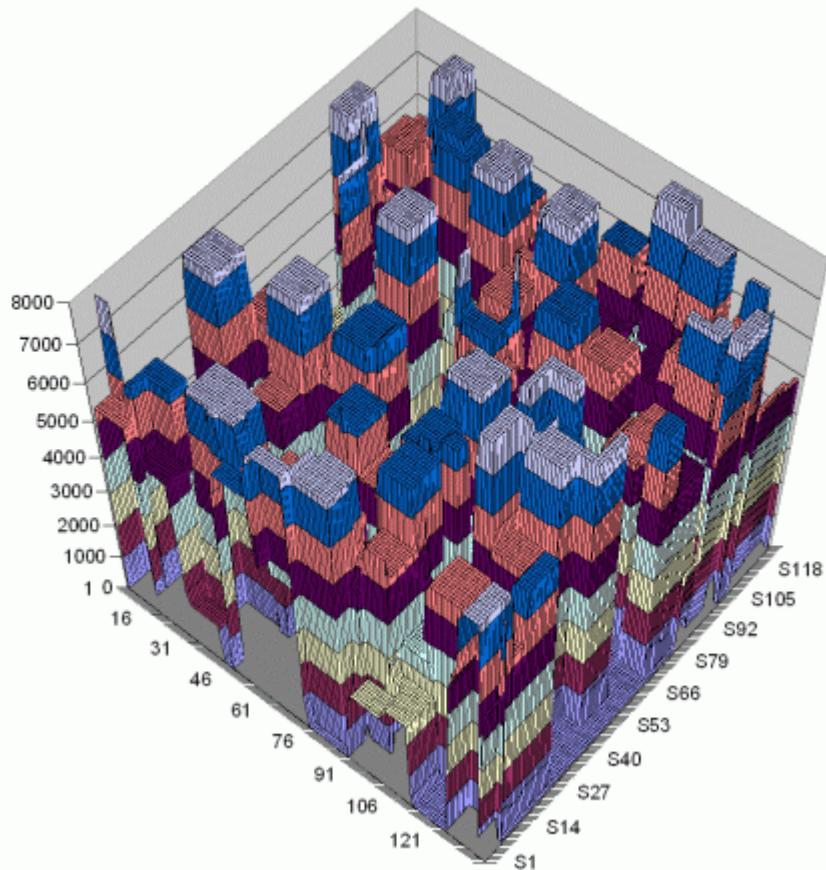
While a number of experiments have been conducted we will summarize some of the results here. This section describes two sets of experiments that were performed during the initial stages of the project. The first set of experiments is intended to test the hypothesis that the vector voting strategy performs better than a random search for food in certain environments. In addition, we wanted to determine what specific patterns of behavior emerge when vector voting is used and how these patterns relate to those patterns observed within primate and human hunter-gatherers. The second set of experiments compares the effect that the sharing of resources among band

members has on the overall performance of the vector voting model as well on the specific patterns of foraging behavior that emerge. That is, which sharing paradigms are most supportive of the vector voting model, and which paradigms are not.

In general, a set of 10 trials has been used for each of the two experiments. For each trial, a pair of simulations was run. The same random number seed was used for each of the trial pair. For example, for the first set of experiments the trial pair will be a vector voting model and a random walk model. The results of each series are averaged to produce a result set for analysis. The food and landmark distributions are held constant for a group of trials so that scoring of landmarks and interpretation of patterns of movement can be compared.

#### 4.1 Random Walk versus the Vector Voting Model

A series of 10 trials was run where every band performed a random search for food. An equivalent set of trials was run against the same food and landmark distributions using the vector voting model to regulate group decision making. All simulations are run using the patchwork food distribution and the set of model parameters given below.



**Figure 2.** An example of the patchwork distribution used in the runs

Parameter Name	Parameter Description	Parameter Value 1	Parameter Value Description 1	Parameter Value 2	Parameter Value Description 2	Mismatch
bandSize	maximum size of each band (number of members)	6		6		
bandSplit	size at which band can split	0		0		
bandStart	size of each band at the beginning of the run (number of members)	3		3		
baseMDReq	base minimum daily food requirement for an adult	1500		1500		
beliefOption	Belief Space Option - unimplemented	0		0		
bounceMode	Behavior mode when a band reaches the edge of the space or collides with another band	1	make new decision on bounce	1	make new decision on bounce	
CAAcceptanceInterval	Cultural Algorithm Acceptance Interval (days) this is the interval at which the CA component accesses the data in the model space	10		10		
CAInfluenceInterval	Cultural Algorithm Influence Interval (days) this is the interval at which the CA component provides feedback to the model space	100		100		
CAMode	Selects the CA algorithm to be used,	1	basic landmark history	1	basic landmark history	
caption	Trial Caption	Memory Based Search		Random Search		YES
consumeMode	mode of consumption during foraging	0	fixed order	0	fixed order	
daysAhead	Number of days of consumption that an	2		2		

	individual can stockpile to recover from shortfalls or to get ahead					
daysPerYear	Number of days per year for seasonal variation	0		0		
daysToDie	number of days without MDR before a guy dies	7		7		
decayDays	number of days for a dead band to decay and disappear from the display until this interval expires, the square occupied by the band is blocked to other bands	7		7		
diagBand	A band number to trigger diagnostic outputs that are limited to a selected band	0		0		
diagLM	Diagnostic Landmark Number	0		0		
diagnosticDest	destination for diagnostic outputs 1 => console 2 => log, 3 => both	0		0		
diagnosticInterval	diagnostic interval in days, interval at which diagnostic outputs will be written	10		10		
diagnosticMask	diagnostic mask, a bitmap selecting various classes of diagnostic outputs	0		0		
flipMode	select a yes no parameter to be flipped for each trial	0	No Parameter Flipping	0	No Parameter Flipping	
foodDistFactor	First computational factor for food distribution. Used differently by each mode. Not always	0.5		0.5		

	used.					
foodDistFactor2	Second computational factor for food distribution. Used differently by each mode. Not always used.	0.5		0.5		
foodDistMode	food distribution function	6	patchwork around landmarks	6	patchwork around landmarks	
foodDistRange	food distribution range in squares usage varies by algorithm currently used only for mode 6, patchwork around Lankmarks to control the size of the patches	12		12		
foodMaxValue	The maximum quantity of food that can exist in a single cell	16384		16384		
foodVarFactor1	Food variation factor 1	0		0		
foodVarFactor2	Food variation factor 2	0		0		
foodVarMode	Vood Variation Mode	0		0		
forageMax	number of squares to forage in Mode 1	0		0		
forageMode	forage mode 0=en mass while traveling 1=individual after traveling (unimplemented)	0		0		
goodMode	Mode for calculating good scores	0		0		
lmDistMode	landmark distribution mode	0	random (seeded corners)	0	random (seeded corners)	

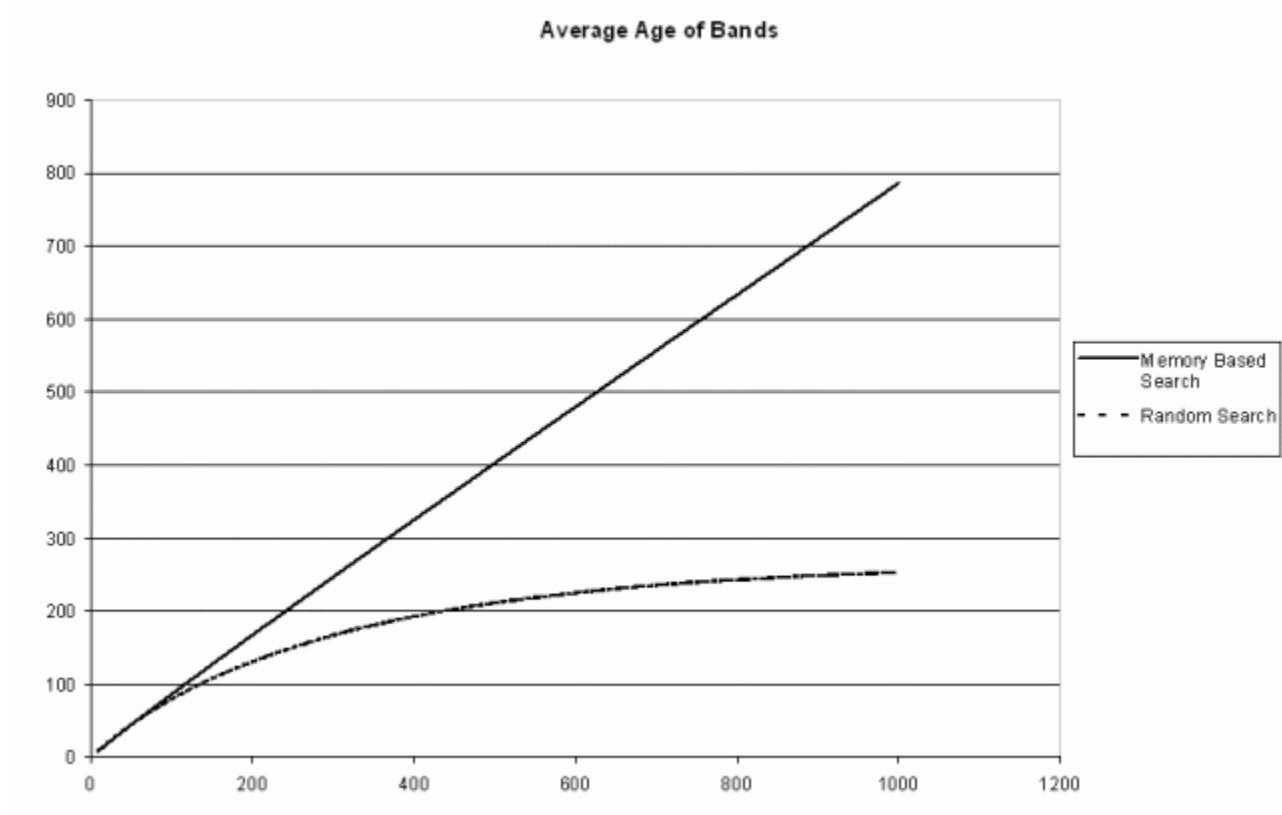
lmPerMemory	controls how many landmarks are attached to a memory when it is formed	2		2		
logDataInterval	interval to write to data log (days)	0		0		
memDepth	memory depth in days, number of days after a memory is formed before it is forgotten	64		64		
memMode	Memory formation mode	0	use n closest landmarks in any direction	0	use n closest landmarks in any direction	
numBands	number of bands to generate	128		128		
numLandmarks	number of landmarks to generate	128		128		
quitAfter	Terminate the mode after this many steps. Results in program failure. Superseded by batch mode operation.	0		0		
randomDays	number of days of random movement before using memory (during this interval, bands do not starve)	0		0		
regenRate	rate of food regeneration per day (<= 1.0)	0.05		0.05		
reincarnateDays	number of days before a dead band is reincarnated	9999		9999		
reproduceDays	Number of days where the band is fully fed before it can add a member.	10		10		
searchMode	search mode	1	memory based vector voting	0	random	YES

seedProb	probability of food in a cell (random distribution) also controls the minimum amount of food not currently used in other modes	1			1		
starveMode	Mode for determining starvation	1	deficit > MDR * days to die	1	deficit > MDR * days to die		
stepsPerDay	maximum number of steps to move in one day	4		4			
stopEvery	interval (days) at which the model will pause	100		100			
trialName	file name prefix for the trial (yyyymmdd_hhmmss) this is not specified externally, but is generated and exported at run time	20010213_2015 44		20010213_2002 22		YES	
visibility	default landmark visibility in squares, for how many squares is a landmark visible	32		32			
vision	default vision in squares, how many squares can a Guy see a landmark	32		32			
voteMode	Voting mode (unimplemented)	0		0			
worldXSize	X dimension of the model space	128		128			
worldYSize	Y dimension of the model space	128		128			
zoomFactor	swarm zoom factor, 0 = automatic	0		0			

**Table 8.** The run parameters for the experiments used here

As we might expect, a memory directed search performs significantly better than random movement in almost all food landscapes. The one exception is that if the food distribution is

completely randomized, there is no advantage to the memory directed search. However, in all of the examples given here, the global distribution of resources is random, but the local distribution of resources is patterned. The question is whether the vector voting model can outperform the random walk model even when there is just local patterning. If it can, then it should do even better when region-level patterns are present. The following diagrams compare metrics averaged over the 10 trials between the random walk (random search) and vector voting search (memory based search) models for the locally patterned patchwork examples.

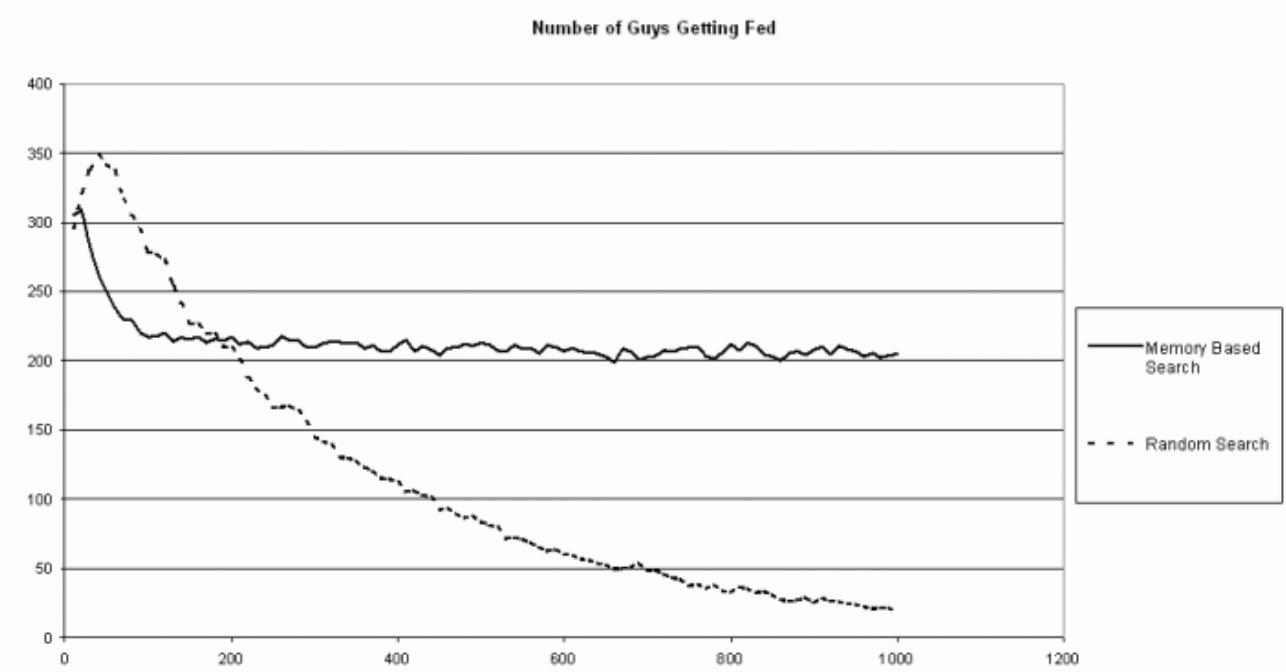


**Figure 3.** The average age of model bands over a 1200 day simulation for the vector voting (memory based) and random walk (random search) models

Figure 3 shows the average survival age of bands during the execution of the model. This calculation is an average of the age of bands at the time they died as well as the age of each surviving band. The curve for the random model flattens out over time indicating that the age of current bands is contributing less and less to the sum. In other words, the number of surviving band is approaching 0. On the other hand, the age of the surviving bands for the vector voting model is producing a relatively constant increment to the sum over the 1000 model days. That is, a consistent percentage of bands are surviving year after year.

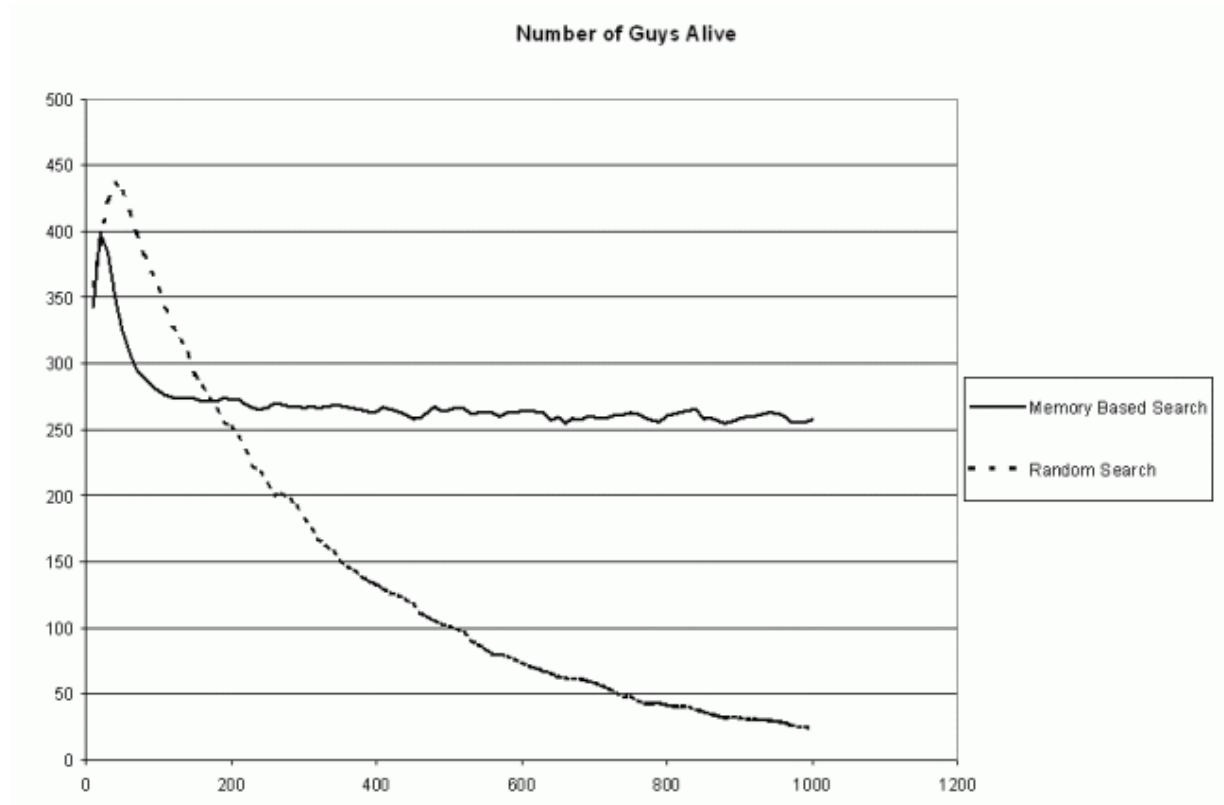
Figure 4 shows the number of individuals getting fed over all bands for each year. The carrying capacity for the environment is approximately 300 individuals. Our initial population was 384, slightly over the maximum amount that we estimated could be supported by the environment on a continuous basis. So, it provides some initial selective pressure on the bands. Initially, the

foraging is random in order to gather information for use in decision making. Since all of the food cells are at the maximum at onset the groups are able to increase the number of individuals getting their maximum daily requirements. However, once foraging has started to deplete the resources, the number of individuals getting fed tends to converge to around 200 for the vector voting model, about 100 less than the total carrying capacity. By the addition of more landmarks and the proper selection of the sharing algorithm, the system can plateau at even higher levels, approaching 300 as the maximum plateau. The number of individuals fed by the random walk model continues to decrease approaches 0 as an equilibrium point. Although it is clear that the environment can support more than that, the system has not been able to acquire sufficient information to exploit those patterns.



**Figure 4.** Average number of band members getting fed in the region over 10 runs each for the vector voting and random walk models.

Figure 5 shows the number of band members that were alive on each day of the model execution. Notice that in both cases the numbers of individuals alive exceeds the number of individuals getting fed. For the vector voting model, this number tends to plateau out around 260 here, about 40 less than the carrying capacity. The key is that in any one point in time not all individuals are getting fed, but enough get fed frequently enough in order to allow the overall population to plateau at a higher level than that for those getting fed. The random model again approaches an equilibrium of 0 as the difference between the number of those getting fed and those alive becomes 0 as does the population.



**Figure 5.** The average number of individuals alive for each model year in the vector voting (memory based) and random walk (search) models for 1200 days.

The results above demonstrate the effectiveness of the memory based search algorithm in finding food in a harsh but consistent environment. The vector voting model is able to exploit sufficient local patterning of resources in order to produce a suggested equilibrium population of around 260, just under carrying capacity. All of the bands without any memory based decision making died out even though the environment was able to support a proportion of them.

#### 4.2 The Effect of Resource Sharing On Vector Voting Model Performance

In this section we perform a series of 10 trials each using 5 different resource distribution models. These were:

**Fixed Order** - the members of a band are fed in the same order on each day. The first member eats as much as it needs to reach the MDR before the second member eats anything, etc.

**Equal shares** - each member of a band receives a portion of the food that is based on the individuals MDR. Each member with the same MDR gets the same amount of the collected resources.

**Neediest first** - the members of the band are fed in an order based on their food deficit. The member with the largest deficit eats first, etc.

**Satisfied first** - the members of the band are fed in an order based on their satisfaction with the direction of travel. Those who voted for the selected direction eat first. Those who voted for adjacent directions eat next, etc.

**Round Robin** - the members of the band are fed in sequence, starting with a different member in rotation on each successive day.

Various success metrics were evaluated for each set. These include:

**Band longevity** - the average age of a band at the point where it died (all members are dead) or at the point in the simulation where the measurement was taken if the band is alive (has at least one living member). This is charted over the duration of the simulation.

**Bands that are dead** - the number of bands where all members of the band are currently dead. In the absence of reincarnation processing, this is equivalent to the number of bands that have died up to the point when the measurement was taken. This is charted over the duration of the simulation.

**Bands that were fed** - the number of bands in which all members of the band received the MDR on the day of the measurement. This is charted over the duration of the simulation.

**Guys that were fed** - the number of individual members of the band who received the MDR on the day of the measurement. This is charted over the duration of the simulation.

Five sets of 10 trials each were run covering 5 food distribution modes with each mode coded as follows:

- Mode 0 - Band members are fed in a fixed order on each turn.
- Mode 1 - Band members receive equal shares of the food on each turn.
- Mode 2 - The band members with the largest food deficit are fed first on each turn.
- Mode 3 - The band members who wanted to travel in the chosen direction are fed first on each turn.
- Mode 4 - The band members are fed in a fixed order, starting with a different member in rotation on each turn.

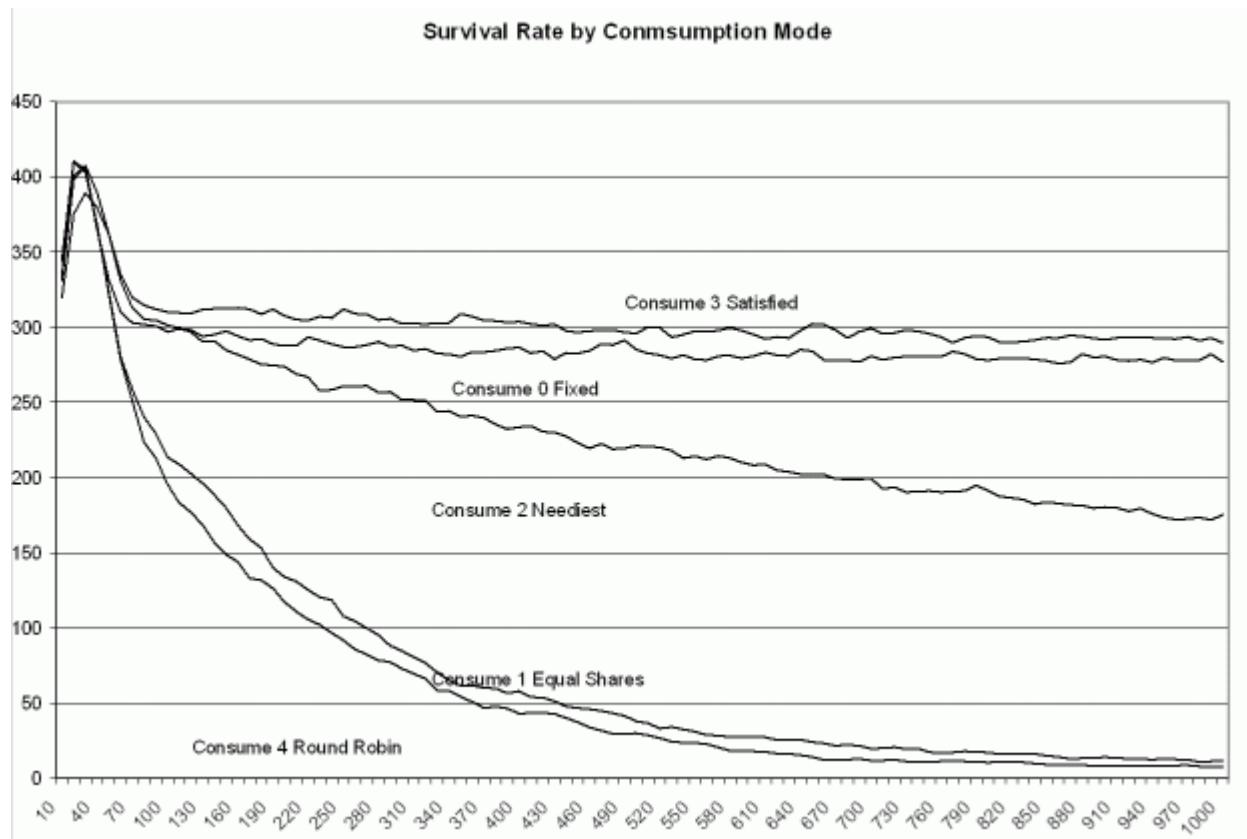
All simulations were run using a memory based search with a patchwork food distribution. Figure 6 shows the rate of survival for each of the five distribution modes. The vertical axis is the number of surviving band members and the horizontal axis is the day number.

One of the key things to note is that the performance of the vector voting model is markedly affected by how the resources are shared among the individuals in a band. This is due to the fact that there is no sharing of information among band members concerning the memories that brought about the decision of each. Thus, if individuals who are making bad choices are rewarded as much as those who are making good ones, it is likely that a member making a good choice will die and their memories will be lost to the group forever. Thus, both the round robin and equal sharing paradigms cause the number of individuals to be supported by vector voting models to converge to 0. This is equivalent to the random walk model. This is because the loss of key individuals is causing the decision making to behave in a random manner.

On the other hand, by rewarding the individuals who contributed to the winning decision we are rewarding the providers of key information and working to preserve the groups information that is stored at the individual level. Thus, satisfied first performs the best of the five, producing overall population levels approximately equal to the carrying capacity of the environment. In this case, vector voting exploits the local patterning of resources and satisfied first supports the retention of this information at the local level. "Survival of the fittest" in the case where there is no shared information between individuals means "survival of the group".

Fixed or rank order access to resources performs slightly less than the satisfied first paradigm, but not much less. It is also easier to enforce since one doesn't have to poll the members and decide who did what, where and when. Thus, fixed rank is a good fit with the vector voting model since it produces performance comparable to the satisfied first approach without the social overhead.

The neediest first paradigm is midway between the two groups. Once the population size is significantly below carrying capacity, the vector voting model allows sufficient exploitation of the environment to allow most individuals to be fed. Thus, the randomness that is introduced into the decision-making process is balanced out by the individuals that are helped through the process. Notice that if the group is allowed to share their detailed memories via language, the relative performance of this more altruistic strategy will improve.



**Figure 6.** The average number of individuals in the population over 10 runs for each of the 5 food distribution modes

## 5 Conclusions

In this paper we augmented the vector voting model of consensus-based decision-making in primate groups with emulative learning. Emulative learning was the basic form for cultural transmission used here. Extensive runs of this model suggest that the consensus based approach based upon icon memories was sufficient to produce regional populations that approximate the carrying capacity of a randomized, resource scarce environment when certain resources sharing strategies as used to distribute resources among a band's members. Thus, while the vector voting model cannot theoretically make the best choice all of the time, coupled with feedback in the form of emulative learning it is able to effectively optimize the number of individuals living in a relatively harsh class of environments, semi-arid environments here. The key is that although the environment exhibits local dynamics, it does not display local dynamics at this point. Therefore, a decision-making approach, such as the vector voting paradigm, that efficiently exploits local dynamics is sufficient to produce an optimum population size here relative to carrying capacity.

The vector voting model, while exploiting local patterns effectively, is very sensitive to the resource distribution system used. This is because individual knowledge is not shared. This means that if resources produced by the vector voting approach are shared without any consideration of the individuals who made the decisions in the first place, the individuals and their attendant information can be lost introducing randomness into the process and ultimately generating performance that can be the same as a random walk. Thus survival of the fittest is synonymous with survival of the group when information is not sharable.

Also, there are two resource sharing strategies, satisfied first and rank order, that tend to produce the results that are closest to the carrying capacity. While the satisfied first approach is the most precise way to reward the individual responsible for a groups success, and in turn preserving the information that they used, it requires a certain amount of overhead to manage and enforce. A fixed or rank ordering of access to resources produces results that are slightly, but not much lower, and with less overhead in terms of how to determine allocation.

The next step in our project is to implement the Cultural Algorithm paradigm which allows the individuals to pool their memories. Theoretically, this sharing of information will allow the group to make regional decisions about certain predicates more reliably than in the vector voting model. In particular, the pooling of information via a belief space will allow the group to decide on predicates that are not tied to particular locations, e.g. norms of behavior, within the environment. These are called position invariant predicates. Being able to do this makes imitative learning, a form of learning largely unique to the human species, possible [5]. Both activities are influenced by the structure of the language used to perform the pooling and within which the learned information can be articulated. This will be the subject of a subsequent paper. Also, it will be of interest to see how the introduction of shared information impacts the relative utility of the resource distribution modes. That is, will equal access to knowledge within the group imply that equal access to resources will provide better support for group decision making than it currently does in a vector voting environment.

## Acknowledgements

Dr. Reynolds research is supported by NSF Intelligent Information Systems grant IIS-9907257.

## References

- [1] Reynolds, R. G., (1978) On Modeling the Evolution of Hunter-Gatherer decision-Making Systems, *Geographical Analysis*, Vol. X:31-46.
- [2] Tomasello, M., (1999) The Human Adaptation for Culture, *Annual Review of Anthropology*, Vol. 28:509-529.
- [3] Kummer, Hans, *Primate Societies*, New York, Aldine Press, 1968.
- [4] Reynolds, R. G., and Zeigler, B. P. (1979) A Formal Model for the Operation of Consensus-Based Hunting-Gathering Bands, in C. Rebfrew, K. Cooke (eds) *Transformations: Mathematical Approaches to Culture Change*, New York, Academic Press, pp:405-418
- [5] Pedone, R., and Conte, R. (2000) The Simmel Effect: Imitation and Avoidance in Social Hierarchies, paper to be presented at Workshop on Cultural Transmission, Sienna, Italy.

# **LANGUAGE CHANGE AS EVOLUTION: LOOKING FOR LINGUISTIC “GENES”**

## **Vienna English Working Paper**

By  
**Nikolaus Ritt**

I would be extremely grateful for comments and criticism. Please send me an eMail.

### **0.0. Prolegomena**

This paper argues for the view that language change and acquisition represent evolutionary processes in a sense that we nowadays tend to associate with Darwinism. This view is not exactly new, but has never really established itself within the community of historical linguists. It goes back, at least, to August Schleicher, to whom we owe the notion of linguistic ‘family trees’ and who viewed languages as similar to organisms with life-cycles in the sense that they are first ‘born’, then grow into fully-developed, or maximally complex ‘adults’, and may finally degenerate, lose their complexity and possibly ‘die’. His conception of ‘evolution’ was essentially pre-Darwinian, though, and his application of the concepts to linguistic study highly metaphorical. Since Saussure and the formalist linguistic tradition he more or less established were opposed to the borrowing of metaphors from other sciences on grounds of principle, it is no big surprise that approaches like Schleicher’s ceased to be taken seriously. Evolutionary concepts only sneaked back in disguise, so to speak, into linguistics via the functionalist approaches developed by Jakobson and the Prague School. There, language change was regarded as functional, ‘goal directed’ and, one could say, ‘adaptive’ in a sense similar to the one the term has in biology. However, the functionalists didn’t explicitly base their explanations on a theoretical framework that could be called truly evolutionary in the Darwinian sense - probably because they were too cautious themselves to borrow metaphors which, they must have felt, belonged essentially to a different science and had no place in theirs. Retrospectively, I feel that by being so cautious, they made things unnecessarily difficult for themselves, as it may have been due to the very lack of an explicitly evolutionary framework, that their approaches to language change were vulnerable to attacks from structuralists (such as Lass, see for example 1980: 64ff.) for being teleological and unable to offer truly causal explanations. - Whatever the exact reasons for the development may have been, however, it is apt to say that nowadays, as April Macmahon puts it, „Evolution [...] has become a ‘dirty word’ in modern linguistic theory” (1994: 314).

It is little surprising, therefore, that more recent attempts at dealing with linguistic phenomena from an explicitly evolutionary point-of-view have originated at the margins, or even outside the linguistic community itself. Here is a survey of books and articles that I have found particularly inspiring.

First, I would like to mention Cavalli-Sforza and Feldman’s, and Lumsden and Wilson’s volumes on cultural evolution. Both include linguistic evolution without giving the latter much space in their arguments or presenting any specifically linguistic cases, though. Similarly, Richard Dawkins’ (1982 and 1989) proposal that cultural evolution might be based on the existence of ‘mental replicators’, which he calls ‘memes’, seems to be straightforwardly applicable to

language evolution, while Dawkins himself does not develop his arguments in that direction. Interesting comments on Dawkins' proposal as well as on the frameworks developed by the other authors just mentioned can be found in Maynard-Smith (1989).

More recently, a group of scientists based at the Sta. Fe Institute have been developing an elaborate theory of 'Complex Adaptive Systems', which presents itself as a metatheory for the study of a large range of phenomena that display 'adaptive' behaviour and which are difficult to deal with on the basis of classical scientific approaches. These include national economies, immune systems, artificial intelligence, cognitive development and also, quite explicitly, language evolution. Due to the bad reputation which 'evolution' seems to have within the linguistic community, however, the workshop on the evolution of human languages, which the Institute organised in 1991, has not lead to the establishment of a permanent research group.

While the failure of the Sta. Fe initiative shows how alien evolutionary concepts might still be to the community of mainstream linguists, contemporary cognitive psychology seems to be adopting evolutionism more enthusiastically. This becomes particularly obvious in Henry Plotkin's recent volume '*Darwin Machines and the Nature of Knowledge*', for example. Since the links between cognitive psychology and linguistics are getting continually stronger, however, it seems that before long the linguistic community will have to give up its established biases as well. In other words, I feel that there are quite good chances that evolutionary theory might, after a long period of neglect, find itself to become the majority paradigm in (historical) linguistics within the not-so distant future. Evolutionism is in the air, so to speak, and will soon manifest itself strongly. And I don't seem to be the only one who thinks so. Thus, I have recently kept running into or hearing of colleagues who openly admit to working on possible applications of evolutionary theory to linguistics, although publications are still missing. Among these colleagues are April Macmahon from Cambridge, Guy Cook from London, or Steve McGill from Exeter. Also, Robert de Beaugrande's forthcoming volume on Discourse Analysis adopts many notions from the theories developed at the Sta. Fe Institute.

Finally, I would like to refer to two papers by myself, namely forthc. (a) and (b), in which I have thrown much caution overboard and been as explicitly evolutionary as I dared.

But back to this paper, then. It introduces some of my personal ideas about evolutionary theory and its relation to language change, acquisition and use. It does so in an informal, associative manner without giving much heed to potential advantages of the advocated approach over other, more conventional ones. It is intended to stimulate the search for such advantages, which I feel can easily be found. The main reason why I am not presenting them in any detail here is that they deserve, in my mind, an elaborate discussion that would necessarily go beyond the scope of this contribution. So, if this paper succeeds in arousing some interest and the desire to play around with some of the concepts and perspectives that it develops, I shall be happy enough. Otherwise, I hope that the reader will at least find the mental gymnastics my attempts at reasoning will force her to perform refreshing and worth the time spent on reading all this.

## 0. Introduction

*Since 'tis Nature's law to change.*

*Constancy alone is strange.*

*John Wilmot. Earl of Rochester*

--- A dialogue between Strephon and Daphne

This motto opens the first chapter of Jean Aitchison's classic volume on language change (1991). A weird choice, isn't it? After all, the book is about language *change*, so why are we told that it's *constancy* that we should really worry about? OK, I'm sorry. Of course, this is just a cheap move of mine, and, of course, Jean Aitchison was just trying to tell us that language *change* was less strange and unnatural than one might be tempted to think as a novice to the subject. - But still, if one takes the motto seriously, one cannot help but wonder. Why is it, quite generally, that historical linguists have always focused their attention on instances of linguistic change, and have not found it worth the trouble to deal with those elements in the world's languages that have maintained their identities and shapes over longer periods of time? Why is it, to mention a concrete case, that people in our field have mostly been trying to explain why Middle English long /i:/ shows up as /aɪ/ in Modern English, while the fact that Middle English short /ɪ/ is still short /ɪ/ in many Modern English words has been taken more or less for granted? At least in the light of the Earl of Rochester's insightful remark, this is almost as strange as constancy itself. In this paper, I'll do what strikes me as the obvious, therefore, and approach language history via constancy rather than change.

So, why is it then that Middle English short /ɪ/ is still short /ɪ/ in many Modern English words? Good question. Or is it? In what sense can we say at all that the /ɪ/ in items such as ModE *middle, it, children* is the same as the - assumed - /ɪ/ in ME *middle, it, children*? Although it is typically assumed that the two are in some sense identical, even a little bit of reflection tells one that things are not so self-evident at all.

On the one hand, we could say of course that the two /ɪ/s are, in some sense, counterparts, because they fulfil similar communicative functions: *middle, it* and *children* mean pretty much the same in Modern and in Middle English, so there is good reason to assume that the purposes for and the situations in which Modern English speakers will use the words are at least roughly comparable to those which Middle English speakers had in mind when they employed them. If, by the same rationale, we then also say that both Middle and Modern English speakers used /ɪ/ to distinguish words such as *middle, it* or *children* from others, it follows that the role which Modern English /ɪ/ plays within Modern English speech communities is indeed similar to the one which Middle English /ɪ/ played within Middle English speech communities.

In order for the question why Middle English /ɪ/ is still /ɪ/ in Modern English to be meaningful, though, the mere observation that the two /ɪ/s are communicative counterparts is not enough. What makes the relationship between Middle English and Modern English /ɪ/ special is that the two do not only play similar roles but that, in some sense, there also exists a kind of genetic relationship between them. In some way, Modern English /ɪ/ seems to be a distant offspring of Middle English /ɪ/, and we feel that Modern English /ɪ/ has not only inherited many of the jobs

of Middle English /I/ but is its direct descendant at the same time. It is both its communicative, or 'functional', and its genetic counterpart. When we want to know why 'Middle English /I/ is still /I/ in Modern English, we are therefore dealing with two questions rather than one. The first would be: how has Middle English /I/ managed to produce offspring that have survived over the centuries? And the second question would be: why and how have the offspring of /I/ managed to take over the communicative functions of their forebears? In the following I shall deal with each of these questions in turn.

Before I go on, however, I guess I owe you a little break to reflect on the course which my argument (or rather my loose associations) seems to have taken. Thus, you will have observed that at the beginning my paper seemed to be about a speech sound, a phoneme, something well defined, you may have remembered, as an element of 'langue' in the Saussurean sense, and now I have come to talk about /I/'s offspring and about /I/'s jobs - just as if I were talking about a person, a human being, or at least a living thing of some sort, capable of reproducing and of doing things. - Obviously, you will think, my discourse has become metaphorical, and whatever the rhetoric or didactic advantages of metaphors might be, you will be aware that one must not let oneself be carried away by them and that, in particular, one must not tacitly endow them with a technical sense, because that way one might wind up in a completely fictional world and solve merely fictional problems that have no bearing on the world out there and thus no truly scientific value.

1. Phonemes (and other constituents of natural languages, for that matter) as active replicators  
So, what about /I/ and its offspring then? Is there a technical sense in which such a statement can be read? In what way, if at all, can phonemes be assumed to increase and multiply? This question may sound weird to anybody whose mind has been framed to think in the categories established within the linguistic community - I agree -, but might it not still pay to take it seriously if only to see what happens? How can phonemes be assumed to 'propagate' then? I guess the first type of answer that will probably spring to the mind of most of us is: 'through language acquisition.'

Children learn phonemes through listening to (more or less) grown-up speakers communicating. The general idea is that some part of the human brain works as a device for language acquisition, into which some of the more basic principles concerning the way human languages work are hard wired, so to speak, and which, when exposed to actual utterances will filter out and store in its more flexible components those pieces of information that a speaker needs to produce such linguistic utterances as are likely to serve her communicative needs in the community she grows up in. Those pieces of information will either be more like elements or more like rules/processes, but the distinction doesn't really matter here, because we are interested in /I/, and /I/ is most probably an element rather than a rule. In any case, when we think of /I/ as being learned by humans and stored within some part of their brains, the picture seems to emerge of /I/ as being transmitted passively, while the active agents in its replication seem to be the speakers. It would seem, therefore, that when trying to answer the question why Middle English /I/ is still /I/ in Modern English, we ought to focus our attention on speakers and rephrase our question as something like 'Why have English speakers over the generations successfully acquired /I/?'

But is this necessarily so? In what way are speakers really more active than, say, phonemes in the replication of the latter? Obviously, speakers do not control language acquisition consciously

and actively. This has been a home truth ever since Saussure. They are not normally in a position to decide whether they like to acquire a particular phoneme, for example, or not. In this sense 'langue' is beyond the control of individual speakers. So, if we wanted to investigate the speakers' role in language acquisition/replication we wouldn't be talking about speakers as autonomous subjects freely determining their own actions, but we would be looking inside them and disregard their personal integrities, so to speak. What we'd be interested in would be the ways their auditive apparatuses and their articulatory organs work, and also, of course, the way in which these interact with their brains and the way in which linguistic elements and processes are mentally stored. (It is assumed here that language ultimately does have - even though it may not be reducible to - physical reality.) The speakers we would be looking at then, would look like a system of muscles, membranes, teeth, assemblies of nerve cells and other such elements - not much like individuals at all, really.

And where would our phoneme /I/ reside in this mass? Well, although there are as yet no ways of verifying this, there is in fact only one reasonably plausible possibility. It must be located within the central nervous system, i.e. the brain. Without knowing much about the way the brain handles information, most neurologists would subscribe to the notion that an element such as /I/ might be located within an assembly of nerve cells that are linked - however remotely or indirectly - to both articulatory and auditory organs. That assembly, which 'represents', or, actually, 'is' /I/ will be excited when it receives input in the form of sounds that are 'recognised' as /I/ or when other parts of the nervous system get excited in such a way that a realisation of /I/ is pronounced. Another way of putting this would be to say that, depending on the actual state of other relevant parts of the brain, the excitement, or 'firing', of a nerve cell assembly '/I/' will trigger either the firing of such other nerve cell assemblies as eventually amount to appropriate movement of the articulators, and/or the firing of such assemblies which 'encode' or 'are' word forms, morphemes, concepts or socially relevant information. A brain can thus be said to host an /I/ assembly, if there exists a set of nerve cells within it that get indeed excited more or less simultaneously when they receive electrochemical input under such conditions as specified above. The existence of /I/ is thus established when the channels through which electrochemical energy flows during a brain's activity come to be set up in such a way that an assembly of them fires in quasi-unison. The acquisition of /I/ can consequently be thought of as a process that results in the establishment of appropriate links among a set of relevant nerve cells.

During its lifetime, then, an /I/ assembly may be in either of two states: when it fires, it is 'on', when it doesn't, it's 'off'. Due to the fact that the /I/ assembly is linked to articulators it may then happen that the firing of /I/ causes the latter to perform a gesture 'expressing' /I/, typically the allophone [I]. During a lifetime, an /I/ assembly can thus be assumed to give rise to a relatively large number of [I]s, and the [I]s that can be observed in actual utterances can thus be thought of as consequences, or expressions of /I/ in a similar way as the phenotypic characteristics and some behaviour patterns of organisms can be regarded as consequences or 'expressions' of genes. It is exposition to such [I]s in appropriate contexts, then, that allows children to acquire /I/s, or that - to stay within the descriptive framework I have begun to sketch - causes such links between certain nerve cells to establish themselves within children's central nervous systems that may be looked at - from the linguist's point-of-view - as (a representation of) /I/s. Through their inherent ability to produce [I]s under appropriate conditions, /I/s can thus place new copies of themselves within other nervous systems. - Looked at from this perspective, then, an /I/ - or indeed any

phoneme - can be thought of as an entity capable of its own reproduction - a true active replicator and perfect mental counterpart of a gene, which expresses itself through creating organisms capable of spreading copies of the gene through reproduction. In the same way as the story of gene reproduction can be told without invoking organisms as central agents (they can be referred to as the 'vehicles' or 'interactors' of genes), the story of the life cycle of a phoneme such as /I/ can be told without referring to 'speakers' as the primary agents in that process. For the life and reproduction of /I/s 'speakers' constitute only the necessary environment and the necessary tools, they play no active role in it - and they cannot, normally, influence it.

The first part of our question can now be partly answered. Middle English /I/ has managed to pass its offspring down to Modern times, because it was turned on often enough in the right way to produce [I]s, which in turn placed new copies of /I/ in the brains of new generations of speakers.

Obviously, this is only a rough approximation to an answer, and raises a lot of questions in itself. In particular, it is obvious that a variety of further conditions must be met if the production of [I]s is really to engender new /I/s. Exposure to [I]s will only lead to the acquisition of /I/ if the [I]s form part of communicatively effective linguistic messages. I will turn to the question how those additional conditions will look within the framework I'm just sketching presently. I would like to do this; however, by taking up the second part of our initial question, namely why it is that Modern English /I/ still does many of the jobs that ME /I/ did? As it seems to me, this question is more closely related to the problem of the conditions under which /I/ will get acquired than might first be suspected.

What are those jobs I am talking about anyway? Let me give a few examples. Well, first of all, there is the primary function of /I/ as a phoneme, which is to distinguish the morphemes in which it occurs from one another. Secondly, there are the functions of making pronunciation and perception possible, which /I/ shares with all other elements and processes that 'inhabit' human phonologies, of course. Finally, then, there are other potential functions which a phoneme such as /I/ may have, including morphological ones such as indicating a particular morphological environment, or social ones such as identifying the social adherence of a speaker, and probably many others as well. Now, how do these functions translate into the framework that I have been sketching so far? If phonemes are regarded as nerve cell assemblies, then - interestingly, but actually quite obviously - so can the so called functions of a phoneme. If recognition of /I/ is viewed as the firing of an assembly, it is equally plausible to assume that if /I/ and /t/ fire one after the other, this event will in turn excite an assembly {/It/} which will automatically excite assemblies for {3rd person} {neuter} and {singular}. Similarly, both articulation and perception crucially involve the excitement of nerve cells along certain 'pathways'. Thus, the perceptibility as well as the pronouncability of /I/ result from /I/'s association with the respective neuronal pathways. In other words and quite generally speaking, then, /I/ fulfils its functions by virtue of its associations with other cell assemblies. In some cases, as in morpheme recognition or in articulation, the firing of /I/ will cause a firing of associated assemblies, while in others the firing of appropriate 'function' assemblies will cause a firing of /I/. The fact that /I/ does certain jobs can thus be viewed, quite simply, as its occupying a certain designated place within a network of associated cell assemblies and standing in mutual triggering relationships with those. Viewed this way, however, the jobs which /I/ does are at the same time the clue to its very existence as an

assembly, since only through being triggered to fire in unison a set of nerve cells emerges as an identifiable assembly in its own right. A number of connected cells that never come to fire in unison are, by definition, no assembly in the sense that we have established above.

It seems to follow, then, that whether or not an /I/ will emerge as a stably connected nerve cell assembly within a particular brain depends on whether it gets excited sufficiently often by the assemblies in its environment - and this, incidentally, represents an answer to the question concerning the conditions under which the production of [I]s can be expected to create new copies of /I/. As we observed, it cannot be enough for some sets of cells to be excited through input from the sensory pathway, because in order for /I/ to be acquired, it is not enough to hear [I] sufficiently often. Rather, [I] must be received and 'interpreted' as part of a meaningful message. In terms of the model I am sketching here, then, this can only be done, if there exist the relevant other 'linguistic' cell assemblies in the mental environment of /I/

Of course, if the acquisition of any linguistic element depends on the presence of other elements, the notorious chicken-egg question seems to raise itself. It can be solved relatively easily, though, on the assumption of repeated boot-strapping as well as of auto-catalytic self-organisation processes. One only needs to assume a very tiny set of linguistic 'universals' to be hard-wired (or 'inherited'), so that these then provide the environment in which further linguistic elements can establish themselves: an assumption that is perfectly compatible, if not equivalent, to the well established notion of a genetically provided 'language acquisition device'. Whatever the detailed mechanisms behind the emergence of fully functional language systems, though, I feel that the picture that emerges from what has been said so far offers a rather cute re-interpretation of Saussure's view of language as a system 'ou tout se tient'. At the same time, the dependence of /I/ on the presence of certain other assemblies represents another beautiful analogy to DNA based evolution, where the evolutionary stability of genes also depends on the presence of other genes within the genome. In the same way - to give just one example - as an /I/ assembly makes sense only within a more or less complete system of other 'language assemblies', there can also not exist a gene for eye colour without the system of genes that manufacture the rest of the eye and the organism around the eye.

Another important aspect of the acquisition of such linguistic elements as the /I/ that this paper focuses on, is that they can be assumed to establish themselves within a new brain only, if their establishment and the activity it causes in associated assemblies gets 'rewarded' and thus 'reinforced'. In other words, the effects of an emerging assembly on its environment must be such that they feed back on the assembly by making its future firing more likely, so that the assembly acquires the necessary stability. 'Reward' and 'reinforcement' must ultimately come from neuronal activities whose effect is that they make people feel good, or, in other words, cause psychosomatic states which represent positive emotions and which are probably hardwired (i.e. genetically determined) into the nervous systems of human beings.) This is just a technical way of saying that the acquisition of appropriate linguistic cell assemblies is probably rewarded through the positive emotions that go along with successful communicative acts, or - in yet simpler terms - that a brain will acquire such elements as help its bearer to get himself across to and understand other members of the speech community. The acquisition of an /I/ phoneme can thus be viewed as just one series among the adaptive process that a brain goes through as it evolves towards a state in which it enables its owner to interact/communicate with her

environment in a sufficiently harmonious manner. At first, an assembly which eventually turns out as an /I/ will be just one of a larger set of many possible neuronal configurations a brain may assume. As its activity, that is its firing, turns out to have rewarding effects more often than the firing of other neuronal configurations, it will probably 'attract' more electrochemical energy and get fired more often than the latter. By virtue of that fact, then, it will establish itself as a stable assembly, while its competitors won't. The emergence of an /I/ assembly within an individual brain can thus be viewed as a process that involves chance variation among brain states and selection of some states over others. The relative fitness of the competing states is determined by the degree to which they affect positive 'emotional' responses, while the general 'tastes' of emotions have probably been shaped through selectional processes acting on the genetic level.

For /I/'s success as an active replicator, this means, that its capacity of producing [I]s is just one aspect determining the success of its replication. Producing [I]s will probably increase the number of times an assembly corresponding to /I/ gets excited within a new acquisitor's brain but it cannot guarantee that new /I/'s stability. However, the very existence of 'adult' /I/s in the speech community also increases the chances for the effects of a potential new /I/ to be 'rewarded' and thus greatly increases the chances of a new /I/'s stability. /I/ is thus an active replicator in two ways: it generates new /I/s through emitting [I]s, and it breeds newly generated /I/s by making its host (or: speaker) react to their effects in a way that gives the hosts of the new /I/s the feeling of being 'understood' or having reached their communicative goals. So much for my view as to why Middle English /I/ is still /I/ in Modern English. /I/s are still around in abundance in Modern English in environments similar to those in which Middle English /I/s thrived, because the latter have managed to replicate themselves successfully through multiple generations of host brains.

## **2. Some reasons why I find it so attractive to view linguistic elements as replicators**

What are the advantages, if any, of viewing phonemes as replicators? After all, it would have been equally possible to say that /I/ and its functions have been successfully acquired by successive generations of English speakers and that's why it is still around? And this sounds more familiar as well, doesn't it? It's much more easily digestible, doesn't demand mental gymnastics and seems to capture, basically, the same facts. So why take the trouble?

Well, I can think of a number of good reasons. Here I will present just two.

First, a statement to the effect that speakers have successfully acquired /I/s over successive generations forces one to change perspectives, because phonemes and speakers belong to different ontological domains. Such a change is intellectually unsatisfactory, particularly in lack of a theory that makes the links between the two domains or levels explicit. Thus, saying that a speaker acquires a phoneme begs the question of how she actually does that, and in structural historical accounts this question is typically never taken up seriously or even answered, because speakers are treated as black boxes at best or even regarded as outside the domain of linguistic science proper. - The view sketched here, on the other hand, is perfectly explicit about the way in which language acquisition is supposed to work, while at the same time describing it exclusively from the point of view of the linguistic elements themselves. No recourse to notions such as that

of 'speakers' as agents in the story of linguistic evolution has to be taken. Speakers figure exclusively as the environments in which the replication of linguistic elements take place, and the story can remain a consistently 'linguistic' one.

Second, viewing linguistic elements as replicators opens the possibility of transferring to the study of linguistics many theoretical insights gained in other sciences that have for a longer period and more explicitly concerned themselves with the study of replicating systems and their evolution. First and foremost these include evolutionary biology, of course, but also other sciences applying generalised Darwinian frameworks, such as evolutionary cognitive psychology, artificial intelligence and particularly the study of systems that are capable of 'learning' on the basis of genetic algorithms, and so on.

In particular, the view that linguistic elements are replicators affords interesting aspects on questions relating to language change. Remember that I argued that the success of /I/s replication depends, among other things, on whether a new brain can integrate /I/ meaningfully within the network of linguistic elements and processes that are emerging within it. One aspect of an emerging linguistic system or competence, is therefore, that it provides slots that need to be filled, and it is such a slot that an element /I/ needs to fill if it is to replicate successfully. It seems to be a characteristic of such slots that there is not only one way of filling them. As the evidence of linguistic variation and change show, it is possible and happens quite frequently that within one and the same speech community corresponding slots happen to be occupied by different elements within different brains. Thus, as a very trivial example, take the variation between /U/ and /A/ in English, where either of the two may occupy the slot provided by the context of such words as *hut*, *but*, *butter*, and so on. It can be said, therefore, that within the pool of linguistic elements which occupy brains of speakers of English, there is a competition between /U/ and /A/ for a certain defined slot, or that - with regard to that slot /U/ and /A/ are *alleles*, to borrow a term from genetics (where it refers to different genes that can occupy the same slot on a DNA strand and are thus competitors). Typically, such a competition for slots - on which any linguistic element depends for its existence - will take place in all instances of linguistic variation, so that variation itself can in fact be regarded as competition. Whenever variation leads to actual change, the cause for such an event must be that for some reason a new competitor for a given slot manages to place more copies of it there than its established rival. In the case of phonological elements, ease of perception and production will obviously play a great role there, in the case of other elements semiotic parameters determining the ease with which links between associated elements are established will probably play similar roles. At the same time, however, the fact that existing elements will tend to reward the establishment of what they 'recognise' as new copies of themselves will be a powerful barrier against the spread of new variants, and if a new variant is to establish itself successfully then it must have sufficient selective advantages over the established competitor. Such selective pressures which favour the spread of a new variant at the expense of an established one may among other things be due to changes in the environment of the slot, where environment includes both a slot's immediate systemic environment and the wider environment comprising the whole set-up of the speech community in which linguistic elements 'live'.

Also, the view of linguistic evolution as a story of items attempting, with varying success, to replicate themselves affords interesting perspectives on the notorious problems concerning the

actuation and the implementation of language change. Change is actuated under this view simply through undirected copying mistakes that occur in the replication process. It may have various reasons, some of them pretty straightforward, such as misarticulations and misunderstandings. More crucially, though, variation is already built into the copying process itself, since an item relies on an acquisitor's brain to produce 'random' varieties of neuronal configurations among which the replicating item will reinforce those most similar to itself and thus make them stable. In other words, the processes by which linguistic elements get copied involves successive stages, during the first of which the average copying fidelity must necessarily be relatively low. The resulting variability of languages is thus a natural matter. - As far as the implementation of changes is concerned, then, it can be regarded as parallel to the implementation of mutational changes within the gene-pool of living species: it results from the relatively more successful replication of a new variant within the brains of a linguistic community.

Finally, this perspective makes one of the more puzzling aspects of language evolution easy to digest, namely the fact that - on the one hand - language changes often seem to be adaptive in the sense that they seem to make things easier for language users, while - on the other hand - the world's languages are full of elements and qualities that must count as clearly suboptimal with regard to those criteria that allow language change to be regarded as 'optimisation' in the first place. As the approach I have sketched here suggests, the items of which languages are made up do not primarily exist 'because of' the purposes they serve their speakers, but rather simply because they have managed to replicate sufficiently well. Language evolution takes place, in other words, primarily because the elements that constitute languages are replicators which strive to propagate. The needs of speakers only represent constraints on the propagation of linguistic replicators, determining which of them will survive more easily in an environment where resources (in this case electro-chemical energy supplied within human brains) are as limited as they are anywhere within our world.

## References

- Aitchison, Jean, 1991: *Language change : progress or decay?*. 2nd. edition. Cambridge: University Press.
- Beaugrande, Robert de, forthc.: *New foundations for a science of text and discourse*. Norwood; NJ: Ablex.
- Bichakjian, Bernard H., 1988: *Evolution in language*. Ann Arbor: Karoma.
- Cavalli-Sforza, Luigi L. and M. W. Feldman, 1981: *Cultural Transmission & Evolution: A Quantitative Approach*. Princeton: University Press.
- Dawkins, Richard, 1982: *The extended phenotype*. Oxford: University Press.
- Dawkins, Richard, 1989: *The selfish gene*. 2nd edition. Oxford: University Press.
- Hawkins, John A. and Murray Gell-Mann (eds.), 1992: *The evolution of human languages*. Redwood, CA: Addison-Wesley

- Hebb, Donald O, 1949: *The organisation of behavior*. New York: John Wiley.
- Hull, D.L., 1988: Interactors vs. vehicles. In: Plotkin, Henry C., ed.: *The role of behaviour in evolution*. Cambridge, MASS: MIT Press.
- Lass, Roger, 1980: *On explaining language change*. Cambridge: University Press.
- Lumsden, Charles J. and Edward O. Wilson, 1981: *Genes, mind, and culture : the coevolutionary process*. Cambridge, MASS: Harvard University. Press.
- Macmahon, April, 1994: *Understanding language change*. Cambridge: University Press.
- Maynard-Smith, John, 1989: *Did Darwin get it right? : essays on games, sex and evolution*. New York: Chapman and Hall.
- Plotkin, Henry C., 1994: *Darwin machines and the nature of knowledge*. Cambridge, MASS: Harvard University Press.
- Ritt, Nikolaus forthc. (a): The spread of Scandinavian 3rd psn.pl. pronouns in English: optimization, adaptation and evolutionary stability. In: Kastovsky, Dieter, ed. *Papers of the 1993 conference on Languages and Contact in the History of English, Tulln, Austria*.
- Ritt, Nikolaus forthc. (b): Mutation, variation and selection in phonological evolution: a sketch based on the case of Late Middle English a > au/ \_l{C/#}. In: Fisiak, Jacek, ed. *Papers of the 1994 conference on Middle English, Rydzyna, Poland*.

# CONTROVERSIES IN MEME THEORY

By  
**Nick Rose**

Department of Psychology  
University of the West of England, Bristol. UK.  
And the 'Meme Lab'.  
Nicholas.Rose@uwe.ac.uk

## Abstract

Meme theory and the notion of cultural evolution present the possibility of a fundamentally new way of understanding human culture. Yet some of the speculation within meme theory has become confusing and ambiguous. Four areas of meme theory are critically reviewed. These are ambiguity in the definition of a meme and confusion regarding the distinction between replicator and phenotype, the problem of inheritance of acquired characteristics, the relationship between memetics and sociobiology, and the selection or mutation of memes being carried out by conscious foresight. Whilst I suggest directions which might resolve these problems, the purpose of this short review is to help generate the wider debate required to settle these important issues.

**Keywords:** meme theory, replicator, phenotype, Lamarckism, sociobiology, self-centred selectionism, Cloak, Dawkins, Dennett

## 1 Introduction

Evolution by natural selection has been controversial since Darwin first published 'The Origin of Species' in 1859, not least amongst those who hold religious beliefs. By demonstrating that through the accumulation of small inherited changes and differential survival, 'nature' could select animals upon the basis of their suitability to their environment, Darwin achieved an explanation of how complexity and apparent design could have arisen in nature without the need for a designer. Perhaps the most controversial aspect of Darwin's theory was that it challenged the view that 'Man' was somehow separate from the rest of nature. Human beings were merely another branch of life, stemming from the same evolutionary process and primitive beginnings as all other animals.

Despite the controversy, Darwinian evolutionary theory has succeeded in becoming a central feature of the biological sciences. The theory has undergone some refinement during the time it has been around; not least with the unification in the 1930's with the genetic mechanism of heredity (e.g. Fisher [22]). Essentially, however, it could be said that Darwin's theory is the centre piece of biology, the basis for explaining the origins and history of life on Earth, and perhaps everywhere else in the universe (c.f. Dawkins [17]; Dennett [19]).

The core tenets of evolution by natural selection 'are as much open to doubt as the theory that the earth goes round the sun' (Dawkins [15], p1). Advances in sociobiology and evolutionary

psychology attempt to use the strength of this theory to generate explanations of human behaviour by seeking possible genetically based evolutionary adaptations (e.g. Barkow, Cosmides, and Tooby [2]). However, it is clear that human beings differ from other animals on Earth by the complexity of their culture. By apparently seeking to explain all human behaviour in terms of biology, sociobiology has found itself accused of what Dennett [20] calls 'greedy reductionism'. It is clear that much of human culture requires a more appropriate level of description. Although the level of the gene may not fulfill that requirement, it remains to be determined whether evolutionary theory could be more widely applied to incorporate human cultural change. It has been suggested that cultural complexity may have arisen through mechanisms analogous to those responsible for the formation of organic life, leading to the attempt to create a theory for the evolution of culture.

Memetics takes the Darwinian revolution one step further. Evolutionary theory explains how complexity and apparent design could emerge within nature without the need of a designer. In the case of culture the 'designer' could be the 'common sense' notions of 'free will', 'intentionality', or 'Self'. In a way memetics appears to attempt to replace the 'Self' (c.f. Blackmore [5]; Dennett [19]) 1 with an algorithm of culture in the same way that Darwin and his successors have replaced 'God' with an algorithm (c.f. Dennett [20], p 50) 2 of life. We can expect the future of memetics to be a difficult and controversial one.

Universal Darwinism (Plotkin [32]; Dawkins [15]) and its application to culture provides a rich and exciting source of speculation from a wide variety of disciplines. Meme theory, in particular, has proved fertile ground for such theorising. Whilst much of this theory has been good, some of it has become confusing and ambiguous.

This paper will offer a critical review of four areas of meme theory. These involve difficulties for the theory which I suspect are also of concern to many people both within, and outside, memetics. I have views as to how each of these problem areas could be settled. However, the purpose of this paper is to give a short review of these areas in the hope of stimulating something of the wider debate required to find solutions for these problems.

## 2 Replicators and Phenotypes

The definition of a meme is currently ambiguous. A meme can be found variously described as; a unit of imitation (Dawkins [15]), a unit of information residing in a brain (Dawkins [16]), culturally transmitted instructions (Dennett [19, 20]), any permanent pattern of matter or information produced by an act of human intentionality (Csikszentmihalyi [14]), roughly equivalent to ideas or representations (Plotkin [32]), a unit of information in a mind whose existence influences events such that copies of itself get created in other minds (Brodie [9]), actively contagious ideas (Lynch [28]), a mental representation (Gabora [23]), a self-replicating element of culture passed on by imitation (Oxford English Dictionary), etc. Without some kind of firm definition the word 'meme' becomes almost meaningless (c.f. Wilkins [35]) applied to instructions in brains, information, behaviour, words, mental states, books and all kinds of cultural artefacts without consistency.

Dawkins [15] used genes as an analogy for memes and recognised that the analogy had natural limitations. It is difficult to recognise anything like a memetic allele (although Durham [21], refers to 'allomemes') or memetic chromosomes, for instance. Dawkins [15] suggested that the meme is still "drifting clumsily about in its primeval soup" (p 206) much like early replicating molecules, and might have a very different mechanism for competition such as memory space in the brain. Thus Dawkins [15] appears to resist any definition of a vehicle of culture or memetic phenotype.

Wilkins [35] attempts to clarify this analogy by tracing Dawkins' conception of memes from Williams' evolutionary gene and examining Hull's (e.g. Hull [26]) distinction between replicator and interactor. I'd like to take an alternative approach and trace Dawkins' conception of memes from Cloak's [12] work on cultural ethology. Cloak attempted to lay the foundation for applying ethological methods to the study of culture-specific human behaviours. Cloak's [12] definitions are much less confused, and thus perhaps more useful, than the definitions for memes. He calls the set of cultural *instructions* people carry in their nervous systems, 'i-culture'. The *material* structures, relationships among material structures, and changes in these relationships which are actually brought about or maintained by behaviours of those cultural instructions, he calls, 'm-culture'. Features of m-culture thus include features of their behaviour, their technology, their social organisation, and when considered as a set of verbal behaviours, ideologies. Cloak [12] summarised the i-culture-m-culture interaction in a way which makes the link to Dawkins' subsequent formulation of memes and their phenotypes quite apparent.

"An i-culture builds and operates m-culture features whose *ultimate function* is to provide for the maintenance and propagation of the i-culture in a certain environment. And the m-culture features, in turn, environmentally affect the composition of the i-culture so as to maintain or increase their own capabilities for performing that function. As a result, each m-culture feature is shaped for its *particular* functions in that environment." (Cloak [12], p 170, original italics).

Dawkins' [15] examples of memes, such as, pots, arches and clothes fashions, would naturally belong to Cloak's [12] 'm-culture'; only the cultural instructions inside the brain which produced these artefacts would be considered the replicator whose 'purposes' are being served by these things. Dawkins [16] later clarifies his position by adopting the i-culture-m-culture distinction; "I was insufficiently clear about the distinction between the meme itself, as replicator, on the one hand, and its 'phenotypes effects' or 'meme products' on the other. A meme should be regarded as a unit of information residing in a brain (Cloak's 'i-culture')." (Dawkins [16], p109).

Despite this helpful clarification the distinction between memes and their phenotypes has been overlooked by subsequent authors, adding to the confusion generated by Dawkins' [15] lack of clarity. For an example of such confusion I would like to use Dennett's [19, 20] now famous maxim;

"A scholar is just a library's way of making another library." (Dennett [19], p 202; [20], p 346) Whilst there is no doubting the shock value upon the reader, Dennett's maxim is quite problematic. One could compare Dennett's [19, 20] statement to the (Batesman [3]) quotation; "a bird is the nest's way of making another nest", in genetics. Dawkins [16] points out that Bateson's remark, which arose from an argument that genetic determinants of development are necessary

but not sufficient, is wrong in Dawkins' view because the nest is not a replicator. The gene is the replicator, and both the bird and the nest are phenotypic expressions of those genes. As Dawkins [16] puts it; "A nest, like a bird, is the gene's way of making another gene." (Dawkins [16], p98)

Dennett appears to suggest that we should consider that a library (presumably the 'information' housed within a library) is a replicator, and the scholar (a person who studies books) a phenotype. However, Dennett [20] goes on to suggest that language (oral or written) is a 'vehicle' for invisible memes, which are the cultural replicators. This would suggest perhaps that Dennett does not hold the position that a library is a replicator, but something like a meme vehicle. In which case we might be better off saying that; libraries, as vehicles for the memes, serve to promote the replication of those memes (among scholars). Not as punchy as the original, but much clearer I think, and also consistent with Cloak's [12] original work, where a library would be a class of m-culture, the function of which is the maintenance and propagation of i-culture.

There are two points I hope to have brought out from this brief review of the problem. Firstly, that some of the problems of defining memes and making a firm distinction between replicator/phenotype, stem from Dawkins' [15] lack of clarity and the overlooking, by subsequent authors, of Cloak's [12] work and Dawkins' [16] correction. Secondly, that Cloak's [12] work is worthy of greater consideration than it currently appears to enjoy, as it generates some rigid definitions and distinctions which much of memetics has so far lacked.

### 3 Lamarckian Heredity

Why did such usually clear and consistent writers fudge the important issue of replicator/phenotype distinction? I suspect one reason was because Dawkins and contemporaries realised that they were proposing an evolutionary system which appeared to work on Lamarckian-like principles - that is, a system which involved the inheritance of acquired characteristics. Given that Darwinists had spent the best part of 100 years trying to explain why biological evolution was not Lamarckian and trying to undo some of the misconceptions generated by this alternative model, it seems possible that they drew back at the last for fear of resurrecting an old 'enemy'.

Lamarckian heredity is now widely rejected as the mechanism for inheritance within biological systems (although there have been occasional scares, see Dawkins, (1982) [17]. Lamarck's theories on evolution and inheritance have been often mis-represented or mis-understood, (c.f. Hull [26]). Within memetics we tend to refer to the fact that culture is acquired by directly copying the phenotype as 'Lamarckian' (e.g. Boyd and Richerson [8]). Put simply, the view that a blacksmith's son will inherit the developed arm muscles of the blacksmith is as simply untrue as the idea that a dog which loses a leg will give birth to three legged puppies. Organisms do not appear to inherit the changes made to the phenotype during the lifetime of the parent organism. Although this is untrue for biological systems, something like Lamarckian inheritance may exist for cultural systems.

Short of some kind of telepathy it is difficult to see how instructions for a cultural behaviour inside one person's brain could be transmitted to another person's brain without the mediation of

a phenotype. The mediation of the phenotype may be enough to call cultural inheritance 'Lamarckian', but of more interest is the question of whether accidental deviations in the phenotype are acquired by the new host. As Blackmore [4] explains, this is not always the case. She gives the example of telling a story with a hoarse voice. When the new host of the story retells the tale, they do not do so with the same hoarse voice. It may be the case that human beings are very resistant to this kind of 'Lamarckian' inheritance, and that changes accidentally acquired by the phenotype are not automatically adopted by potential hosts.

Although some theorists (e.g. Boyd and Richerson [8]; Gabora [23]) appear certain that cultural inheritance is 'Lamarckian', others (e.g. Dawkins [16]; Hull [26]; Blackmore [4]) seem to have reservations, and question the relevance of the term when applied to cultural evolution. In one sense there is a causal arrow from phenotype to replicator in that the synthesis of a cultural instruction proceeds from the interaction of a potential host and the phenotype of that instruction; and this may be called 'Lamarckian'. In another sense the system appears remarkably resistant to accidental noise within the phenotype, often synthesising the cultural instruction without that acquired characteristic of the phenotype. It is not clear whether cultural inheritance is Lamarckian (in the sense given above), or merely occasionally Lamarckian. Therefore, it seems sensible to conclude that cultural inheritance is mediated through the phenotype, and that the possibility exists that characteristics acquired by the phenotype can also be inherited. It may be better to drop the term 'Lamarckian' altogether (c.f. Blackmore [6]).

Does this imply a causal arrow between the phenotype and the replicator? (Dawkins [16] concern). It does, but only in that changes to the phenotype may be synthesised as 'mutant' instructions in the brain of a host that comes into contact with it. In this way it appears that the mechanism of transmission is simply another example of the difference between memes and genes. Copying the phenotype (when the phenotype has been accidentally changed) could be said to sometimes cause the creation of mutant memes. This form of 'Lamarckism' could be seen as just another potential source of copying infidelity in meme transmission. The 'meme-line' for the construction of the original meme phenotype has ended, and a new mutant 'meme-line' has sprung up. Whilst this is certainly a 'problem' for memes (as all copying infidelity is a problem), it does not necessarily cloud the distinction between memes and their phenotypes or prevent cultural evolution from being considered Darwinian.

Indeed the problem of copying infidelity may give us insight into the deeper mystery of why humans have evolved such complex culture whilst no other animal on Earth has done so. Nature abounds with examples of cultural transmission (e.g. Bonner [7]), but humans appear unique in the complexity and variation of their culture. Many writers have speculated as to the cause of this disparity for example; creativity (e.g. Gabora [23]) or Machiavellian intelligence (e.g. Byrne and Whiten [10]).

Dennett [20] suggests that for cultural evolution to occur the fidelity of transmission must lie within a particular range. The fidelity cannot be 100% because culture would have no variation; which is required for differential survival to drive evolutionary change. The fidelity must still be very high, or else culture would have no continuity between individuals and 'Good Tricks' (Dennett [20]) would be lost as quickly as gained, and again evolution would not occur.

One might speculate that at some stage of evolution human imitation progressed to the point where it crossed a fidelity threshold. It became high enough fidelity for 'Good Tricks' to survive for some time in the system (c.f. Maynard Smith and Szathmary [29], who model permissible genome size as a function of replication accuracy for viruses). It is easy to forget just how hard imitation is, humans are so good at it. Few other species have been shown to truly imitate one another, and none are anywhere near as proficient at the task as even human babies (c.f. Heyes [25]; Meltzoff and Moore [30]). If transmission fidelity is a key factor responsible for the complexity of our culture we would expect normal adults to be well within the expected range of fidelity, whereas humans (e.g. autistic children perhaps) and animal groups which do not readily develop complex culture would not.

## 4 Sociobiology and Memetics

Evolution by natural selection, with the gene fulfilling the role of the replicator, provides formidable explanatory power when applied to biological systems. Dawkins [15, 17] demonstrates how the idea can not only account for complex biological structures, such as the eye, but also social behaviours, such as patterns of aggression. Altruistic behaviour, which appeared a thorn in the side of 'nature red in tooth and claw' models of evolution, received special attention. Biologists, (e.g. Hamilton [24]; Trivers [35]; Dawkins [15]; Cronin [13]; Axelrod [1]; Ridley [33]) have used the 'gene's eye-view' to provide convincing explanations of altruistic behaviour among animals through mechanisms such as kin selection and reciprocal altruism. Considering evolution from the point of view of the selfish gene has also given new life to explanations of flamboyant features among animals, such as the peacock's tail, by allowing a better understanding of the mechanisms of sexual selection (e.g. Cronin [13]).

Despite the success of sociobiology in illuminating possible evolutionary mechanisms which lead to particular forms of human behaviour, many aspects of culture remain beyond the aegis of neo-Darwinian explanation. It is apparent that the evolution of biological phenotypes cannot completely explain the panoply of behaviours which humans exhibit. Kin selection, sexual selection and reciprocal altruism may well describe the mechanisms which operated to form, for example altruistic behaviour, but much of the specifics of those behaviours are not adequately explained by sociobiology. The mechanisms sociobiologists examine allow us to understand why altruistic behaviour evolved in the first place, but examples like blood-donation, where the donor appears to receive no (biological) benefit in return, or adoption, where parents invest in non-kin, appear to require additional explanations.

There appears to be a feeling, particularly among psychologists and sociologists that perhaps sociobiology has gone too far (e.g. Rose [34]), a resistance to what some see as 'greedy reductionism' (Dennett [20]). It is apparent that there needs to be a theory which attempts description of human behaviour at a more appropriate level of analysis.

However, some proponents of meme theory often appear to undervalue sociobiological explanations for behaviour. Meme theory should only be applied where Occam's razor allows. A theory of cultural evolution should incorporate sociobiological findings, only calling upon the addition of cultural mechanisms where they are necessary in order to accurately describe or explain behavioural phenomena. An example from Lynch [28] illustrates this point.

"Eating abundant high-fat food causes people to gain weight as they age. So women's body-fat percentage now correlates with their age. But men who prefer young women can lead longer reproductive careers, replicating their memes into more sons. By thus favouring young women, the lean-partner preference outreplicates the fat-partner preference in modern well-fed societies." (Lynch [28], p 87)

Whether or not this example is true (and there are no sociological references in support of the claim made), Lynch [28] gives us no clue as to why we should ascribe lean-partner preference to purely cultural factors. In the example simply replacing the word 'meme' with 'gene' could make it a sociobiological argument. Surely genes which promote men to have sexual relationships with younger women (because of their longer reproductive careers) could lead to a 'lean-partner preference'.

As well as over-relying upon the power of 'vertical' transmission of memes ('vertical' refers to cultural transmission from parents to children, c.f. Cavalli-Sforza and Feldman [11]), this kind of memetics does not tackle the possibility of other mechanisms for culture (e.g. 'epigenetic' rules; Lumsden and Wilson [27]), and overlooks possible genetic biases in cultural selection and the explanatory power already available from sociobiology. We should not ignore or attempt to rewrite the advances made by sociobiology. It seems likely that there will be many instances where memes and genes have co-evolved (e.g. Durham [21]) 3, particularly when the majority of cultural transmission has been 'vertical'. If we wish to look for examples of apparently 'independent' cultural effects on human behaviour, it might be more profitable to look at biologically maladaptive behaviours which have arisen from predominantly 'horizontal' transmission (i.e. cultural transmission between non-kin, c.f. Cavalli-Sforza and Feldman [11]). Where cultural practices appear to directly increase the chances of biological reproductive success we should acknowledge them as co-evolutionary systems (e.g. Durham [21]) and recognise the important contribution of biology to many of our behaviours.

## 5 Self-centered Selectionism

Finally I would like to examine what I believe is a major misconception which has found its way into memetics. I shall call this misconception Self- centered selectionism. Perhaps Dawkins himself provides the most famous example;

"We are built as gene machines and cultured as meme machines, but we have the power to turn against our creators. We, alone on earth, can rebel against the tyranny of the selfish replicators." (Dawkins [15], p 215)

On the very last page of *The Selfish Gene*, Richard Dawkins closes his argument with a call to overthrow the tyranny of the replicators. In doing so it could be claimed that he committed a common, but fundamental, error; he assumed there was 'someone' beyond the constructs of the memes and genes who could do the overthrowing.

The idea that a Self, beyond the constructs of the genes and memes, can select or design memes is what I call 'Self-centered selectionism'. Self- centered selectionists claim to accept the idea

that memes are the unit of cultural heredity, and even that memes have 'power' to influence behaviour; but then contradict themselves by claiming that it is the 'conscious Self' which selects which memes a person has in 'their' brain. This is the sentiment echoed throughout Csikszentmihalyi [14], and Brodie [9]. The 'someone' beyond the construct of the memes and genes is 'consciousness', they claim. It is apparent that both of them believe that consciousness has the power to select memes in order to fulfill some life goal.

"After all, our genetic program, laid down before our ancestors achieved consciousness, dictates that we place all our efforts into what it takes to replicate our own genes. ... Yet there are also many people for whom the goals of survival and reproduction are not sufficient. It is for these individuals that the possibility of contributing consciously to evolution might be a very attractive proposition." (Csikszentmihalyi [14], p 168)

"You can consciously program yourself with memes that help you with whatever you're up to in life. That's one of the main strategy-memes in the memetics paradigm. It goes against that strategy to believe religious dogma without having consciously chosen it as empowering to your own life." (Brodie [9], p188)

The other idea they both appear to support is akin to the idea of 'directed mutation'; a paradoxical position whereby mutation that occurs is not random, but somehow directed towards some goal. The claim that variation is directed by human intentionality appears utterly incompatible with the idea of culture as an essentially Darwinian evolutionary system.

"At the moment of creation, the meme is part of a conscious process directed by human intentionality." (Csikszentmihalyi [14], p 120)

"In the not-so-distant future, the bulk of our culture will be composed of designer viruses. Why? Because now that we know how to do it, we will. We will conquer the conceptual landscape as surely as we conquered the wilderness. At first, designer viruses will compete with cultural viruses for a share of our mind. Soon the old cultural viruses will lose, because the natural selection with which they evolve is not as quick as the intelligence-directed creation of designer viruses." (Brodie [9], p 200)

The claim that we can intentionally design and choose memes begs the question; 'why do we need an evolutionary theory at all?'. If the processes of meme selection are not 'natural', one is perhaps left with 'conscious foresight' picking and choosing memes towards some goal. Even if we ignore the very real issue of consciousness doing anything, this is still entirely at odds with the proposal that culture is an evolutionary system.

Given the complexity and variety of life found on Earth it is not surprising, perhaps, that in the past people have looked to that complexity as evidence for the existence of a God or supreme designer. William Paley [31] used the example of finding a watch on a deserted heath, and that from the complexity and apparent design of the watch one could infer the presence of a designer. By analogy, he claimed, because complexity and apparent design exists within nature, nature too must have had a designer. Paley was firm in the claim that; "Arrangement, disposition of parts, subserviency of means to an end, relation of instruments to a use, imply the presence of intelligence and mind", and his belief that; "There cannot be design without a designer". It is this

view which has been roundly defeated by the Darwinist revolution within the biological sciences. In biological evolution it is unnecessary to posit some external influence, or guide, or God which directs evolution along. Natural selection can account for the evolution of complexity and design without a designer, without foresight.

If memes are not selected by the same sorts of 'natural' processes which affect genes, it becomes difficult to understand how memes could be said to have 'phenotypic power', or any other kind of power for that matter. If one believes God meddles in biology, one need not posit evolution to explain the forms we find in biology. Likewise, if one believes that 'consciousness' has the foresight and independence to select and direct behaviours towards some goal, one need not posit evolution to explain the forms we find in culture.

People choose memes, but people are constructs of memes and genes. If meme theory is to be given serious consideration then we must first reject the notion that some 'central executive Self' can pick and choose among the memes, and refer instead to the sorts of filters (c.f. Dennett [19]) which memes and genes have constructed. Dennett [19] suggests it is the memes which already occupy the brain which influence the brain to accept or reject new memes which come along.

"But if it is true that human minds are themselves to a very great degree the creation of memes, then we cannot sustain the polarity of vision with which we have started; it cannot be 'memes versus us' because earlier infestations of memes have already played a major role in determining *who or what we are*. The 'independent' mind struggling to protect itself from alien and dangerous memes is a myth; ..." (Dennett [19], p207, original italics)

Understanding the 'natural' selection pressures upon memes is of fundamental importance to meme theory. To attribute selection to conscious foresight, ignoring that (short of some kind of precognition) foresight does not really exist, not only undermines the value of meme theory as an evolutionary process, it is also a kind of giving up. Meme selection is already a big enough mystery without introducing the complex and enigmatic concept of consciousness. We might as well attribute meme selection to magic, for all the progress that *will* achieve.

## 6 Conclusion

I have briefly, and critically, reviewed four important areas which I believe are central to meme theory. The difficulties which I have covered generate concern both within and outside memetics and which need some kind of resolution before major progress can be made towards a science of memetics. Whilst I have aired my views as to how the difficulties in these areas might be settled, these aspects of meme theory can only really be resolved through further debate and eventually research. It is my hope that these short reviews will go some way to stimulating that process.

## Notes

1. I mean 'Self' in the same sense that Blackmore [5] refers to. Blackmore suggests that there is no 'Self' beyond a construct of memes, no one to 'have' a belief, or 'own' an idea. She suggests that the illusion of a 'Self' is itself a meme-complex (i.e. a collection of co-adapted memes)

which is 'unzipped' by Zen training. This general line of thinking is similar in direction to some of Dennett's ideas in *Consciousness Explained* [19].

2. An 'Algorithm' in the sense that Dennett [20] uses the term is broadly characterised as a 'foolproof and somehow "mechanical" procedure' (Dennett [20], p 50). Here I use the term to underline the essential 'mindlessness' of the process involved (i.e. not directed by foresight or a 'Self').
3. Durham [21] uses the idea of memes in his book '*Coevolution*'. He provides some excellent examples and analysis where genes and memes have co-evolved in particular cultures. In his view memes can alter the behaviour of humans in such a way that certain genes are favoured over others; e.g. cultural beliefs about dairying and fresh milk consumption affect the reproductive fitness of genes for adult lactose absorption. Durham comes to the conclusion that evolution involves a dual inheritance system, genetic and cultural; which interact with each other in predominantly co- supportive ways. Whilst this may be true when memes and genes share the same vector of transmission (i.e. from parents to kin), it is likely that when memes begin to use lateral transmission (e.g. to non-kin) they can begin to behave more independently of biological advantage, sometimes leading to biologically maladaptive behaviours (c.f. Cloak [12]).

## References

- [1] Axelrod, R. (1990). *The Evolution Of Cooperation*. New York: Basic Books.
- [2] Barkow, J.H., Cosmides, L. and Tooby, J. (1992). *The Adapted Mind. Evolutionary Psychology and the Generation of Culture*. Oxford: Oxford University Press.
- [3] Bateson, (1978). Review of the *Selfish Gene*. *Animal Behaviour*, 26:316-318.
- [4] Blackmore, S (1996). Memes, Minds and Selves. Seminar paper for the LSE lecture series 'About Biology'. Nov 28th 1996, (also at: <http://www.memes.org.uk/lectures/mms.html>)
- [5] Blackmore, S (1998). [in press] Waking from the Meme Dream. In: G. Watson., G. Claxton., and S. Batchelor. (Eds), *The Psychology of Awakening: Buddhism, Science and Psychotherapy*. Prism Press. (also at: <http://www.memes.org.uk/meme-lab/DART96.HTM>).
- [6] Blackmore, S (1999). [in press] The Meme Machine. Oxford: Oxford University Press.
- [7] Bonner, J.T. (1980). *The Evolution Of Culture In Animals*. Princeton, NJ: Princeton University Press.
- [8] Boyd, R. and Richerson, P. (1985). *Culture and The Evolutionary Process*. Chicago: University of Chicago Press.
- [9] Brodie, R. (1996). *Virus of the Mind*. Seattle: Integral Press.
- [10] Byrne, R.W., and Whiten, A. (1988). *Machiavellian Intelligence*. Oxford: Clarendon Press.

- [11] Cavalli-Sforza, L.L. and Feldman, M.W. (1981). *Cultural Transmission and Evolution: A Qualitative Approach*. Princeton, NJ: Princeton University Press.
- [12] Cloak, F.T. (1975). Is a cultural ethology possible? *Human Ecology*, 3:161-182.
- [13] Cronin, H. (1991). *The Ant and The Peacock*. Cambridge: Cambridge University Press.
- [14] Csikszentmihalyi, M. (1993). *The Evolving Self*. New York: Harper Collins.
- [15] Dawkins, R. (1976). *The Selfish Gene*. Oxford: Oxford University Press. (2nd Edition, 1989).
- [16] Dawkins, R. (1982). *The Extended Phenotype*. Oxford: Oxford University Press.
- [17] Dawkins, R. (1988). *The Blind Watchmaker*. London: Penguin.
- [18] Dennett, D. (1991). *Consciousness Explained*. New York: Little, Brown & Co.
- [19] Dennett, D. (1995). *Darwin's Dangerous Idea*. London: Penguin.
- [20] Durham, W.H. (1991) *Coevolution: Genes, Culture And Human Diversity*. Stanford, CA: Stanford University Press.
- [21] Fisher, R.A. (1930). *The Genetical Theory of Natural Selection*. Oxford: Clarendon Press.
- [22] Gabora (1997) The origin and evolution of culture and creativity. *Journal of Memetics - Evolutionary Models of Information Transmission*, 1, [http://jom-emit.cfpm.org/1997/vol1/gabora\\_1.html](http://jom-emit.cfpm.org/1997/vol1/gabora_1.html)
- [23] Hamilton, W.D. (1964). The genetical theory of social behaviour, *Journal of Theoretical Biology*, 7:1- 32.
- [24] Heyes, C.M. (1993). Imitation, culture and cognition. *Animal Behaviour*, 46:999 - 1010.
- [25] Hull, D.L. (1982). The naked meme. In: H.C.Plotkin (Ed) *Learning, Development, and Culture*. Chichester: John Wiley & Sons Ltd.
- [26] Lumsden, C. And Wilson, E. (1981) *Genes, Mind And Culture*. Cambridge, MA: Harvard University Press.
- [27] Lynch, A. (1996). *Thought Contagion*. New York: Basic Books.
- [28] Maynard Smith, J. and Szathmary, E. (1995) The Major Transitions in Evolution. Oxford: Oxford University Press.

- [29] Meltzoff, A.N. and Moore, M.K. (1977). Imitation of facial and manual gestures by human neonates. *Science*, 198:75-78.
- [30] Paley, William (1812). *Natural Theology*. [15th Ed. 1815 Printed by S.Hamilton].
- [31] Plotkin, H. (1994) *Darwin Machines and the Nature of Knowledge*. London: Penguin.
- [32] Ridley, M. (1997). *The Origins of Virtue*. London: Penguin.
- [33] Rose, S. (1978). Pre-Copernican Sociobiology. *New Scientist*, 80:45-46.
- [34] Trivers, R.L. (1971). The evolution of reciprocal altruism. *Quarterly Review of Biology*, 46:35-57.
- [35] Wilkins, J.S. (1998). What's in a Meme? *Journal of Memetics - Evolutionary Models of Information Transmission*, 2, [http://jom-emit.cfpm.org/1998/vol2/wilkins\\_js.html](http://jom-emit.cfpm.org/1998/vol2/wilkins_js.html)

# **OKAY, BUT EXACTLY “WHO” WOULD ESCAPE THE TYRANNY OF THE REPLICATORS**

## **Reply to the commentaries on the author's paper: Controversies in Meme Theory**

**By  
Nick Rose**

Department of Psychology  
University of the West of England  
Fishponds, Bristol, BS16 2JP  
Nicholas.Rose@uwe.ac.uk

The section of my paper entitled Self-Centered Selectionism actually dealt with several issues, which was clumsy of me. To help rectify this I shall respond to the commentaries by attempting to divide up the issues.

### **1 Self-Centered Selectionism and Self-Centered Variation**

In a Darwinian system the random variation among the replicators allows differential survival to drive evolutionary change. The meme 'eye view' of this process has been characterised by Blackmore (1997) "...imagine a world full of hosts for memes (e.g. brains) and far more memes than can possibly find homes. Now ask - which memes are more likely to find a safe home and get passed on again?" The problem, as I see things, with the ability to consciously choose or reject which memes we have is that it shifts the focus of cultural selection away from the characteristics of the 'meme' and instead concentrates upon the properties of consciousness (with all the associated problems). I am happy to accept that people choose memes - and that they do so in a non-random or directed way. However, I would insist that whatever is directing the selection of memes we are better off examining genetic predisposition and the memes already in situ and trying to understand the differential survival of memes - before resorting to the more enigmatic properties supposed of the mind. Just as the environment for genes includes other genes (and their vehicles), the environment in which memes differentially survive includes other memes. It is these genes and other memes, perhaps in the form of 'filters' (c.f. Dennett, 1991), which I believe make up the 'self' which selects memes.

At the last I have no great argument against the possibility of conscious foresight or intention selecting memes, and I have no trouble admitting this. For me it simply focuses the debate in a rather unpromising direction - and represents a kind of giving up. To that extent I find myself in general agreement with David Hull and Derek Gatherer. One point of contention with Hull is the question of Artificial Selection. Hull suggests that because I rule out intentionality in memetic evolution - that I reject also Darwin's argument that the goals of conscious agents play a crucial role in artificial selection. This is ironic because Darwin spends a large section of his first chapter in 'The Origin of Species' undermining the extent of conscious selection - for his argument builds towards natural selection from the premise of unconscious artificial selection having a crucial role in the form and character of domestic animals. But, nit-picking aside, the example of artificial selection (and modern genetic engineering) is a misleading one. Cultural

practices which have effects upon the phenotypes of animals are (we might claim) the result of memetic evolution. Selective breeding is not random, but then neither is natural selection. The selective environment for domesticated animals also includes humans who, through the course of cultural evolution, have discovered ways of exploiting them and even manipulating their characteristics at the genetic level. I still feel that this does not require conscious foresight to guide meme selection - merely evolved cultural practices to guide the selection of genes in animals. I find this compatible with Darwin's observations and preferable to introducing conscious foresight.

There are many aspects of John Wilkins's commentary with which I agree, but he insists at the last upon the intending self as a social agent. He writes; "*If memes are selected for or against intentionally on the basis of prior experience and propensities to estimate the likely success of a given strategy, that will deform the fitness landscape for these memes and meme-bearers.*" If I could strike the word 'intentionally' I would find myself in perfect agreement. I find the word 'intentionally' utterly redundant here, adding nothing to the description of the process. The 'intentional stance' (c.f. Dennett, 1991) is an excellent heuristic - it allows us to form quick and ready predictions about the behaviour of another - whether that other is Wilkins's lion or another person. Whilst such a 'theory of mind' is undoubtedly useful, I would say (probably along with Dennett) that it is only a convenient short-hand. I might use a similar short-hand to describe other non-intentional processes, for instance; the salt molecules seek areas of low concentration, or the phototropic plant wants to grow in the direction of light. But, at the end of the day it is merely a short-hand for a more complex, and essentially passive, process. Wilkins suggests that the 'self' (as a social agent) has some control over the kinds of memes it hosts, not unlike the control our immune system has over the sorts of micro-organisms that a body hosts. However, we don't require 'immuno agency' to understand the immune system - so why introduce 'social agency' to understand culture? Whilst it may be 'bleeding obvious' to Wilkins that intentions play a part in evolution, I would say it is a short-hand which is 'bleeding redundant' to understanding selection.

Hull is right; I don't care very much for such 'selves'. I find attributing meme selection to a 'self', which exists beyond the constructs of memes and genes, redundant, misleading and lazy. Whilst George Modelska may be right that there is no 'inherent contradiction' between self-centered selectionism and Darwinian evolution, I would agree with Gatherer and Hull who suggest that it is a position unhelpful to the development of memetics. However I maintain that what is essentially fatal to the argument that culture evolves along Darwinian lines is the suggestion that meme variation is a product of a process directed by conscious intentionality and foresight; which I shall dub 'Self-Centered Variation'. If variation among memes is somehow directed by consciousness towards some goal then it is not a Darwinian process. None of the commentators appeared to argue against this position, so I shall hesitantly assume that they were in general agreement.

## 2 Tyranny and Freedom

Memetics poses some difficult questions about free will and the self. Francis Beer describes the uneasy coexistence of free will and determinism within our thinking and theorising and perhaps it is this underlying tension which so often produces impassioned argument. I have nothing

against using the term 'free' in a legal or political sense, indeed these are definitions of freedom which make some sense to me. Certainly a slave has, in a very practical sense, less freedom than I do. However, whatever freedoms I enjoy only exist by convention of the society in which I live. Whenever we speak of freedoms, in a political or legal sense, we also recognise responsibilities. In many instances 'responsibilities' represent things which we are not 'free' to do. As Beer rightly points out, these meanings are themselves 'fluid' and subject to the selective forces we seek to use to explain other aspects of our culture.

Ilfrynn Price attempts to find a compromise in this debate and seeks to soften the dramatic claim made by Dawkins (1976) which I felt typified the Self-Centered Selectionist error (c.f. Rose, 1998). He writes; "*Perhaps it is only through the memes / genes for the inquiring mind and an appreciation of the power of selfish replicators that we may at last escape at least some of the tyranny.*" I would add with a touch of cynicism; 'and run headlong into other tyrannies'. However, to describe the genes or memes as 'tyrants' in the first place is certainly a 'literary excess' - for we are the products of these memes and genes. I find the distinction between us and our genes or us and our memes distinctly odd (and one which smacks of Dualism). At the end of the day who would escape the tyranny of the replicators? To escape the tyranny of memes we should surely have to lose the ability to imitate and instruct and culturally transmit information to one another - and have minds unshaped by culture. To escape the tyranny of genes we should surely have to lose the ability to reproduce - and have bodies and brains unshaped by biology. Not only would such a revolution be impossible, but who would be left after the revolution to enjoy the victory?

## References

- Dawkins, R. (1976). *The Selfish Gene*. Oxford: Oxford University Press. (2nd Edition, 1989).
- Dennett, D. (1991). *Consciousness Explained*. New York: Little, Brown & Co.
- Blackmore, S. (1997). The Power of the Meme Meme. *Skeptic*, 5:43-49.  
<http://www.memes.org.uk/meme-lab/SKEP97.HTM>
- Rose, N. (1998). Controversies in Meme Theory. *Journal of Memetics - Evolutionary Models of Information Transmission*, 2. [http://jom-emit.cfpm.org/1998/vol2/rose\\_n.html](http://jom-emit.cfpm.org/1998/vol2/rose_n.html)

## DARWINIAN SOUP

Review of *The Meme Machine* by Susan Blackmore · Oxford, 264 pp, £18.99

By

**W.G. Runciman**

LRB | Vol. 21 No. 12 dated 10 June 1999 | W.G. Runciman

The word 'meme', popularised by Richard Dawkins in *The Selfish Gene*, has recently gained entry into the *OED* as 'an element of a culture that may be considered to be passed on by non-genetic means, esp. imitation'. But the idea that culture is transmitted by imitation and learning in a manner analogous but not reducible to natural selection has been around for a long time, and many other terms have been used in describing it. Darwin himself, when he came to consider 'the causes which lead to the advance of morality', concluded that 'natural selection effects but little' and looked instead to such things as 'the approbation of our fellow-men' and 'example and imitation'. Much later, when the mechanism of natural selection had come to be understood in greater depth and detail than had been possible for Darwin himself, the American psychologist Donald Campbell and others began to develop the notion that cultural evolution (including the emergence of such things as theories of cultural evolution) is also driven by a competitive process of variation and selection. Campbell called his Presidential Address of 1975 to the American Psychological Association, 'On the conflicts between biological and social evolution and between psychology and moral tradition'. But how does cultural evolution actually work? Any reader sufficiently interested in that question to read this review is likely to be aware that current discussion of it is permeated by controversy. The controversies are of two kinds. For some opponents of 'memetics' in any guise, the mere idea of it is totally misconceived, whether because (as the Reductionists argue) cultural evolution is really all about the workings of our genes, or because (as the Creationists argue) it's really all about the souls implanted in us by God, or because (as the Post-Modernists argue) no so-called theory of cultural evolution will ever explain it any better than any other so-called theory. There are also not a few anthropologists and sociologists for whom the mere association of 'memes' with the name of Dawkins is enough to rule them out of mention in correct academic society. But to those, including myself, for whom the ideas of heritable variation and competitive selection are much the most serious game in town when it comes to explaining, and not merely narrating, the course of human sociocultural evolution, the issue is not whether to incorporate culture as such within neo-Darwinian theory, but how.

The disanalogies with natural selection are obvious, and Susan Blackmore is well aware of them. There aren't, literally speaking, lineages of memes as there are of genes, and memes, unlike genes, can be discarded by the people who carry them and replaced by others. Moreover, the relation of the information encoded in memes to the rituals and artefacts of a given culture is not at all the same as that of the information encoded in genes to the organisms which carry them. But it doesn't follow that memes can't therefore be replicated and diffused by what is still a process of heritable variation and competitive selection. Some neo-Darwinian theorists, although they agree that in the human species genetic is supplemented by cultural inheritance, insist that the function of memes is still to serve the interests of genes: cultural mutations will continue to be replicated only if they somehow enhance, on average, the inclusive reproductive fitness of their carriers. But although that may be true in the long run, it demonstrably isn't true in the short

- and the short run covers some of the widest and most significant divergences between the cultures of the different human populations in the ethnographic and historical record. Blackmore firmly, and in my view rightly, sides with those who maintain that 'memetic' mutations can be culturally adaptive - that is, can be such that their continued replication will be favoured by the relationship of their carriers to their environment - even when they are at the same time genetically maladaptive - that is, such as to reduce the probability of replication of the genes of the people who carry them.

But what is - or isn't - a 'meme'? Blackmore sensibly eschews any attempt to find precise 'memetic' analogies for genetic concepts like alleles, meiosis and so forth. But having reviewed some of the many conflicting opinions in the current academic literature, she concludes that the best thing to do is to 'use the term "meme" indiscriminately to refer to memetic information in any of its many forms; including ideas, the brain structures that instantiate those ideas, the behaviours these brain structures produce, and their versions in books, recipes, maps and written music'. This is not only confusing, but defeatist. It is true that there is no single model which can capture everything involved in the transmission of culture from adjacent or successive populations to one another. But if 'memetics' is, broadly speaking, about the way instructions affecting phenotype are transmitted from mind to mind by imitation and learning, it is presumably worth drawing an explicit distinction between the instructions transmitted, whether visually, orally or in written form, by parents, teachers, peer-group members or role-models, and the behaviour which the instructions dictate. Nobody can deny that such transmission is happening every day. Why, then, should even those whom Daniel Dennett calls 'Darwindreaders' object to the adoption of a single word for whatever specific bundles of instructions affecting phenotype - whether encoded in gestures, speech, writing or any other medium - do in fact guide the behaviour of the members of different cultures or sub-cultures?

The carriers (or as Blackmore prefers it, 'vehicles') of 'memes' are then the minds in which the instructions are held, and the behaviours which result from following the instructions - cooking soups, building churches, singing songs, wearing hats, performing plays, giving birthday parties - are the memes' phenotypic effects. The instructions do not, however, have to be strictly deontic in logical form, or strictly prescriptive about every detail of behaviour. 'Memes' can include discrete units of memorable information (whatever they are) within the totality of values and beliefs which guide our cultural behaviour, as well as mental representations (however they function) of recipes, drawings, musical scores, prayer-books, folk maxims, manuals of etiquette and so forth. The different individual 'cultures' of the human populations in the ethnographic and historical record then consist of sets of related 'memes' (Blackmore calls them 'memeplexes') which have been handed down sufficiently widely and accurately over a sufficient number of successive generations to amount to a definable cultural tradition.

Not even the blessing of the *OED*, however, has converted to 'memetics' more than a handful of the increasing number of behavioural scientists, from archaeologists to linguists to economists, whose studies of one or another aspect of sociocultural evolution are conducted within a recognisable neo-Darwinian paradigm. This may in part be due to an innate suspicion of neologisms. But it is also due to a widespread feeling that 'memes' sound too much like 'genes' for the significance of the disanalogies between them to be allowed for, that the use of the word provokes more definitional quarrels than it resolves evidential disagreements, and that until we

know much more than we do about where and how instructions affecting phenotype are stored in the human brain it is premature, if not seriously misleading, to talk about their replication and diffusion as if we did. Blackmore's answer is to remind us that the theory of natural selection made enormous advances long before the underlying chemistry was understood and to claim that for all our ignorance about how memes are stored and transmitted 'we certainly know enough to get started.' But even those of her readers who are willing to assent to that proposition may be startled by her headlong gallop into problems as complex and intractable as language, altruism, religion and human self-consciousness itself.

The puzzling question about language, as Darwin observed, is why linguistic ability hasn't evolved in other apes as well as humans. In both humans and other primates, there is a correlation between increasing group size and increasing size of the neocortex. But what were the ecological conditions which allowed the benefits of progressively bigger brains to outweigh the heavy energetic costs? And what exactly were those benefits? As Blackmore acknowledges, we have almost no idea when our remote forbears started using language, let alone what language was first used to say. Her claim that their brain size was 'created by memes' - 'memes' here standing for the newly evolved capacity for genuine imitation, as opposed to mere stimulus enhancement or operant conditioning - is purely speculative. It isn't even as if genuine imitation hasn't been observed in other species. Primatologists are divided about what sort of evidence is conclusive for this purpose, but Darwin had no doubt that 'apes are much given to imitation.' Might not language have helped the replication of memes more than memes helped the replication of genes responsible for the 'language instinct'? Without a more detailed model of the feedback between the greatly enhanced capacity for imitation which Blackmore assumes to go with language and the physiological and anatomical changes which language required, it is difficult to see how the strength of her claim can be assessed. Her fellow-psychologists would probably agree that there is strong evidence for a link of some kind between the acquisition of language and the acquisition of a theory of mind which enabled our remote forebears to conceptualise other people as such. But nobody (yet) knows what it is, and invocation of 'memes' isn't enough to furnish the answer.

Similarly, it is far from clear in Blackmore's discussion of altruism how the strength of her claim that 'meme-driven' altruism spreads because people 'infected with altruistic memes . . . are copied more than other people and so these memes spread more widely' is to be tested. There is no lack of evidence that human beings will, under some conditions, behave generously towards others who are not their genetic relatives, and will sometimes do so even where they have no reason to expect that the recipients of their generosity will at some later stage do them a favour in return. But equally, there is no lack of evidence that under other conditions, human beings will cheat and exploit others who *are* their genetic relatives, and that once some members of a group or community start to renege on social contracts it is the behaviour of those who do so which the others will increasingly imitate. Particularly in large groups whose members are not genetically related, it is not at all easy to see how altruism can become the evolutionarily stable strategy unininvadable by defectors who will be imitated in preference to the altruists. So under what conditions can Blackmore confidently predict that altruistic strategies will be sufficiently widely imitated that they will come to dominate selfish ones (both being, presumably, 'meme-driven')? There is by now an extensive literature on this, much of it grounded in evolutionary game theory, including not merely formal models but field-studies and social-psychological experiments. It

would be unreasonable to expect Blackmore to cover it in a single chapter of a short book which has in any case been written more as a manifesto than an encyclopedia. But the implication that it can all be absorbed in, or has been overtaken by, a 'memetic theory of altruism' whose 'main assumption is that people preferentially copy the people they like' may well leave the authors of that literature somewhat underwhelmed.

About religion, where sociobiology has the hardest time linking 'meme-driven' behaviour to inclusive reproductive fitness, Blackmore is on firmer ground. It is altogether more plausible to suppose that religious memes are replicated and diffused because of their selective function at the level of the group than because their carriers are individually more likely to have more children (although there are, admittedly, some well-documented religious communities where this does hold good). Instructions affecting phenotype which are underwritten by beliefs in supernatural sanctions, are encoded in sacred books, and enjoin the performance of elaborate ceremonies may well give the group a competitive advantage over other groups and thus enhance the probability of replication of memes carried by its members relative to theirs. Moreover, it may well do so by at the same time reinforcing altruism within the group through a mixture of punishments and rewards and permitting or even encouraging deceit, manipulation and aggression in relationships with the members of other groups.

Blackmore seems inclined to regard all religions as 'memetically selected' bundles of 'clever tricks' which fool people into believing things which are either untestable or false (as, let us remind ourselves, science often does, too). But this is to ignore the distinction between genuine conversion and mere 'adhesion', as Arthur Darby Nock called it. No doubt religions have often been helped in their quest for converts by apparently successful prophecies, answers to prayer and miraculous interventions in the natural course of events. But for many of its (mere) adherents, religion is not a matter of intellectual commitment so much as emotional involvement in the mourning of the dead, the celebration of victory, the re-enactment of myth, the collective expression of hope or fear, and the symbolising through rites of passage of transition from one stage of life to the next. If 'memetics' is to contribute anything significant to what is already well known to anthropologists, sociologists and historians of religion, it will have to be by identifying specific components of different religious 'memeplexes' whose probability of replication can be shown to have been enhanced by specific features of their carriers' environment - a rewarding task, but a very difficult one.

When it comes to self-consciousness, Blackmore appears to assume that her readers will all be unreconstructed Cartesian dualists still clinging to what Gilbert Ryle famously called the 'ghost in the machine'. But how does 'memetics' resolve the disjunction between the acknowledged findings of psychology and brain science on the one hand and the persistent phenomenological sense of self on the other? For Blackmore, the sense of self is another 'clever trick', which works because the various memes by which we are 'bombed' can, once they get into our heads, 'gain an advantage by becoming associated with a person's self concept'. But if our sense of self, illusory as it may be, is intrinsic to any and all our beliefs, values and preferences, how does it influence the relative probability of replication of some memes as against others? On a 'memetic' view, competition between memes is what drives the various cultures between which the human species is distributed down one rather than another evolutionary pathway. But what actually happens in a culture where anti-Cartesian memes propagated by psychologists and philosophers

start to invade minds in which what Blackmore calls the 'selfplex' has previously been undisturbed? If memes 'associated with a person's self concept' have, on average, an advantage, are anti-Cartesian memes a mutation doomed to extinction? Or will they somehow combine with the memes constitutive of our phenomenologically unreconstructed beliefs, values and preferences in such a way as to be replicated along with them? And if they do, what difference will it make to which of the bundles of instructions by which our cultural behaviour is directed? Will we cook soups, build churches, wear hats, perform plays or give birthday parties any differently as a result?

Despite these reservations, nothing in this review is meant to imply that Blackmore's emphasis on the significance of imitation in cultural evolution is necessarily misplaced. The relation of imitation to individual learning is itself a matter of some dispute. But the acquisition of an entire repertory of cultural behaviour by individual trial-and-error would be prohibitively costly, and the human capacity to absorb and retain rules of conduct observed in the behaviour of other people goes far beyond what could ever be achieved by operant conditioning and stimulus enhancement alone. On the other hand 'socialisation', so called, is not just a matter of children doing what they see their elders doing. It is also a matter of their doing what their elders tell them to do even if the elders don't even pretend to do it themselves: 'don't do what I do, do what I say!' And what about cultural behaviour which is the result not of imitation or learning but of enforced obedience to instructions from lawgivers and other people in positions not merely of influence but of power?

It may be that I am here disclosing a professional bias, since I am not a psychologist or anthropologist but a historical and comparative sociologist more at home in topics like class conflict, the formation of states and the displacement of slavery by wage-labour. But there is more to cultural evolution than just the relative psychological attraction of different memes. Blackmore sweepingly asserts that 'social scientists study the way that people's lives and selves are constructed by their roles, and by the texts in which they are embedded. But they have no evolutionary theory within which to understand the processes going on.' That, however, is simply not so. I am not the only sociologist in the world who sees social roles as part of the same continuous evolutionary process out of which first genes and then memes have emerged and which stretches all the way from the chemical evolution of the nucleic acids to 'intelligent' machines. But social roles involve a further mode of information transfer affecting phenotype. As well as instructions transmitted genetically by inheritance and instructions transmitted memetically by imitation and learning, there are instructions encoded in formal rules underwritten by institutional inducements and sanctions which define the practices that make our roles what they are; about these Blackmore has little or nothing to say.

How successful, then, is *The Meme Machine* likely to be as, under Blackmore's definition, a meme designed to further the replication and diffusion of the 'meme' meme? The answer, I suspect, is that it is unlikely to do for public understanding of the meme-centred approach to cultural selection what Dawkins's *The Selfish Gene* has done for the public understanding of the gene-centred approach to natural selection. This isn't just because Blackmore doesn't have Dawkins's altogether exceptional gift for accurate synthesis and lucid exposition. It is because she attempts too much too soon. No author of a book like this one can yet draw on the large volume of well-tested hypotheses and closely-documented examples available to writers on

biological theory. Only detailed and wide-ranging re-examination of the relevant ethnographic, historical and archaeological evidence for the heritable variation and competitive selection of clearly identifiable 'memes' whose spread can accurately be traced and functions convincingly specified will turn 'memetics' from a project into a science.

**W.G. Runciman** is the author of *The Social Animal*.

# **DARWINIAN PROCESSES AND MEMES IN ARCHITECTURE: A MEMETIC THEORY OF MODERNISM**

**By**

**Nikos Salingaros**

Division of Mathematics,

The University of Texas at San Antonio,

San Antonio, Texas 78249, USA.

[salingar@sphere.math.utsa.edu](mailto:salingar@sphere.math.utsa.edu)

**Terry Mikiten**

Department of Physiology, Graduate School of Biomedical Sciences,

University of Texas Health Science Center at San Antonio,

San Antonio, Texas 78284, USA.

[mikiten@uthscsa.edu](mailto:mikiten@uthscsa.edu)

## **Abstract**

The process of design in architecture parallels analogous generative processes in biology and the natural sciences. This paper examines how the ideas of Darwinian selection might apply to architecture. Design selects from among randomly-generated options in the mind of the architect. Multiple stages of selection generate a design that reflects the set of selection criteria used. The goal of most traditional architecture is to adapt a design to human physical and psychological needs. At the same time, however, a particular style of architecture represents a group of visual memes that are copied for as long as that style remains in favor. Darwinian selection also explains why non-adaptive minimalist forms of the modernist style have been so successful at proliferating. The reason is because they act like simple biological entities such as viruses, which replicate much faster than do more complex life forms. Simple visual memes thus parasitize the ordered complexity of the built environment.

**Keywords:** adaptivity, architecture, Darwinian process, design, images, memes, minimalism, modernism, parasitic, selection, styles, viruses.

## **1 Introduction**

The world of the architect is created in an architect's mind according to certain principles related to the biology of the brain. According to one theory of the thinking process, an idea arises out of the competition among similar and dissimilar ideas occurring simultaneously in adjacent neural circuits of the brain [6; 7]. The same principles of competition and selection might be said to apply to the general public in accepting architecture. Things in the built environment originate and endure (not just in the tectonic sense, but in their survival value in a society's common language) because they 'make sense' in some way. Competing ideas in a society eventually suppress or reinforce each other to produce one or more dominant themes. In other words, creativity and survival work in ways that are compatible with the cognitive machineries that make up the mind.

Nevertheless, sometimes the mind works against the body by acting in a harmful manner. An architect's mind has the power to either create designs that adapt to human needs and emotions, or to impose arbitrary images on the environment. A Darwinian selection process in architecture takes place among competing ideas in the mind. A second process, also Darwinian, occurs in the society of consumers. This second selection process is between styles, where certain styles win over others. In both of these selection processes (i.e., in the mind, and in society), the criteria are a mixture of human needs and irrelevant stylistic fashions. Meme propagation and encapsulation explain why these two disparate sets of selection criteria can coexist, and when one set can displace the other.

This paper applies the theory of memes to the field of architecture. Two main points are argued: (1) Darwinian processes (combining variation and selection) are important to architecture; and (2) the specific case of modernist architecture corresponds to a 'parasitic' meme, which has spread in spite of its being non-adaptive for the people that make use of modernist buildings. These two theses are logically independent, though both are necessary to present a picture of how architectural styles propagate. The first thesis may seem rather obvious to many readers, yet it has important implications for the design process in general. The second thesis is more controversial, and is discussed in greater depth. Explaining the unlikely success of modernism by other than subjective criteria is our eventual goal.

Design is a problem-solving activity. Human intelligence allows both the generation of possible alternative solutions and selection among them to take place mentally. That neatly summarizes our intellectual advantage over other animals: our imagination is a profoundly useful virtual reality simulator. A more intelligent system will have a more efficient mental representation and selection process. The architect's mind is impacted by the problem space and various 'memes' (conceptual entities that propagate among human minds) from a variety of sources [14]. These could come from one's own memory; visual templates from the environment; the influence of other architects; etc. Competing forces such as engineering constraints, a desire for creativity, and the unique need to express oneself drive the design to its final state.

A Darwinian process in the mind of the designer depends on a set of selection criteria. Traditional societies such as pre-industrialized people, and the industrialized nations up to and including the nineteenth century used a wide range of selection criteria that, among other practical constraints, enhance emotional well-being for the user. Specific architectural styles, however, replace the selection criteria of traditional adaptive design by a matching to visual templates, or 'memes'. Once adaptive design is abandoned, the spread of architectural styles depends strictly on factors governing meme propagation in a society. A minimalist style then possesses an unbeatable advantage over more complex styles, because of its low information content. It is possible to explain in this fashion an important event in the evolution of mankind: the drastic change in the visual character of the built environment during the twentieth century.

The successful spreading of modernist design is interpreted in terms of the replication of memes. After a design style is introduced and is accepted for whatever reasons by a group of people, then it is subject to Darwinian selection from among the pool of competing styles. This is where consumers, the construction industry, and the architectural establishment come into play by exerting selection forces. A second selection occurs entirely outside the architect's mind, within

the arena of human society [13]. Some architectural styles die out, whereas others survive and become popular. Their success has little to do with their fitness for human habitation; the criteria for success in Darwinian selection are abstract and are not based directly on human needs, even though it is human beings that do the selecting.

Studying how architectural memes spread in a society, and how competing memes are selected requires a knowledge of the factors affecting meme propagation. Francis Heylighen has identified a list of these. We will discuss seven of his factors: SIMPLICITY, NOVELTY, UTILITY, FORMALITY, AUTHORITY, PUBLICITY, and CONFORMITY in the context of architecture [17; 18]. With the exception of UTILITY, none of these factors serves actual human needs. We will argue, therefore, that the spread of a design style occurs in a society more because of mass media than for practical reasons. Even UTILITY will be shown to obey memetic transmission, as often the mere *promise* of UTILITY is responsible for the success of an architectural style that creates buildings impractical in actual use.

We will propose an eighth factor that aids meme propagation: ENCAPSULATION describes how memes link with other memes. This process confers an advantage to the encapsulated meme because: (a) it increases the meme's virulence by making it appear more attractive; and (b) it protects the meme from external challenges by insulating it inside a complex of other, beneficial memes. An encapsulated architectural meme manipulates our emotions in order to propagate. ENCAPSULATION embeds a meme or collection of memes into a meaning structure [22]. Through this mechanism, visual memes acquire an emotional and physical basis. At that point, they cease to be regarded as mere ideas open to debate, but assume the fundamental character of beliefs defining one's consciousness.

It is also possible to condemn an architectural style by deliberately encapsulating it within a shell of negative associations. By using ENCAPSULATION as a weapon to discredit competing styles, a useful idea can be tainted (whether there is any basis for that association or not). A society's collective unconscious from that point on automatically rejects such an idea or style without question, even though it may offer excellent solutions to urgent problems. In contemporary architecture, destructive encapsulation is used to discredit new buildings in the Classical and Nineteenth-century styles. This has happened despite the fact that earlier buildings in those styles are among the most comfortable and best adapted to human needs. We will argue that by encapsulating them with pathological memes, those styles have effectively been placed in quarantine.

Success in the spread of social memes is measured by how far they establish themselves as basic beliefs in a society. This paper explains their spread from the point of view of Darwinian processes. A group of memes achieves its greatest success when it becomes part of the establishment; i.e., it is institutionalized. We are first going to deal with those factors that increase the spread of memes, and thereby help in their chances for institutionalization. In the final section, we explain how once memes have been institutionalized they acquire a rigidity that makes them extremely difficult to remove. The institutional perspective offers some strong explanations for the remarkable persistence of modernist architecture and urbanism in spite of their negative aspects.

## 2 The modernist style and military architecture

The modernist style of architecture, otherwise called the *International Style*, has been the overriding building style from the 1920s until now. The style is instantly recognizable by its geometry of cubes and rectangular slabs; flat plain surfaces; the lack of thick connective boundaries; the use of steel, glass planes, and concrete panels; and in many cases the elimination of color and structure on the human range of scales 1mm-1m [23]. Representative buildings and architects include the *Bauhaus building* (1926) by Walter Gropius; the *Pavillon Suisse, Cité Universitaire* (1932) and *Carpenter Center for the Visual Arts* (1961) by Le Corbusier; the *Casa del Fascio* (1936) by Giuseppe Terragni; the *UN Headquarters* (1950) by Wallace Harrison and Max Abramovitz; the *Seagram building* (1958) and the *Neue Nationalgalerie* (1968) by Ludwig Mies van der Rohe; and the *National Theatre* (1967) by Denys Lasdun.

Modernists claim their buildings to be 'functional'. Nevertheless, simply looking like a machine from the 1920s doesn't guarantee functionality in a building. Those machines were housed in smooth metal shells, following cubist aesthetic principles, so their 'look' had nothing to do with their function: it merely conformed to a passing artistic fad. A culture that substitutes images for the real thing risks losing its accumulated knowledge. Many authors claim that this has already happened, since our generation has lost innumerable adaptive architectural traditions stretching back several millennia. More recently, 'high-tech' has become the fashionable international style of corporate architecture, simply because its superficial appearance of metal pipes, glass, mirrors, and plexiglass links it to modern technology; this goes on despite high-tech's extremely high cost and low user comfort. Representative high-tech buildings and architects include the *Centre Pompidou* (1977) by Richard Rogers and Renzo Piano, and the *Hong Kong and Shanghai Bank* (1986) by Norman Foster.

Two contradictory movements in twentieth-century architecture work against adaptive selection towards the human needs of users. The first is an attitude taught in recent decades by our schools: that an architect has artistic license to ignore certain practical constraints — indeed, that it is necessary to do so — in the pursuit of a 'great work of art'. The second is a standardized approach to buildings, behind which is a conviction that shaping design to particular individual needs is simply being self-indulgent and therefore socially irresponsible. Early modernists set up standards for minimal dwellings that had little relation with the living needs of real human beings, and incredibly, most of them are still applied today. A central idea in German social housing of the 1920s, the *Existenzminimum* [4] codified the minimum space in which a German blue-collar worker and his family could be housed. That is where oppressively low ceilings and cramped, tiny kitchens in today's apartments originate.

Although traditional architecture for human use adapts to human needs and sensibilities, military architecture is the exception. A well-defined typology has been used throughout the ages to construct deliberately uncomfortable environments. These include defense installations and castles (experienced from the outside), and dungeons, prisons, crematoria, etc. (experienced from the inside). Such environments lack texture, color, and decoration, preferring damp, grey surfaces that are usually punishing for human beings. Their forms and surfaces are meant to oppress and frighten us: they communicate danger and evil directly through architecture. Where possible, a grandiose scale dwarfs the role of a human being in the environment. To achieve a

forbidding, hostile exterior, a building must reveal a minimum of information. This makes sense when defensive fortifications protect against attack by infantry.

There are obvious stylistic similarities between military and modernist architectures, since many modernist buildings look forbidding, ominous, stark, alien, faceless, and present a generally hostile appearance. The reason for this impression is that they utilize some of the same typology from military and prison architecture. Here we face a paradox: how could society select an architectural style for human use that has a similar typology as the military style, which was developed specifically to make people feel uncomfortable? Our explanation is that modernist architecture is a 'parasitic' meme group that is non-adaptive to human use and sensibilities. At the same time, however, the group of memes defining the modernist style of architecture has memetic advantages that helped it to take over. It is for this reason that modernism won out over competing styles.

### 3 Design as a Darwinian process

Design ought to begin by understanding a building's particular uses. A designer is aided by recalling built examples that work under similar circumstances; this is the idea behind *Alexandrine Patterns*, which distill working solutions from widely different cases [1]. Alexander *et. al.*'s *A Pattern Language* provides a collection of design constraints extracted from traditional architecture the world over that are meant to anchor and guide an emerging design [24]. These aim to make the designed structure adaptive to human needs, while leaving the form and visual aspect unspecified. It really doesn't matter what triggers one's creativity, as long as the generated alternatives cover a broad enough range, and the selection is adaptive. The possibilities of a Darwinian process of design are tied to the system of options within which it operates, and the richer the system is, the broader the field.

Each design competes in the mind of the designer with other conceived possibilities, and the fittest ones (those that partially solve the problem as posed) survive to the next generation. More detailed designs generate further alternatives, which are culled by selection in the subsequent round. The cycle starts with the creation of variants, which then get culled by using a set of selection criteria; the survivors are used to create a new generation of variants, which get culled in turn; and so on. This represents a typical Darwinian process [6; 7; 8]. Visual inspiration can fix the entire gestalt of a project in a single initial image. Often, it is precisely such a *conceived* image that, through the emotional feedback it generates in the mind of the architect, sustains the design and drives it towards completion.

When architects turn for inspiration to fixed images from a set vocabulary defining a style, images displace the adaptive component of design by changing the selection criteria. Design then becomes a process of comparison with certain visual stereotypes, which radically affects the end product. Matching to currently popular images takes priority over all the other design constraints. The new selection criteria may not aim at adapting a design to human needs. The selection process itself ceases to be recursive because selection occurs only on the first level, which is derivative of memory and stored images. If structural, functional, and practical constraints are abandoned in the interest of maintaining images, however, such a design method acquires advantages of economy over more complex approaches that are adaptive.

By accepting input from natural objects (e.g., Le Corbusier from a crab shell he found on the beach and later used to model the Pilgrimage Chapel *Notre-Dame-du-Haut* at Ronchamp); man-made artifacts; buildings originally intended for another use (e.g., the fascination of Walter Gropius and Le Corbusier with American grain silos); modernist architects chose not to follow an adaptive process for turning their inspiration into a practical design. Copying an image is very easy to do and gives a superficial sense of understanding while ignoring the complexities of both the copied structure, and the needs of what is being designed. Grain silos were the end-result of adaptive design for agricultural storage, not for habitations. Copying the 'look' of a structure developed for something else, and applying it to a use for which it was never intended, is not adaptive. A crab shell is beautifully adapted to house a crab, but not for its magnified shape to house human beings wishing to worship in a church.

## 4 Memes and architecture

The word 'meme' denotes any idea that endures and propagates [5; 11; 12; 14]. Memes have a lifetime; they can 'die' when they cease to be of interest to the population for whatever reason. If memes die, then in a given collection of them, one can speak of the survival of some, and the death of others. Survival in an environment, coupled to forces that promote mutation and change leads to Darwinian selection. The concept of memes thus has explanatory value. It also has heuristic value because it forces us to examine how ideas persist and propagate. An image will stick in memory if it is encapsulated in a meaning structure. An image will be more likely to be transmitted to others if it is easy to remember, among other factors to be discussed below.

While our topic is architecture, it is instructive to discuss for a moment a parallel situation in biology where these ideas are routinely useful. In considering how microbes attack tissue, as for example those in the oral cavity that cause tooth decay, the scientist studies the tendency of a microbe to adhere to the tooth surface. Microbes that have the greatest stickiness are also likely to have the greatest virulence; i.e., cause the most serious disease. The logic is straightforward: the stickier the microbe, the greater the number that will adhere to the tooth at any one time. Research shows that the surface of tooth enamel has a certain chemical structure, and the virulent microbes have a corresponding chemical structure that binds to it; rather like the two mating surfaces of Velcro.

Individual memes, or images, are the equivalent of agents that 'infect' memory. Each image has a set of attributes that makes it more or less likely to stick in memory and to be transmitted to others. In the universe of Art and Design this mechanism is readily apparent. The volatility of design themes drives the world of fashion, where the business and sales force creates a strong pressure for selection that is Darwinian at its core. New mutations arise with regularity, and these are tested against the environmental forces in which they appear. The life and death cycle can be swift for unsuccessful fashion styles. The same is true in architecture, where there is an undeniable and changing 'fashion'. Nevertheless, a fashion arrests the adaptive design process, in which selection evolves specific solutions to individual problems that are exquisitely suited for their job and surroundings [24].

Architectural memes are more nearly analogous to physical replicating entities such as viruses, than to more general memes representing only ideas. The reason is that the former are encoded

as actual structures (other than neuronal circuits). It is only their replication that occurs through memetic transmission; the artifact in this instance has a physical existence outside the human mind. An architectural style thus exists in two very different forms: (i) as an ideology codified in books and taught as a tradition in architecture schools, which perpetuates a group of memes in people's brains; and (ii) as images represented in the built environment. Each aspect reinforces the other. The built environment serves as a source of continuous re-infection by visual architectural memes. The image/building/image cycle has positive feedback, and can lead to an exponential rate of infection.

While architecture is often classed along with music, poetry, and the fine arts as a vehicle for individual artistic expression, it is actually far more than that. Humanity needs to house itself, and architecture represents a world-wide building industry that is forever looking for prototypes to copy. The vast majority of buildings, be they commercial or vernacular, require a typology of reproducible patterns. Clearly, the process by which architectural styles spread through copying is one that lends itself to a memetic explanation. This is seen in practice, where throughout history, a single example was often sufficient to establish a new style of architecture. Even though the early buildings defining a new style could number only a handful, their true impact lies in their easy repeatability. Conversely, a style that is difficult to reproduce will die out. The style succeeds not because its original examples are either attractive or useful, but because it infects the vernacular building tradition.

## 5 Explaining the unlikely success of Modernism

In 1922, Le Corbusier exhibited a series of drawings labeled "A Contemporary City" at the *Salon d'Automne* in Paris; he built the *Pavillon de l'Esprit Nouveau* for the International Exposition of Modern Decorative and Industrial Arts, held in Paris in 1925; Walter Gropius built the *Bauhaus building* in Dessau in 1926 as a visual example; Ludwig Mies van der Rohe organized and contributed to the mass housing projects for the Stuttgart *Weissenhofsiedlung* in 1927, which consisted of very similar white, rectilinear, flat-roofed temporary and permanent buildings. All of those buildings and drawings provided images for young architects to copy. The reason anyone would even consider such plain prototypes was the promise of inexpensive housing for all made possible by modular design, bolstered by proclamations of links to a 'new' society.

The rate of transmission of a visual style among human minds depends on several factors. Considered simply as information, the success of an architectural style is governed by the speed at which the associated memes can propagate. The situation is akin to percolation or diffusion: copies of an object (a piece of information encoding the style) have to pass from one human mind to another. This resembles the mechanism by which infectious agents spread in a population. Individuals in the population have little control over the process. Propagating agents are obviously not selected by the host, since they parasitize their more complex hosts. The process is infection rather than competition. An epidemic occurs when a virus has evolved an unbeatable advantage over its hosts.

Francis Heylighen has identified factors contributing to meme propagation [17; 18]. The meme could be an image, or a set of rules defining an architectural style. We will examine four factors now: **SIMPLICITY**, **NOVELTY**, **UTILITY** and **FORMALITY** are relevant for the initial spread

of modernism, and three more of Heylighen's factors help in the institutionalization of modernism (there are other factors that are not discussed). One of Heylighen's criteria is SIMPLICITY. A simple idea is easier to reproduce and has a competitive edge over ideas that are more difficult to grasp; it poses a lesser burden on our cognitive system [18]. Therefore, an architectural style that is simpler to encode will propagate more successfully than one that is difficult to encode. In an analogy with life forms, viruses reproduce much faster than more complex organisms because of a reduced structural investment. Biological and computer viruses take advantage of their host's structural complexity, using it to propagate themselves, and without which they could never replicate at all.

The early modernists introduced images of geometrical emptiness with enormous replicative power. The modernist vocabulary of plain, featureless surfaces in a flat geometry of cubes and rectangles eliminates substructure; eliminates borders; eliminates contrast and color in design by using only plain white or gray; and finally, tries to eliminate the building material itself through its replacement by glass panes [23]. We see some or all of these features used in a majority of buildings throughout the second half of the twentieth century. Design richness and complexity in other architectural styles was eliminated in the drive to reach forms with minimal information content. Architects in the 1920s working in a traditional style originally dismissed this effort as perverse and not worthy of notice; they little realized that it satisfied the SIMPLICITY criterion for memetic propagation.

Another criterion is NOVELTY, where standing out and thereby attracting one's attention facilitates a meme's assimilation. New, unusual, or unexpected ideas arouse one's curiosity [18]. Twentieth-century architecture used novelty of a deliberately shocking kind. The early modernist prototypes looked strange to people used to Nineteenth-century architecture. Indeed, the modernist style is arrived at by reversing elements of previous traditional styles [23]. The spread of those novel images occurred primarily through the media before any significant number of examples was actually built. Le Corbusier was remarkably successful at propagating modernist memes through the journal *L'Esprit Nouveau*, which he controlled [9]. That was the age when picture magazines became a popular medium for visual information, helped by technical advances in photography, printing, and distribution. People were eager to read about new ideas, especially if they were accompanied by futuristic-looking illustrations.

UTILITY plays a double role here. First, the architectural media declare (without justification) that a minimalist structure is somehow more efficient or is better adapted to the functions it is supposed to house. The opposite is true: many modernist buildings are dysfunctional because their imposed form and impractical materials hinder human activities. Criticisms of modernist buildings and their materials include the impossibility of effective temperature control in a glass-walled structure; the tremendous energy waste in attempting to do that in a sealed building; the 'sick building' syndrome; the social damage of living in skyscrapers (most severe for children and the elderly); the dangerous wind shear created on the ground by smooth-faced skyscrapers; flat roofs that invariably leak; the staining or cracking of large, plain surfaces; a general problem of joints when connective interfaces are eliminated in the interests of style; psychological alienation produced by dead gray surfaces and concrete slabs, which give an unpleasant 'hard' echo; etc. [3]. Still, the mere *promise* of UTILITY is often enough for propagating spurious ideas [18].

Second, the modernist style represents a genuine advantage for the construction industry that can build cheap, minimalist box-like structures without having to worry much about either structural quality, or accommodating human physiological and psychological needs [2]. A visually simplistic architectural style thus offers a commercial benefit via UTILITY that counts as a major factor in its propagation. Modernist memes found a ready environment after the Second World War, when buildings had to be produced in large numbers and at low cost. Never before in history had such building efforts taken place. This was also in the period that the industrialization process was at full speed, penetrating more and more economic sectors of society. The construction industry eagerly embraced the UTILITY offered by modernist memes. Philip Johnson (the first American architect to convert to modernism) frankly admitted that: "*The International Style did sweep the world because it came along at the same time developers wanted to make cheap buildings, and this was cheaper than other architectures*" [19].

Yet another factor is FORMALITY: the more formally an idea is expressed, the more likely it survives in transmission [17; 18]. The intention of modernist design is to be context-independent. Adaptation requires selection on the basis of local climate, materials, culture, and relationship to adjoining buildings and specific human needs. Since its inception, however, modernism has been 'universal' because it is based on a small set of simple images. Different individuals in different contexts can interpret modernist rules in the same way. A modernist building can be put up anywhere in a city, anywhere in the world, because the style is independent of locality or particular circumstances. Materials of choice are pre-formed panels, glass, steel, and reinforced concrete; these are industrial materials that are detached from any region. Modernism imposed the universal visual language of abstract cubism to come up with "*one single building for all nations and climates*" [3].

Non-adaptivity to human needs, which helps in memetic propagation, is rooted in modernist ideology. The philosophical origins of modernism in Germany of the 1920s reveal a parallel between modernism and totalitarianism [25]. The German art historian Wilhelm Pinder (a supporter of Hitler) and his student Nikolaus Pevsner (an architectural historian who was one of the strongest promoters of modernism as a guide for social and political ideals) argued that great architecture is the product of the *Volk*, during periods when ideology triumphs. Adolf Hitler, Josef Goebbels, Walter Gropius, and Ludwig Mies van der Rohe all shared the conviction that architecture was an expression of the central spirit of an epoch, and thus justified idealism, absolutism, and arrogance [25]. In this view of the world, the individual is insignificant, and the needs of the human user are thus of little consequence [25]. Philip Johnson complained of the futility of trying to discuss the aesthetics of modernism with Walter Gropius: "*Talking to Gropius was a dead end because he would still mouth the Giedion-esque platitudes of social discipline and revolution*" [9].

## 6 Competition among early twentieth-century architectural styles

The unbiased human mind applies selection criteria that give the most positive emotional feedback from a built structure. This is something we have evolved to do: we instinctively avoid pain and discomfort and seek pleasure. If a design (and, by extension, the building when finished) provides joy to the architect, then one can expect the user to share that experience. The same does not follow, however, when purely intellectual selection criteria replace those based on

emotions. What one person believes in ideologically is not necessarily shared by others. Modernism was very successful at convincing people to forgo sensual pleasure from built forms, as minimal surfaces and spaces offer less visual stimulation than human neurophysiology is built to handle [22]. Memes help us to understand why architectural styles that give emotional satisfaction were replaced by those that don't.

Once built, structures survive or not according to Darwinian selection. Here we are no longer talking about Darwinian processes in the mind of the architect, but survival in the outside world. Occasionally buildings get destroyed by natural or human acts; most often conscious decisions are made on whether to repair the inevitable wear in an existing building, or to build a new building altogether in its place. In biology, the survival rule for a species is to procreate before death. Culling of organisms is determined by survival in the environment. In architecture, survival of a particular style depends on whether the buildings representing that style are preserved, or are replaced by those of another style. Architectural survival therefore depends upon decisions that are heavily influenced by stylistic concerns.

Different styles competed with each other at the beginning of the twentieth century. Any architectural style that contained traditional elements was doomed to extinction because people now demanded NOVELTY. Styles that had comparable NOVELTY were further selected on the basis of SIMPLICITY, UTILITY, and FORMALITY. Art Nouveau is very high in information content. The convolutions, curves, and complex colors on which the style depends ought not to propagate rapidly, and that's exactly what happened. Despite an initial flourish, Art Nouveau didn't last for more than about a decade. Its markedly plainer successor, Expressionism, was equally short-lived because of its curvature. Art Deco abandoned the curves of Art Nouveau and Expressionism, adopting a more rectangular geometry, and was much longer-lived. One could surmise that, by lowering its information content, it acquired greater staying power. Finally, Modernism got rid of the visual richness of Art Deco, reducing its information to an absolute minimum; it won out over its competitors by spreading around the world and surviving until today. These events in architectural history support a memetic theory of architectural styles, with selection on the basis of SIMPLICITY.

Looking at both UTILITY and FORMALITY leads to the same conclusion. Unless there is a strong societal demand for information-rich buildings and environments, the construction industry will select those that are visually plain (since they are often cheaper to build, though not to maintain). As far as FORMALITY is concerned, a set of context-independent rules was never given for either Art Nouveau or Art Deco. We have no formal set of symbols that can generate an Art Nouveau building. The style depended upon the individual creative genius of say, Louis Sullivan or Victor Horta, who drew their inspiration freshly from each new architectural context. It is worthwhile noting that a highly successful style in architectural history, the Classical Style, also depends on rather precise formal rules that can be applied in any situation regardless of context. Nevertheless, its much higher complexity compared to modernism allows adaptive design.

A large number of Art Nouveau and Art Deco buildings were built in the early decades of the twentieth century, before the modernist selection criteria took hold. Many of those buildings did not survive, precisely because the selection criteria used in the 1960s for preserving buildings

were the same as those for designing new buildings. The stereotyped visual template of a glass box determined which buildings to save from demolition, and those buildings that did not match were destroyed. In effect, structures were categorized according to modernist images, thus providing a mandate to eliminate those judged to be 'misfits' One reads that the architecture of the twentieth century is founded on rational laws as opposed to base emotions, which is quite true. Nevertheless, this statement ignores the incredible persistence of modernist memes, which is fundamentally emotion-based. This emotional dimension of memetic transmission will be discussed next.

## 7 Encapsulation of images in the mind

Entities with a finite lifetime will survive in the sense of propagating their information only if they are favored by selection forces. An encapsulating shell that surrounds a meme by an attractive verbal explanation endows it with a meaning system that helps in transmission. Once inside a mind, a meme will lose its boundaries as it is sacrificed to a larger meaning structure that is expressed as a physical grouping of neurons. The mind shows itself to be a multiply-connected network, where ideas, opinions, factual knowledge, and prejudices are all interlinked into what may be called one's 'consciousness' [16; 22]. In this way, memes influence an individual's thoughts and actions. This is the idea behind advertising: embedding a commercial product into a person's consciousness will guarantee the use or purchase of that particular product as the result of a subconscious decision [5].

Architecture and advertising act in much the same way. After they are taught in Architecture school, the memes of an approved style become a permanent part of one's thinking patterns. They are encapsulated into meaning structures such as metaphors. The group of neuronal circuits consisting of images, their encapsulation, and their interconnections defines some domain of one's consciousness. Those regions of the brain provide bases for meaning structures, which are used to interpret the world throughout one's life [16; 22]. We are thus programmed to automatically replicate memes whether or not they are good or useful; that's because they are part of a person's inner belief system [5]. This also explains why visual icons can rarely be dislodged by scientific arguments. A simple but irrational belief can displace an accurate but more difficult description of the world.

This section proposes another factor affecting meme propagation: ENCAPSULATION. A meme boosts its virulence by linking itself to other attractive memes, which then shield the original meme. (This is related to but distinct from Heylighen's criterion of COHERENCE, wherein the assimilation of new ideas depends on their being consistent with existing knowledge [17; 18]). The advertising industry is founded upon techniques of encapsulation: either physical packaging, or the packaging of products within ideas. A commercial product sells just as much because of an attractive package as for any other factor. An effective marketing strategy links a product via emotional appeals to self-esteem, sex, status, power, individuality, etc. It is not a coincidence that modern advertising techniques developed alongside modernist design, and early modernist architects showed a keen interest in psychological manipulation as it was then being incorporated into the advertising industry [9]. Le Corbusier actually made a living from mass media and commercial promotion independently of his work as an architect [9].

This concept applied to architecture reveals an unexpected yet major reason for why architectural design evolves the way it does. A change in encapsulation comes from societal discontinuities, which affect architecture just as much as practical matters such as the introduction of new materials and novel methods of construction. For example, immensely powerful social forces unleashed between the two world wars led people to adopt modernist design memes as a reaction to class oppression. They identified decorated buildings as visual symbols of what was wrong with the past. With all the old values discredited by the horrors of the First World War, people eagerly embraced new ideas, thus linking desirable social aims to encapsulated memes. They willingly sacrificed pleasure from their surroundings for the promise of a better future. The *1918 Manifesto* of the Dutch group of modernist architects known as "De Stijl" states [10]:

*"The war is destroying the old world with its content ... The new art has brought to light that which is contained in the new consciousness of the age ... Tradition, dogmas and the predominance of the individual stand in the way of this realization. Therefore the founders of the new culture call upon all who believe in reform of art and culture to destroy those obstacles to development ... The artists of today, all over the world, impelled by one and the same consciousness, have taken part on the spiritual plane in the world war against the domination of individualism, of arbitrariness."*

Bruno Taut, a key member of the German group of modernist architects, had this to say in his *Frühlicht* of 1920 [10]:

*"Oh, our concepts: space, home, style! Ugh, how these concepts stink! Destroy them, put an end to them! Let nothing remain! Chase away their schools, let the professorial wigs fly, we'll play catch with them. Blast, blast! Let the dusty, matted, gummed-up world of concepts, ideologies and systems feel our cold north wind! Death to the concept-lice! Death to everything stuffy! Death to everything called title, dignity, authority! Down with everything serious!"*

These extracts give an indication of the rage against traditional styles in art and architecture prevalent at that time. They reveal the profound societal discontinuity that was to provide a breeding-ground for any ideology, mixing political as well as artistic memes, which promised radically new solutions to the problems facing mankind.

A biological virus remains infectious against the continuous development of antibodies by host organisms. The way it does this is to change its encapsulation so that it is no longer recognized by the host. This is said to be one of the mechanisms for the resistance of the HIV virus to therapy [20]. In exactly this fashion, modernism successfully changes the shell in which its memes are packaged. Modernist ideologues accomplish this switch with great dexterity: almost a sleight-of-hand. As soon as one of the encapsulations is identified, and it is realized that it does not lead to the promised benefit, the shell is changed to a new one. The central core — containing images that erase information and complexity from the environment — remains the same. We list seven encapsulations for modernist memes:

- (i) 'Progress and economic prosperity from technology';
- (ii) 'Freedom from class oppression through new design';

- (iii) 'Social equality and housing opportunities for all';
- (iv) 'Moral superiority from using honest materials that express the spirit of the age';
- (v) 'Improved health and hygiene through smooth surfaces';
- (vi) 'The mathematical principles of pure form';
- (vii) 'Cost benefits resulting from modular production'.

Today, the modernist style predominates in architectural practice, and is taught in our schools to the exclusion of most other styles; the above encapsulations are therefore presented and discussed at length as part of the standard architectural literature. It is not useful to repeat that material here. What is of immediate interest is that the seven slogans listed above are very successful at encapsulating modernist memes, thus helping their propagation. Our point is that, in the absence of either a scientific or sensory basis, modernist architecture justifies itself solely by its memetic encapsulations. For detailed criticisms of the unfounded claims of modernism and the weakness of the usual arguments trying to justify those encapsulations, see [1; 3; 23; 24].

An architectural meme that has become part of our meaning system is protected by its ENCAPSULATION. Attempting to revise the meme pulls at the entire meme complex, which is attached to the rest of the mind's associational network of concepts. Writings by modernism's proponents link the visual images representing the style to other, beneficial memes, so that questioning modernist design appears to question the technological, scientific, economic, and social progress of the twentieth century. This often triggers a strong emotional reaction that is reminiscent of religious intensity [25]. We suspect that certain memes such as these become encapsulated into our belief systems in places traditionally occupied by a religious credo. From the first author's experience, this impression is reinforced by the irrational manner in which contemporary architects engage in discussing the foundations of their discipline.

When confronted by criticism based on scientific reasoning, many architects react by interpreting everything in modernist terms: they base all arguments on what the modernist 'masters' said, as if that were some sort of revealed truth. This is indicative of religious fundamentalism. An automatic reliance on modernist dogma as part of one's basic belief system is consistent with a memetic infection; i.e., the justification for a belief is the infecting meme itself. It is pointless to argue against ideas and values that people accept unquestioningly, or have adopted in the struggle to better their lives [5]. The reason is that people are physically, viscerally, and emotionally attached to their beliefs, regardless of how they acquired them, and irrespective of their absolute validity. No-one wants to have to reach back and re-wire their brain into new habits of thinking, because such a process can be traumatic. It is far easier to hold onto one's ideas and values, and when challenged, the natural reaction is to defend them emotionally without thinking about their origin [5].

## 8 The two faces of encapsulation

ENCAPSULATION has also been used to discredit traditional architectural styles and throw them out of favor. The meme here is a negative association, which spreads independently of whether the accusation is true or not. This happened to the Beaux-Arts style, which was tainted by association with pre-World-War-I society in supposedly 'decadent' western Europe. The same is true of the Victorian and Edwardian styles in England [25]. The Classical style, after surviving for more than two millennia, was discredited because neoclassical buildings were erected during the Second World War in Germany, Italy, and in Stalinist Russia [25]. The absurdity of this argument does not however undo the remarkably effective use of ENCAPSULATION to further the agenda of modernist architects. As a result, there exists a violent resentment today against traditional styles; although no-one who feels that way can explain logically why that should be so.

Destructive ENCAPSULATION is well known in the political arena, where it is used for character assassination. In the world of art, the Iconoclastic movement declared figural representation to be unholy, despite the complete absence of any such restriction within Christianity. This happened around the 9th century, and led to the wanton destruction of religious paintings and mosaics before it was reversed. Early Christian icons dating to before the 11th century are as a consequence extremely hard to find. A brief resurgence of Iconoclasm occurred in Italy in the 15th century, instigated by the deranged monk Savonarola, which prompted the burning of several of Botticelli's paintings. History is unfortunately replete with examples in which individuals, groups of people, races, ideas, or artifacts are eliminated, after being branded by association within a destructive ENCAPSULATION.

One of the twentieth century's most successful memes is: "*Ornament is a crime*", coined by the Austrian architect Adolf Loos in 1908 [10]. This phrase is impossible to forget; it goes straight to one's memory whether one agrees with its message or not, thus ranking it with the most successful advertising jingles ever. Because of the NOVELTY criterion, the more outrageous social memes are often the most virulent [5]. This particular ENCAPSULATION identifies anyone who dares to enjoy architectural ornament with persons who by creating ornament supposedly become criminals, and "*inflict serious injury on people's health, on the national budget and hence on cultural evolution*" [10]. Infection by this meme continues to this day, since Loos is presented as a pioneer of the modernist movement, and his overly plain buildings (though built with the most expensive materials) feature in histories of European architecture. Something occurring outside established architecture may eventually prove far more damaging in the long term. For millennia, people have built modest structures such as pieces of wall, a raised flower bed, a veranda, or an addition to someone's house, etc. This vast "architecture without architects" is simple, functional, often ornamented, and made out of available materials. Some of mankind's most endearing artifacts are produced within this tradition. They possess an emotional appeal and mathematical coherence that is lost when such structures are replaced by rigid industrial objects trying to emulate a crystalline geometry. People infected with modernist memes are eager to erase their heritage, since it reminds them of the past. Because of inner fear and feelings of inadequacy, people are terrified to risk losing what they believe to be modernist progress. In many societies, it has actually become illegal to build anything that doesn't 'look'

modernist. Something wonderful and complex — a tradition of building small things to please one's emotions — is becoming extinct as a result of a memetic infection.

## 9 A complexity threshold

The rapid spread of modernism is reminiscent of the spread of biological and computer viruses. What links them is their reduced complexity overhead (i.e., the minimum structural complexity they have to maintain during transfer from site to site). By sacrificing the structural complexity needed for metabolism, viruses gain an unbeatable advantage over more complex, metabolizing life-forms that they infect [20]. There is a parallel here with modernist design as it competed with more complex architectural styles such as Art Nouveau and the Classical style. Any style that attempts to adapt itself to human physical and emotional satisfaction, as well as to local materials and climate, will necessarily exceed a certain complexity threshold. In neglecting those needs — indeed, in making it its explicit aim to ignore them — modernist architecture crossed the complexity threshold going downwards. This brought it an unprecedented advantage, but removed an essential quality that we associate with 'life'.

Although 'life' has not been rigorously defined as a concept, biological life consists of two components: metabolism, and replication [15; 21]. The apparatus for metabolism represents much of what we observe as biological structure in every organism. The machinery for replication, on the other hand, occupies only a limited portion of an organism's structure. A virus replicates its encoded genetic information without being able to metabolize. It is the simplest possible life form, and by this definition, it is not 'alive' in the sense that a more complex metabolizing organism is. In an analogous manner, modernist structures, though immensely successful at replicating in the built environment, do not possess the same degree of 'life' (measured in terms of organized complexity) as do more traditional architectural styles that adapt to human use and emotional needs.

There is a debate going on in evolutionary biology as to whether viruses developed before, concurrently, or after metabolizing life forms [20; 21]. The third option argues that parasitic replicators have to have a population of more complex organisms to parasitize before evolving. A probable scenario for this third option is that some incomplete pieces from the replicating apparatus of an organism found it possible to lead an independent existence outside the metabolizing structure. Whatever the actual case, this third option is intriguing for its parallel to modernist architecture. With the above analogy, modernism could not have taken root before society became complex enough to support it. The intuitive perception of modernist buildings as 'alien' forms invading our cities (and minds) makes more sense in a society that is so morally and ideologically confused as to be in no position to stop the invader.

Evolution relies strongly on the organization of complexity. The metabolizing structure of all life forms exceeds a certain complexity threshold. Natural selection pushes many organisms to become more complex. It is true that some species reach a plateau when their structural complexity provides a reasonably good chance for survival and reproduction. Those that do this have no need to change as long as their environment or ecological niche remain stable. Nevertheless, the direction of evolution as defined by the progress from elementary life forms to humans is one of increasing complexity. A sudden *decrease* in organized complexity thus

appears as a catastrophic reversal akin to species extinction. Just as when viruses kill off a population of mammals, or when computer viruses erase a host of hard disks full of organized data, so the organized complexity of the built environment is decreased when Nineteenth-century buildings are replaced by modernist ones.

The low information content of minimalist design distinguishes it from other, more traditional styles of architecture, as well as from more recent stylistic trends. We want to clarify a misunderstanding in discussions of complexity in architecture. Biological forms are characterized by their extraordinary high degree of organized complexity. A high degree of organized complexity (visual as well as structural) is also found in the great buildings of the past such as mediaeval cathedrals and mosques, and in vernacular architectures. This property should be contrasted with a high degree of *disorganized* complexity that is seen in detailed, busy, but disorganized buildings such as postmodernist and deconstructivist structures. Disorganized complexity is also encoded in the visual cacophony of signs and materials in the suburban commercial strip, and the jumble of neon signs of the Las Vegas casinos. Our age appears incapable of organizing spontaneously-generated complexity, and we believe modernist memes contribute to this deficiency.

## 10 How architecture perpetuates modernist memes

By a remarkable confluence of historical events and circumstances, selection on the basis of images of emptiness has succeeded in displacing a variety of architectural traditions based on adaptation to human needs. Those who promote modernist structures agree that the style's lower organizational complexity is meant to deliberately contrast with the higher complexity of traditional architecture. Is it possible now to re-establish traditional adaptive design methods in practice? We don't believe that significant changes can come from within mainstream architecture, because it is itself a product of (and is totally dependent upon) modernist memes. Additional insight comes from seeing how three more criteria proposed by Heylighen: AUTHORITY, PUBLICITY, and CONFORMITY, contribute to the propagation and eventual institutionalization of memes [17; 18].

AUTHORITY from famous architects and their sponsors legitimizes design memes in people's minds. The backing from a recognized expert or institution boosts the acceptance of a particular idea [18]. After the Second World War, the United Nations built its headquarters in New York City as a validation of the modernist style. Several progressive governments reinforced this example by building new capital cities in a modernist style: India (Chandigarh); Brazil (Brasilia); Bangladesh (Dacca); and Australia (the post-war buildings in Canberra). The U.S. Government adopted modernism for its international trade missions and exposition spaces, projecting images of prosperity from a superpower, while corporations competed to outdo each other in occupying modernist headquarters. In our times, the administrative buildings of the European Community in Brussels embody modernist memes. (People conveniently forgot that modernism was the official architecture of Fascist Italy).

Setting up a cycle of positive feedback, acceptance of modernist memes by governments and organizations elevated their architects to a position of AUTHORITY. The 1932 exhibition on Modernist Architecture at the Museum of Modern Art in New York was a highly influential

event, using the museum's AUTHORITY to promote the so-called 'International Style'. (After New York, the exhibition travelled for seven years around the United States [9]). Two former directors of the German Bauhaus school were subsequently made heads of Architecture schools in the United States when they emigrated from Europe. Those architects then used their positions to promote modernist memes through their teaching, and by praising each other in the media. Their positions of AUTHORITY also guaranteed them more commissions to erect modernist buildings. The public rarely feels confident enough to challenge the AUTHORITY of individuals presented as the world's experts on the topic, even if what they say runs contrary to people's basic feelings and intuitions.

Professors at prestigious universities such as Sigfried Giedion and Nikolaus Pevsner — the first enormously influential as the Secretary of *CIAM (Congrès International d'Architecture Moderne)* — wrote scholarly "histories" of architecture that twisted facts to promote an ideology [25]. Modernism was falsely presented as the inevitable end result of the continuous evolution of historical architectures, instead of the radical negation of traditional styles that it represents. By claiming that modernism is not a style, and thus not subject to stylistic competition, they implied for it an AUTHORITY above and beyond architecture. Styles that modernism competed with and displaced (e.g., Neoclassical; Edwardian) were either dismissed as irrelevant, morally reprehensible, and were ignored altogether, or they were misleadingly appropriated as ancestors of modernism (e.g., Art Nouveau; Expressionism). An invented architectural history thus endowed modernism with false historical and moral AUTHORITY. Those treatises, along with others bearing the same misleading message, became the standard textbooks for more than one generation of architecture students.

An essential feature of evolution is that complex systems build upon existing complex systems: each new development adds something to what already works. New layers of functionality develop on top of older structures, without altering them radically. This summarizes both the advantages and disadvantages of cumulative design by selection [14]. We can trace evolutionary ancestry by looking for features in common with less developed organisms on the evolutionary scale; some of which survive in an inactive or useless form. For an architectural example, the Classical style retains features of its ancestral wooden construction, although they make no structural sense when building with stone. Nineteenth-century styles retained much of what had developed up until then. As modernist architecture was intent on destroying and replacing all past and existing styles, however, it cannot be termed an evolution of those styles.

PUBLICITY is the effort to spread an idea; often an ideology includes explicit injunctions that believers should engage in propaganda [18]. In architecture that is taken care of by a wealth of picture-filled books and architectural magazines, films, television documentaries, and the press; all of which promote modernist memes. These offer a platform from which often confused ideas are endowed with visual legitimacy. The 1932 'International Style' exhibition was conceived as a publicity campaign for modernist architecture, and its catalogue as a propaganda tool for disseminating the new style in the United States [9]. Modernist architectural memes spread through advertising techniques coupled with proselytizing in architecture schools. Since its inception in 1979, the Pritzker architecture prize has been awarded to architects who best embody the latest trend in design; such prestige and accompanying PUBLICITY in turn helps to perpetuate those fashion trends. The same is true for numerous other architectural prizes of lesser

prestige. Those prize-winning built examples are publicized by the media, and influence the design of new buildings.

CONFORMITY guarantees that newcomers into a group will be infected by an accepted meme, even though it rejects sound knowledge and contradicts established beliefs [17]. CONFORMITY pressure establishes and maintains an invariant belief over a group of people [17; 18]. Peer pressure from the architectural community maintains approved architectural images, with the threat of ostracism for apostates [25]. Many cases are known of ridicule heaped upon architects who stray from the official design style. Architectural magazines tend to publish only articles featuring buildings that maintain the *status quo*. Architecture students are infected with modernist memes by their teachers, and are under pressure to conform to the accepted style. The teaching of architecture has changed since most architecture schools adopted the Bauhaus concepts, so that today design is almost entirely image-driven. It is very difficult even to discuss adaptive design such as Alexandrine patterns [1; 24]. Academic architects still invoke the tired 'modern versus traditional' argument, which ultimately thwarts creative endeavor.

## 11 Modernism has become an institution

This paper described how modernist memes spread in society, to the point where they displaced most other architectural styles. We now wish to discuss the opposite process: how to remove modernist memes from our society. A group of memes achieves ultimate success by becoming institutionalized. The rigidity of institutionalized memes then makes it extremely tough to get rid of them. This section explains the great difficulty in displacing modernist memes from today's architectural establishment. Nevertheless, we are optimistic that people (most likely users rather than architects) will begin to consider alternatives once they see the connection between modernism and memetic infection.

An institutional system will take actions to protect its political base, which results in a conceptual bias. Decisions are made as to which information is relevant to it (i.e., relevant to architecture as defined by those in power), whereas other conceptual models will be ignored. The political agenda favors specific issues, deciding as well how dialogue with competing issues is to be addressed, if at all [13]. The same rules apply to any other meme group that has become institutionalized, and are not specific to either architecture, or design [13]. The institution defines both the importance and the interpretation of concepts, so that it controls whether a particular idea will be discussed, and whether any action will be taken. Not surprisingly, arguments and actions that do not fit into the conceptual framework of those in power are not going to be pursued.

This analysis is confirmed by how architecture has progressed since adopting modernism. The institutionalization of modernism in our society acts as a filter for innovation within architectural design. Despite well-publicized reactions against the empty forms of early modernism, most of the original visual memes have been retained in post-modern design. The overall forms may vary, yet the basic 'look' is still familiar. Architecture today remains at its core non-adaptive to human needs; that basic aspect certainly has not changed. Neither has the proscription against detailed ornamentation and color that the early modernists imposed. Materials tend, on the

whole, to be those preferred by the early modernists because of their universality, and even when traditional materials are used, they are used in ways so as to mimic "pure" modernist surfaces. Because every institutional structure acts as a memetic filter, innovative concepts may be able to evolve only outside it. In the case of architecture, the evolution of design that is adaptive to human needs is taking place mostly outside the establishment: either spurred by architects who have been evicted from the establishment, or by other professionals who have discovered that the existing complex is too rigid to deal with societal problems. The establishment is reacting in a predictable manner by ignoring innovations. Instead, it underlines its absolute control by allowing limited debate within certain boundaries. The debate is very tightly controlled, however, and is never allowed to endanger the institutional basis. Thus, the declaration of the postmodernists that 'modernism is dead' should be interpreted for what it really is: a smoke-screen meant to protect the architectural establishment from any serious attack.

The proponents of modernist design sought to eliminate competing styles by employing two tactics: aggressive attack, and ridicule. The switch from one to the other marks the point when modernism became an institution. In the years up to the 1950s when modernism was peripheral and was trying to gain the upper hand, destructive encapsulation was used with great effect. By the time this tactic succeeded, the modernist establishment found it more appropriate to express its strength by ridiculing its competition. The word 'pastiche' was henceforth used to make fun of any architects who tried to incorporate traditional elements into their designs. 'Pastiche' is the artistic equivalent of 'plagiarism' in the sciences, and implies that such an architect is not being original. In fact, those who follow the establishment typology are doing exactly the same thing. Developments in architectural design touch upon the commercial benefits of the modernist style in comparison to other factors. Modernist designs, because of their simplicity and context independence, are designs that can be produced industrially for a reasonable price. The institutionalization of modernist architecture has taken place in the entire construction industry, driven by the rise of industrial construction. This institution is separate, and more powerful than architectural education, institutes, and magazines. Some authors have identified this as the dominant factor for modernism's success [2]. In accepting a Faustian bargain, architects provided ever plainer designs that could be built more cheaply, until by now the threshold is so low that it is extremely difficult to raise. Any change that threatens to increase construction costs while lowering productivity in the construction industry is going to be fought by a massive establishment.

## 12 Conclusion

The idea of design as a Darwinian process that relies on selection has interesting ramifications for architecture as a whole. This explanation of how design emerges in the human mind reveals a split between design methods based on stereotyped images, and those based on adaptation to human needs. Both architectural and popular literatures come back to the theme that a majority of twentieth-century buildings provide neither the physical nor the emotional comfort for their users that older buildings — which are built in a freer, more adaptive style — almost invariably do. Nevertheless, despite such strong criticisms, certain visual styles continue to dominate construction and design practice today. An answer to why this is so comes from visual memes: self-sustaining conceptual entities that become encapsulated in human memory. Originally introduced in discussions of evolutionary biology, memes serve well to explain why architectural

fashions survive and propagate. In particular, memes explain why the modernist style has achieved such remarkable success in displacing traditional architectural styles.

## References

- [1] Alexander, C., Ishikawa, S., Silverstein, M., Jacobson, M., Fiksdahl-King, I. and Angel, S. (1977) *A Pattern Language*, Oxford University Press, New York.
- [2] Benedikt, Michael (1999) "Less for Less Yet", *Harvard Design Magazine Winter/Spring 99*, 10-14.
- [3] Blake, Peter (1974) *Form Follows Fiasco*, Little, Brown and Co., Boston.
- [4] Broadbent, Geoffrey (1990) *Emerging Concepts in Urban Space Design*, Van Nostrand Reinhold, London.
- [5] Brodie, Richard (1996) *Virus of the Mind*, Integral Press, Seattle, Washington.
- [6] Calvin, William H. (1987) "The Brain as a Darwin Machine", *Nature* **330**:33-34.
- [7] Calvin, William H. (1990) *The Cerebral Symphony*, Bantam Books, New York.
- [8] Calvin, William H. (1997) "The Six Essentials? Minimal Requirements for the Darwinian Bootstrapping of Quality", *Journal of Memetics: Evolutionary Models of Information Transmission* **1**. [http://jom-emit.cfpm.org/1997/vol1/calvin\\_wh.html](http://jom-emit.cfpm.org/1997/vol1/calvin_wh.html)
- [9] Colomina, Beatriz (1994) *Privacy and Publicity: Modern Architecture as Mass Media*, MIT Press, Cambridge, Massachusetts.
- [10] Conrads, Ulrich (1964) *Programs and Manifestoes on 20th-century Architecture*, MIT Press, Cambridge, Massachusetts.
- [11] Dawkins, Richard (1989) *The Selfish Gene*, (New Edition) Oxford University Press, Oxford.
- [12] Dawkins, Richard (1993) "Viruses of the Mind", in: *Dennett and His Critics*, Edited by: B. Dahlbom, Oxford, Blackwell, 13-27.
- [13] de Jong, Martin (1999) "Survival of the Institutionally Fittest Concepts", *Journal of Memetics: Evolutionary Models of Information Transmission* **3**, [http://jom-emit.cfpm.org/1999/vol3/de\\_jong\\_m.html](http://jom-emit.cfpm.org/1999/vol3/de_jong_m.html)
- [14] Dennett, Daniel C. (1995) *Darwin's Dangerous Idea*, Simon & Schuster, New York.
- [15] Dyson, Freeman (1999) *Origins of Life*, (Revised Edition) Cambridge University Press, Cambridge.

- [16] Edelman, Gerald M. and Tononi, Giulio (2000) *A Universe of Consciousness*, Basic Books, New York.
- [17] Heylighen, Francis (1993) "Selection Criteria for the Evolution of Knowledge", in: *Proceedings of the 13th International Congress on Cybernetics*, Namur, Belgium, Association International de Cybernétique, 524-528.
- [18] Heylighen, Francis (1997) "Objective, Subjective and Intersubjective Selectors of Knowledge", *Evolution and Cognition* **3**:63-67.
- [19] Kunstler, James Howard (1993) *The Geography of Nowhere*, Touchstone, New York.
- [20] Levine, Arnold J. (1992) *Viruses*, Scientific American Library, New York.
- [21] Maynard-Smith, John and Szathmáry, Eörs (1999) *The Origins of Life*, Oxford University Press, Oxford.
- [22] Mikiten, Terry M., Salingaros, Nikos A. and Yu, Hing-Sing (2000) "Pavements as Embodiments of Meaning for a Fractal Mind", *Nexus Network Journal* **2**:63-74.
- [23] Salingaros, Nikos A. (1995) "The Laws of Architecture from a Physicist's Perspective", *Physics Essays* **8**:638-643.
- [24] Salingaros, Nikos A. (2000) "The Structure of Pattern Languages", *Architectural Research Quarterly* **4**:149-161.
- [25] Watkin, David (2001) *Morality and Architecture Revisited*, University of Chicago Press, Chicago.

## **GERMS AND IDEAS**

By  
**André Siegfried**

### **The spreading of germs and ideas**

There is a striking parallel between the spreading of germs and the spreading of ideas or propaganda. On the one hand we are dealing with a virus which can be transported and transmitted under certain conditions which favor or limit its transportation or transmission: on the other hand with ideas, religions, and doctrines, which can be described as germs, benevolent or malevolent, according to the point of view one takes up. These germs can either remain at their source and be sterile, or emerge in the spreading of infection. The vocabulary which normally comes to mind is that of medicine. Conditions of infection, besides, will also be the same as in the domain of health. For diffusion to take place we must necessarily have a germ, a carrier, and receptive surroundings. It is then, as in the case of epidemics, a matter of contacts, made easier by the techniques of communication or delayed by obstacles arising from custom or from administration. Sociology and biology are curiously akin in this.

### **Conditions necessary for the initial germ to spread**

The initial germ cannot travel on its own; it needs a carrier, which may be a man, a journal, a book, a pamphlet, not to mention the radio or the cinema or the television screen.

The human being is in the circumstances the most natural transmitting agent. A particular man, on moving from one place to another, will take with him a doctrine, a religion, a germ of opposition or revolt. It may be that he feels conviction about it, like a missionary, an apostle, or a propagandist, but it might also be the case that such an agent remains unconscious of the contagion he spreads, carrying a germ of which he does not suspect either the harmfulness or even the existence: just as there are carriers of non-apparent infections, who know nothing of the virus they harbor, which is undiscovered even by doctors. These are the most dangerous. The carrier who is admitted to be such will preach, harangue, teach, discuss, insinuate, openly sowing the seed along his path, which means that he can be stopped or refuted; but with the unconscious or camouflaged carrier diffusion will be infinitely more insidious: it may be an army soldier in a distant garrison or in the battle-field, a peddler on his rounds, a laborer recruited for some distant scheme, or even the valet of some millionaire touring round the world. The fantasy, the unpredictability of human contacts is without limit, surpassing the scope of even the wildest imagination. We have only to consider the pollen spread to the uttermost ends of the ancient world by those bees, the Tartar conquerors; the fecundation which those kill-joys, the nomads, periodically brought to the owners of solid hereditary estates; the incessant temptations of neighboring idolatries for the faithful of the true God; or the unforeseen spiritual transports brought about by the Parthians, the Mongols, and the Arabs. The world is infinitely more permeable than one could believe, and in order to be so it has not had to wait for the telephone, the telegraph or the aeroplane to conquer distance and time for us.

We have said that infection may take place through the reading or the hearing of something from a distant source, or through the sight of some image, but, all the same, human intervention will have been necessary: the germ may be laid in a book, in a screen production, or in a purely formal rite, yet a human intermediary will have been necessary; the source remains inevitably human, for one cannot imagine an idea without a human reservoir, comparable to the germ reservoir in the preservation and transmission of diseases. As far as the virus reservoirs in animals are concerned, however, the comparison stops short, for one can hardly imagine the monkey or the rat playing the part in mind communication that they play in the case of the yellow fever or the plague, even though they be the library rats of the Fable.

## Routes followed in the diffusion of ideas

Here we again find the routes which we have already come across in the spreading of germs, which are naturally a subdivision of the large class of world communications in general. Thus we have to remember the caravans, those ancient instruments of trade, pedlary, pilgrimages, sea itineraries, immigration routes, and missionary routes, which are often indistinguishable from routes of conquest - a list that goes back over the ages. If we are dealing with modern routes, conditions of contact remain the same, but transport techniques have supplied the railway, the motorway, and the airway. The nature of communications does not alter, but their intensity increases, the route becoming the instrument *par excellence* of penetration, now as strong as a battering ram able to knock down any obstacle, now supple., subtle, and capable of insinuating itself everywhere: the work of Gothard, by introducing a foreign labor force into a swiss canton which was still living in the Middle Ages, changed the political and social balance. In his novel *Les hommes de la route*, André Chamson has described a similar revolution brought to the Cévennes in an offensive on the part of the Highways Department. In building roads, in laying down railways, in setting up strings of naval bases across the seas, and in forming a network of substructures for the needs of air travel, pioneers bring with them not only their techniques but also their ideas, and then, after them, other men use these paths to convey products of all kinds, including those of the mind. The part that chance then plays in these contacts surpasses the imagination and one must have recourse to the demoniac fantasy of the association of ideas to find any image for it. The insinuation of idea germs can come about through the arrival of tourists, through a detachment of foreign workers, through an army of occupation, through the enrollment of foreign students at a university, through then introduction to non-immunized surroundings of a single individual, or sometimes of a journal or a book, or through the unexpected conjunction of something read with a state of mind which happens to be receptive at that particular moment. My grandfather, who was a notary in the chief town of one of the districts of Ardèche, while sailing along the Rhône towards Lyons to see a performance of *Rachel*, was engaged during the journey in reading a Bible that a friend had put into his hands --- it was during the time of the Revival; he did not see *Rachel*, came back to his town, and after conversion, became a minister. the example is typical of many. Infections, epidemics, are born in this way. ``There are forms of madness which can be caught in the same way as infectious diseases," says the moralist. The types of human carrier are innumerable - each of us being one, on occasion - but certain categories of men are so by profession, or even simply by temperament: As examples we can quote the missionary, the traditional peddler, the political agitator, the leader feared by his employer, the canvasser in an election, the publicity agent, and the American [travelling] salesman: from the sociological point of view the possibly endless list would be most

interesting to draw up. In any case, the "evil spirit" introduced by the purest chance into given surroundings, is sure to breed trouble and disorder, and one can easily understand the employer, whoever he may be, shielding himself from it *as from the plague*.

## Environmental conditions

The environment, however, must be suitable. For diffusion to take place the environment must necessarily be receptive, that is, suitable for the development of the germ, without which conditions it will remain sterile, or else there may only be a localized and temporary infection. It may also happen that the human carrier disappears of his own accord, or is simply suppressed as being harmful: this radical suppression has the same result as it has on infections. In the geographical studies of public opinion that I made on the borders of the Central Plateau in contact with the Mediterranean world, I found many examples of similar infection reaching upwards and then stopping and becoming sterile. Extreme leftist ideas thrive and multiply on the coast in the form of radicalism, socialism, or communism, according to the times, and they have a tendency to travel up towards the mountains along the valleys, just as Mediterranean qualities work their way up towards the plateau along the routes most exposed to the sun. Thus between Hérault and Aveyron, socialist propaganda between the two wars raised a small extremist center at Séverac-le-Château in Aveyron, among the railway workers; but the atmosphere was evidently not propitious, for its virulence remained limited and localized. If extremism thrives in low-lying country, altitude is not favorable to it. Before 1914, as I remember, I observed an analogous transmission of germs in Hérault, Gard and Lozère, with no consequences. In the distant time of the Dreyfus Affair, when a new generation in forward-looking circles was substituting the socialism of Jaurès and de Guesde for the old radicalism of Clemenceau and Pelletan, there were students in the Faculty of Medicine of Montpellier University who were responding with enthusiasm to this propaganda. One of them, called Phalippoux (I still remember his name) going along the valleys of the Basses Cévennes, had carried the revolutionary good word, which was then Jaurèsism, as far as the parishes of the district of Florac, where the protest element declared itself left-wing in virtue of a tradition of resistance stemming from the Desert and persecution. Many of the good Huguenot electors of the time voted therefore "to the left" and "always to the left" --- as a young greenhorn who had only just attained his majority said to me: "I am a Marxist, and I vote on principle for the most advanced side." The agitator from Montpellier had thus created small revolutionary centers at Pont-de-Monvert and Barre-des-Cévennes. But the altitude, physical as well as moral, no doubt was unfavorable, and the movement never developed; it is yet another example of infection setting up a small center and then dying a natural death. We could quote numberless examples, in the sphere of religion as of politics: the rare and sporadic Christian conquests in Asia suffer from these limitations: they form small centers, active and not without vitality, but discontinuous and unable to spread, and becoming endemic only through a kind of adaptation to their surroundings, as is the case with Portuguese Catholicism on the Malabar Coast. Sometimes all that is needed is that the merest chance, or else the founding of a new transport route, should carry far afield some bearer of religious conviction: this is how some negro employees travelling in the Southern Pacific Pullman between New Orleans and Los Angeles, once brought to South California the germ of some sect or other which held that Jesus and the Apostles were black; one can imagine that the infection would not spread very far beyond the railroad.

## **Defense measures and their degree of efficacy**

In order to destroy and sterilize a germ, or simply to circumscribe its diffusion, one can, as in the struggle against disease, proceed in different ways.

The most decisive action is to destroy the germ at its source. But this is often difficult. Either one does not know the germ or, if one knows it, one does not know exactly where to find it. It may be that its possibilities of diffusion have not been properly estimated. It may also be that the germ has multiplied in such a fashion that one is defenseless, tempted to say with the combatant who is completely surrounded: "They are too many!" We can well understand the reaction of Simon de Montfort at Béziers in his extermination of the Cathars: "Kill them all, God will recognize his own!" The Albigensian germ was in fact completely destroyed, and since then it has only existed in the memory.

Another defense, which is the most usual, consists in suppressing the carrier, or at least in preventing him from penetrating the territory that one wishes to preserve. Thus all the agents of diffusion of this or that doctrine will be systematically destroyed. At all events their entry will be prevented by an impassable barrier at the frontier. That is the traditional Russian system, favored by the Czars long before it was used, as today, by the Soviets. It is also the American system of the immigration quota, which tends to limit the entry of immigrants of Latin or Slav origin, whose influence, it is assumed, might compromise the orthodox Anglo-Saxon tradition. After the First World War America thought that she could in this way keep out a European infection that she considered insidiously Communistic. McCarthyism, after the Second World War, was but a temporary paroxysm of this defensive obsession. If however the infecting agent somehow succeeds in getting in (and how can one really prevent this?) every effort will be made to watch him, to limit his activities and to render him harmless. There may be some doubt as to his degree of harmfulness, and, in this case, he will be subjected to some sort of quarantine. I have been assured that, in the Province of Quebec, the orthodoxy of whose Catholicism is well known, a Canadian student returning from any French university is placed under observation for a certain period, to make sure that no infectious disease develops in him.

The procedure which is most prudent is naturally to take preventive measures: censorship of the press or of books, examination of luggage to make sure that it contains no subversive publication, police supervision of suspects, religious persecution, to say nothing of these so-called "psychological" operations of re-education, de-fanaticizing and re-adaptation which are the shame of our century. The Russians and the Chinese are past masters at this type of treatment, but it certainly seems to have been used occasionally in the American army during the Korean war, and it would be most reassuring to know that it is not practiced in North Africa. Hitler caused this technique to progress in a truly horrifying way and since then the demoniac progress of neuro-surgery has placed in the hands of experimenters in the matter instruments capable of working miracles in the domain of evil.

One wonders, however, how far such procedures are effective against that winged germ, often invisible, which is the mind. If one "kills them all," as in the case of Albigensian Manicheism, the enterprise is successful, but if as much as one remains propagation starts again, as is the case with the majority of persecutions: whether it is a case of the Protestants after the revocation of

the Edict of Nantes, or the Jews throughout the ages, or of the Armenians assassinated *en masse* by the Turks, the operation has proved itself useless, for there are still Protestants, Jews and Armenians. The sterilization of the surroundings by the suppression of all contact with the exterior has produced the expected effect in Russia: the state succeeds in preventing the Soviet citizen from knowing what is going on outside, and besides, if any outside propaganda evoked the slightest response this would soon be suppressed by a police force which could, in this context, be described as ``sanitary.'' The iron curtain and the bamboo curtain are, from this point of view, as effective a protection as the ancient wall of China, which protected the refined civilization of the Yangtze Kiang against the raids of the Tartar hordes.

So far the comparison between the epidemics of the body and those of the mind has proved fruitful. It remains so as long as the question is one of immunization or vaccination. The laws of nature have given the organism a mechanism for self-defense, corresponding to a fundamental instinct for restoring the balance. That is the significance of immunity, of which vaccination is but a derivative. In the defense measures that human societies adopt against the attacks of propaganda or the penetration of new ideas, immunization and vaccination likewise exist. The virulence of a given doctrine is attenuated by its adapting itself to its surroundings or simply through habit. It can even happen that the injection of a reduced dose can bring with it immunization against a more active form. I seem to have observed several times in Latin America, especially in Argentina, that a particular form of demagogery is the equivalent of a social vaccine against communism. The same thing can be observed in the United States with respect to Irish demagogery, which never goes as far as revolution. A given reform, subtly inspired by the menace itself, will often prevent the disease from declaring itself or at least from spreading when a local infection appears. A healthy society has its own defenses against the infections which might endanger the integrity of its personality. On the other hand an unhealthy society will catch all the illnesses that are going, and will react to all the infections. Rome, in its decline, was not able to defend itself against Christianity. Bismarck said of Russia that she had no business in Europe, and would just catch all its diseases. Freudian psychology spread easily over the Protestant countries, particularly the United States and Switzerland, while Catholic countries put up an effective resistance.

Certain doctrines and certain religions spread like epidemics, following the same lines of least resistance, coming up against the same obstacles, first striking out - depending on the transport available - towards the main centers of communication and distribution, from which they then spread in all directions. Conditions are the same as those we analyzed above with regard to the great pandemics of the nineteenth century, halting places and ports of call being the first to be affected, remote places only being reached after some delay. This explains why certain isolated districts hidden in the mountains or forests escape infection for so long, acting as strongholds of resistance to the infection of progress, and as the last refuge of conservatism.

A typical example of a spiritual epidemic is to be found in the spread of Christianity over the ancient world, which was exceptionally permeable from a spiritual point of view in Roman times. What do we find if we describe this in biological terms? The initial germ is at Jerusalem, but it might have given rise to a merely local infection, limited to the Jewish society in which Christ carried out his apostleship: the apostles, good Jews, who no doubt meant to remain within their own religion, would probably not have spread their faith to any great distance. But there

were proselytes among the first Christians, Hellenized Jews or Judaized Greeks, and in their ranks was Stephen, the first martyr. After he had been stoned to death by the fanatics of orthodoxy, his closest disciples, fleeing from Jerusalem, thus fled from the capital to the coasts of Asia Minor, where Jews and Greeks met in the synagogues. Here they were in surroundings which lent themselves admirably to the diffusion of the doctrine which Jerusalem refused to accept, and so, by degrees, Christian centers were constituted in Antioch, Tarsus, Ephesus, Thessalonika and Corinth. The carriers, in this instance, were the disciples of Stephen, the converted Judaeo-Greeks, the missionary apostles in whose front rank we find Saint Paul, the carrier *par excellence* through his immense missionary activity. Where the atmosphere was favorable, churches sprang up; where it was not, the Christian seed remained sterile: this was the case at Athens, where the Apostle preached as if in the wilderness. But in other places it was successful, and in the Mediterranean where everything tended towards Rome, it was inevitable that sooner or later the capital of the world should be affected. From Rome the diffusion became generalized and, everything considered, took on for the people of the time, the nature and extent of a pandemic.

The itineraries of Saint Paul are, from this point of view, singularly revealing. The apostle chose the big towns placed along the most frequented lines of communication, because there he found the people less fixed and established, and so more accessible to his preaching; and also because, no doubt, as a tentmaker he found better possibilities there of earning his living. The pattern of roads, the network of sailing routes, obviously also counted for something, for there is a curious mixture of settled plans and irrational impulses in the missionary method. It was natural that he should finally end up in Rome, and, had he not perished under Nero's persecution, doubtless he would have gone on as far as Spain and Gaul. The pressure of East upon West which, in the second and third centuries reached its maximum intensity, meant that all theological developments taking place in Western Asia had an immediate or almost immediate repercussion in Rome, where there were important Asiatic colonies. Maurice Goguel, the great historian of early Christianity, to whom I am indebted for these thoughts on the geography of religion, believed that an exciting study could be made of the geography of primitive Christianity and the manner in which early Catholicism was built up by a sort of concentration, fusing types of Christianity which at first had differed considerably. The first Christians traveled about a great deal, and their journeys and highly developed practice of hospitality did much to unify them. Christianity thus moved westwards, the center of gravity moving from Jerusalem to Antioch, then to Ephesus, and finally to Rome after a fruitless attempt on the part of Corinth to compete with the capital.

Another example of the geographical extension of a religion can be found at more or less the same time in the spread of the mysteries of Mithras, as appears from the classical research of Franz Cumont. The main Mithraic monuments have been found on the frontiers of the Empire, strung out along, and protected by the *Limes*. It is clear that the Roman legions were the means of transmission here; it is a case, and a very curious one, of an occupational infection.

We have made use of the same vocabulary to speak of the spread of diseases and of ideas and propaganda: virus, germ, source; carrier, soil, surroundings; contact, contagion, infection, contamination, endemic, epidemic, pandemic; prevention, inoculation, sterilization, immunization, vaccination, quarantine. Surely we have more than a superficial coincidence here:

both in the domain of biology and in the world of ideas, certain reactions are shown which are common to all living beings.

From *Germs and Ideas: Routes of Epidemics and Ideologies* (1965; translation of *Itinéraires de Contagions: Epidémies et idéologies*, 1960, American title *Routes of Contagion*): Part Four, "The Spreading of Ideas and Propaganda," Chapter 7, "Conditions under which ideas spread, and factors determining the choice of route."

# **MEMECOSYSTEMS: ARE ANIMAL MINDS SUITABLE HABITATS FOR MEMES?**

**By**

**Brent Silby**

Department of Philosophy  
University of Canterbury  
New Zealand

## **Abstract**

Milk-bottle opening behavior in a species of bird known as the British tit has been put forward as an example of a meme in a non-human animal. The existence of this type of case has led some thinkers to believe that non-human minds can acquire memes. I believe that the British tit's behavior has been misinterpreted as memetic. In this ePaper, I argue that milk-bottle opening in the British tit can be explained by appealing to its innate behavioral repertoire. I then suggest that the question of memes in animal minds should be considered on a case by case basis.

## **1. Introduction**

In their paper "Do Animals Have Memes?" (1999) Simon Reader and Kevin Laland suggest that animal minds are suitable habitats for memes. Their ideas come about after a critical examination of points that Susan Blackmore raises in her book *The Meme Machine* (1999). Blackmore argues that humans are the only animals with minds capable of supporting memes. Reader and Laland disagree and attempt to find some examples of memetic transfer in non-human animals.

In this ePaper, I will examine Reader and Laland's claims. I will focus mainly on their suggestion that the British tit's milk-bottle opening behavior is a meme. My objective will be to show that Reader and Laland's claim is wrong, and that the British tit's behavior is not a meme. I will start by outlining Reader and Laland's argument and will show that their claims rely on an incomplete definition of what a meme is. I will then offer a firm definition for the meme, after which I will provide an alternative explanation for the milk-bottle opening behavior exhibited by the British tit. My next point will be that the question of memes in animals should be answered on a case by case basis. My conclusion will be that examples such as milk-bottle opening in the British tit do not provide evidence of memes in animals.

## **2. Memes in the Minds of Animals**

According to Susan Blackmore (1999), imitation is the primary mechanism by which a meme replicates and finds its way into a new mind. If she is right, it would seem that memes are exclusive to animals that have the capacity to imitate behavior. So far this ability seems to be found only in humans and some species of birds, members of which can mimic certain sounds and calls.

Reader and Laland agree that there is limited evidence of imitation in non-human animals, but argue that imitation should not be the defining feature of a meme (Reader & Laland 1999). They suggest that limiting memetic transfer to acts of imitation is restrictive and cuts out a significant range of behavioral transmission. For Reader and Laland, the psychological process underlying the transmission of information is not a determining feature of memetic replication. Instead, they argue that *transmission fidelity* is the key feature of memetic replication. This is to say that regardless of the means by which a behavior is replicated, a memetic transfer can be said to have taken place if the resulting behavior is a *high quality copy* of the original behavior. Reader and Laland are right to make this point. If acts of imitation were the only way that memes could spread, then it would seem senseless to speak of memes entering human minds from written material or from musical notation. In these cases information that produces behavior is transferred with no observation of the original behavior -- it is not mimicked. Music can, of course, be imitated, but most of the time musical memes are encoded in a musical score and are assimilated to new minds when the score is read.

To support their claims, Reader and Laland point to evidence of behavioral transmission in non-humans animals that does not come about through imitation. The examples used are cases of *social learning* -- learning that occurs when behavior is *influenced* by observation of other animals (Reader & Laland 1999). Perhaps the most significant example for Reader and Laland is milk-bottle opening in British birds. In some areas of Britain, a species of bird known as the British tit has acquired the behavior of opening milk bottles to get cream. The tit finds milk bottles at house doors and pecks away at the foil bottle top in order to get access to the cream that sits on top of the milk. When other tits observe this behavior, they try it out themselves and soon learn that if they open milk bottles, they are rewarded. It is important to note that the tits are not *imitating* each other. They are simply being attracted to objects that other tits are pecking at. According to Reader and Laland, the spread of this behavior throughout the tit population in Britain shows that memetic transfer has taken place. Opening milk bottles is not innate to the tit -- it is learned -- and the behavior replicates with high fidelity showing that it is a meme. Since the British tit can have memes, Reader and Laland conclude that animals have memes.

### 3. Is this stretching the Meme too far?

To effectively assess Reader and Laland's claims we need to know exactly what a meme is. Without a firm definition, it is possible for theorists to make any claims about memetic transmission. Reader and Laland want to relax the criteria that past thinkers have placed on the meme, but I think they are relaxing the criteria too far. If we travel down the road that Reader and Laland are opening up, we will end up describing all conditioned behavior as memetic. On the surface, this might seem acceptable, but it is only acceptable if we have a very loose definition of a meme.

Why is such a loose definition unacceptable? The reason is that it makes the term 'meme' redundant. The best way to understand this is to consider the parallel situation in genetics. A gene is a packet of instructions encoded in DNA, which directs the development of cells. Now, if someone comes along and claims that all chemical processes are genetic, then the term gene would apply to everything that goes on in the body. Not only would we describe an animal's physical and mental characteristics as genetic, we would also describe the chemical process

underlying digestion as genetic. But this is not what genetics is about. Genetics attempts to explain things at a higher level than common chemical process. The same is true of memetics. Memes are used to describe high level behavior that is not an innate part of an organism's behavioral repertoire.

In this section, I will offer a firm definition of the meme and will use this definition to show why the spread of milk-bottle opening in British birds is not memetic.

### 3.1 What is a Meme?

Memes are best thought of as sets of instructions that can be followed to produce behavior (see Silby 2000c). Instructions can be encoded in a number of formats, including:

- 1) Musical notation,
- 2) Written text,
- 3) Visible (or vocal) action,
- 4) Connectionist networks such as the neural structure of the brain.

A set of instructions that produces the behavior of, say, whistling the first 4 notes of Beethoven's fifth symphony (a well used example) can be encoded in any of these mediums. When the instructions are followed, the same behavior will result. When a mind encounters an instruction set that produces behavior (say a musical score), it can reproduce that behavior by creating an appropriate neural "program". To understand this, consider a similar situation in the world of robotics. Imagine that a robot is developed that contains a number of built in programs. These programs provide it with behavior essential for its survival. The robot's behavior might include walking, avoiding obstacles, grasping at objects, and the production of certain vocal sounds. Suppose that engineers also give the robot a program that gives it the ability to write small behavioral routines. In effect, this program gives the robot a means by which it can alter its own behavior by writing new programs. Suppose that a feature of this innate behavioral program is that it allows the robot to *observe* the behavior of other robots and write programs that produce the same behavior. In other words, the robot can *imitate* behavior. Now, the question is: what, exactly, are the programs that the robot writes for itself? They are not a part of the robot's innate behavior and they are produced primarily through imitation. They can be translated into different languages and written down on paper. They can also be transmitted to other robots who read the instructions or who imitate the behavior and write their own programs. These behavioral programs are what we are talking about when we speak of memes.

This is precisely the sort of process that goes on in humans. At some point in our history, biological evolution provided our ancestors with a capacity to imitate behavior. This meant that when humans observed the behavior of others, their brains would create the neural wiring needed to produce the same behavior. Such neural wiring patterns are lists of instructions, which can be translated into other mediums such as written language, outward behavior, or computer code. A list of instructions that produces behavior is the thing that spreads into the minds of others. A list of instructions that produces behavior is a meme.

### 3.2 Why do British Birds not have a milk-bottle opening meme?

Given the above definition of the meme, it would seem that British tits do not have an 'open milk-bottle' meme. This is because British tits have not acquired any new behavior. Blackmore (1999) suggests that the British tits already knew *how* to peck and it was simply a matter of one tit being attracted to another -- who happened to be sitting on a milk bottle -- and then carrying out its innate pecking behavior. I think it is also possible that a tit would be attracted to a milk bottle even if no other tit happened to be around. Most birds have an innate attraction to bright shiny objects, and milk-bottle-tops are usually made out of a silvery foil. Milk-bottle-tops must stand out like bright beacons for passing birds, who feel compelled to land and carry out their innate pecking behavior. Encountering cream beneath the silver foil would *reinforce* the behavior and make it more likely to occur in the future. However, because no *new* information has entered the bird's mind and no *new* behavior has been produced, milk-bottle opening in birds cannot be counted as a meme.

Reader and Laland believe they have an answer to this objection. Their answer draws a parallel between the British tit's milk-bottle opening behavior and the behavior of a human tennis player. For Reader and Laland, a tennis player acquires tennis playing memes by observing other tennis players. These memes carry information about the appropriate location to play, the objects a player uses, and how the player should move in a certain way to hit the ball. They go on to suggest that the tennis player does not learn any new motor behavior such as running, holding, waving arms, or hitting, because these are already a part of their behavioral repertoire. The tennis player is carrying out innate behavior in order to achieve a non-innate goal. Reader and Laland then state that "*Exactly the same logic applies to milk-bottle-top opening birds*" (Reader & Laland 1999). The tits are not learning how to peck at milk-bottle-tops, just as the tennis player is not learning how to run or swing her arms around. They are, however, learning to peck a *particular* object to get cream, and this is a behavior and goal that is not innately defined. Therefore, milk-bottle opening in the tit is a meme.

The problem with Reader and Laland's suggestion is that they are glossing over important differences between the two cases. It is true that tennis players are carrying out innate movements, and it is true that the British tit is carrying out innate behavior. But the difference is that the tennis player's movements are combined in a non-innate fashion, whereas the tit's behavior is exactly the same as it would be if it was landing and pecking at any object. We would not be surprised if we saw a tit land on an apple and start pecking at it. Nor would we be surprised if the tit started pecking at the ground. So what is so special about a tit pecking at a milk bottle? Pecking is simply something that the British tit does. Also, consider the different reasons behind the tennis player and tit's behavior. The tennis player behaves in a specific way in order to win the game, which is a *socially defined* goal. The goal of the tit is to feed, and regardless of what food it happens to find, feeding is an innate objective.

There are also important differences in how the behavior is learned. The tennis player learns through imitation and verbal instruction, which are methods of passing on information that the tennis player does not already possess. The tit's primary method of learning is through behavioral conditioning. When a desirable outcome occurs (getting the cream), the behavior of pecking milk-bottles is strengthened. But this is not a memetic transfer of information, because no *new*

behavioral information has been assimilated by the tit. No information (or instruction set) is being moved from one mind to another so no memes have been copied. At most, the case of the British tit is an example of innate behavior being *triggered* through observation and *reinforced* through reward.

### 3.3 Does this mean that animals don't have memes?

The quick answer to this question is, no! All I have shown is that the spread of milk-bottle opening in British birds is not an example of memetic transfer in animals. There may be a number of animal minds that can support memes, but I think caution is needed before we make any definite claims. Because of the diverse range of cognitive ability among species on Earth, we should consider the evidence on a case by case basis. The case of milk-bottle opening in the British tit is not a meme, but there are cases of bird behavior that definitely point to the existence of memes. For example, certain species of bird can imitate sounds ranging from the calls of other birds to the sound of a telephone ring. The production of such sounds is not innate to the bird, so we must conclude that some sort of memetic transfer takes place. Something in the bird produces the neural wiring required to enable it to reproduce the sounds it hears in the environment.

One way to check if memes are involved in animal behavior is to see if the behavior is spread through imitation. For example, if a dog (Fred) watches a person (Roger) jump through a hoop and then copies the behavior and jumps through the hoop himself, we could suppose that the *idea* of jumping through the hoop was passed from Roger to Fred. The hoop jumping meme would have copied itself into Fred's mind through the process of observation and imitation. I think it would be unlikely for this to happen. Dogs do not seem to have the capacity to observe and then copy behavior. Getting a dog to jump through a hoop requires a long period of behavioral conditioning. Furthermore, it is possible that coordinated jumping is actually innate to dog behavior, so we might not be correct in suggesting that jumping through a hoop is a meme.

There are some examples of animal behavior that look memetic. Some dogs can be trained to balance and walk on a moving ball. I find it difficult to believe that such behavior is innate, so it might be accurate to consider 'walking on a ball' behavior to be memetic. Of course, a dog cannot learn to balance on a ball by watching another dog. The only way that a dog can learn to walk on a ball is through extensive training and reinforcement. It seems that if walking on a ball is a meme, then it is a meme that has a very limited potential for replication. Since dogs cannot copy each other, the 'walk on a ball' meme can only replicate when the behavior is viewed by a human and then implanted into another dog's mind through a similar process of training.

Other possibilities of memes in animals include the choreographed walking of dressage horses, the transmission of directional information in honey bees, and food washing in Japanese macaques (as mentioned by Reader and Laland). The important point to note is that the *possibility* of memes in these animals is not enough to conclude that there *are* memes in these animals. Each case must be assessed independently. Perhaps the best way to test whether a behavior is innate or memetic is to isolate a member of the species in question. If, for example, a Japanese macaque was isolated from birth we could observe its behavior and find out if it has the compulsion to wash its food. The result of this test would settle the question for the Japanese macaque, but would leave the question of other animals open.

## **4. What have I shown?**

My goal has been to show that milk-bottle opening in the British tit is not a meme. To accomplish this goal, I offered a firm definition of the meme and worked with this definition to show that the spread of milk-bottle opening behavior in the tit does not exhibit the features of memetic transmission. According to the definition I offered, a meme is a collection of instructions that, when followed, give rise to behavior. These instructions can be encoded in a number of mediums including the neural pathways of the brain. The most important feature of memetic transmission is that the instructions transferred are new, and did not already exist in the mind of the receiver. In this way, we can say that the acquisition of a meme gives rise to a new, non-innate behavior. Since the British tit has an innate attraction to other tits, and since pecking at objects is innate to the tit, we must conclude that milk-bottle opening is not the result of new behavior and is therefore not memetic.

I cannot show that non-human animals do not have memes. In fact, I have offered some possibilities of memes in non-human animals. I have however, shown that milk-bottle opening in the British tit is not memetic, and that the existence of this behavior is not sufficient to conclude that animals have memes. Because of the diversity of cognitive ability among animals, I do not believe the question of memes in animals can be answered definitively by pointing to a few cases. Some animals may have memes while others do not. We must proceed cautiously and take a case by case approach. By doing this we will be able to construct a catalog of animals whose minds have memetic compatibility.

### **Memes were assimilated from the following sources:**

Blackmore, Susan. (1999). *The Meme Machine*. Oxford University Press.

Dawkins, Richard. (1976). *The Selfish Gene*. Oxford: Oxford University Press.

Dawkins, Richard. (1986). *The Blind Watchmaker*. London: Longmans.

Dennett, Daniel C. (1991). *Consciousness Explained*. Penguin Books, 1993.

Dennett, Daniel C. (1995). *Darwin's Dangerous Idea*. Penguin Books, 1996.

Dennet, Daniel C. (1999). "The Evolution of Culture" in *Edge* online journal,  
[<http://www.edge.org/documents/archive/edge52.html>](http://www.edge.org/documents/archive/edge52.html)

Kendal, Jeremy R. & Laland, Kevin N. (2000). "Mathematical Models for Memetics" in *Journal of Memetics - Evolutionary Models of Information Transmission*, Vol 4.  
[<http://www.cpm.mmu.ac.uk/jom-emit/2000/vol4/kendal\\_jr&laland\\_kn.html>](http://www.cpm.mmu.ac.uk/jom-emit/2000/vol4/kendal_jr&laland_kn.html)

Reader, Simon M. & Laland, Kevin N. (1999). "Do animals have memes?" in *Journal of Memetics - Evolutionary Models of Information Transmission*, Vol 4.  
[<http://www.cpm.mmu.ac.uk/jom-emit/1999/vol3/reader\\_sm&laland\\_kn.html>](http://www.cpm.mmu.ac.uk/jom-emit/1999/vol3/reader_sm&laland_kn.html)

Silby, Brent. (2000a). *Revealing the Language of Thought*.  
<<http://www.def-logic.com/articles/RevealLanguageOfThought.html>>

Silby, Brent. (2000b). "The Evolution of Technology: Exposing the Myth of Creative Design".  
<[http://www.def-logic.com/articles/evolution\\_of\\_technology.html](http://www.def-logic.com/articles/evolution_of_technology.html)>

Silby, Brent. (2000c). "What is a Meme?"  
<[http://www.def-logic.com/articles/what\\_is\\_a\\_meme.html](http://www.def-logic.com/articles/what_is_a_meme.html)>

# **WHAT IS A MEME?**

**By**

**Brent Silby**

Department of Philosophy, University of Canterbury  
Christchurch, New Zealand

Memetics is rapidly becoming a discipline in its own right. Many web-sites are being devoted to the study of memetics, and new e-papers are appearing every day. With this in mind, I want to step back and have another look at what it is we are talking about. What is a meme?

In the first section of this small e-paper, I'll get back to basics and will offer a tangible definition of a meme. I will then move on to the next section and ask "what can we do with our knowledge of memes?"

## **What is a Meme?**

Richard Dawkins first came up with the idea of a meme in his 1976 book "The Selfish Gene". Essentially, memes are ideas that evolve according to the same principles that govern biological evolution. Think about all the ideas that you have in your head right now. They are all memes, and they all came from somewhere. Some of them will have come from friends and some will have come from the internet or television. Examples of memes are musical tunes, jokes, trends, fashions, catch phrases, and car designs. Now, the memes that inhabit your mind are in competition with all the other memes in the *memepool* (the collection of all existing memes). This means that they are all competing to get themselves copied into other people's minds. Some of these memes do quite well. Every time you whistle your favorite tune or utter a useful catch phrase, you are facilitating the spread of those memes. Every time you wear something that is "in fashion" you are helping the idea of that fashion enter other people's minds. Consider the first four notes of Beethoven's 5th symphony, or the "Happy Birthday" song. These are ideas that inhabit our minds and have been very successful at replicating. Not only have these memes found their way into literally millions of minds, they have also managed to leave copies of themselves on paper, in books, on audiotape, on compact disks, and in computer hard-drives (Silby 2000).

There is a limited amount of memetic storage space on this planet, so only the best memes manage to implant themselves. Memes that are good at replicating tend to leave more copies of themselves in minds and in other mediums such as books. Memes that are not so good at replicating tend to die out. We can imagine what sorts of memes have become extinct. Ancient songs that were once sung and never written down are one example. Another example is the many stories that were once told but have since slipped into oblivion. A Story is a vast collection of memes that have come to rely on each other for replication. Such a structure is known as a memplex. Stories are memplexes that are in direct competition with other memplexes. If a story replicates through story getting told and read by people, then it will survive. If it stops getting read, it will become extinct. Libraries are full of memetic fossils in the form of books that contain a multitude of ideas that are never looked at (Silby 2000).

You will see that memes behave in a similar way to genes. Furthermore, you will notice that like genes, memes are subject to selection pressures. Whenever you have a situation that contains a number of unique entities that are competing for limited resources, the entities that are better at reproducing will leave more copies of themselves. In the case of memetics, memes are competing for minds to inhabit, and those that are better at reproducing are those that manage to get expressed in behavior (for example, behavior such as whistling).

Defining memes as ideas is standard, but it gives rise to an objection. The objection goes like this:

*All this talk of memes and memetic evolution is meaningless unless we can identify exactly what a meme is. Ideas can come in all shapes and sizes, but there seems to be no way to identify their composite memes. How can we point to a memetic unit? How big is a meme? What is the difference between competing memes? How can they be distinguished from each other?*

These are good questions. To further highlight the problem with memetics, consider the first 4 notes of Beethoven's 5th symphony. This is a meme that has found its way into most people's minds. But how about the entire symphony? It too has found its way into the minds of many people. Is the whole symphony a meme? And if so, then what about the first four notes? What about the first 3 notes; or the first 5 notes? Are these all memes?

The best way to think of a memetic unit is to consider it to be the smallest idea that copies itself completely while remaining intact. So the first four notes of Beethoven's 5th is a meme, but the first 3 is not. The 4th note is always there making up the memetic unit. The entire symphony is a huge collection of small memetic units -- a memeplex. The memes that make up Beethoven's 5th might have been good individual replicators in Beethoven's day. Or they may have been attached to other memeplexes. Beethoven's mind collected these memes and somehow they got connected giving rise to his famous symphony. Now they depend upon each other for continued replication. Of course, the question remains. What is a memetic unit? How can we point to a meme? What are we talking about when we say that a meme is the smallest idea that can copy itself while remaining self contained and intact? The answer to this is quite simple. Memes are essentially sets of instructions that can be followed to produce behavior. Instructions can be encoded in either:

- 1) Musical notation,
- 2) Written text,
- 3) Visible (or vocal) action,
- 4) The neural structure of the brain.
- 5) Digitized structures in a computer

A meme that produces the behavior of whistling the first 4 notes of Beethoven's fifth can be encoded in any of these systems and it will give rise to the same behavior. When a mind encounters an instruction set that produces behavior, it can reproduce that behavior by creating an appropriate neural "program". The best way to think about this is to consider an analogy in the computer world. Imagine that a robot is developed which contains a number of built in programs. One of these programs gives it the ability to write small behavioral routines. Essentially the robot can alter its behavior by writing small programs. A feature of this program is that it allows the

robot to observe the behavior of other robots and write programs that produce the same behavior. In effect, it can imitate other robots. Now, these programs are memes. They are not a part of the robot's innate behavior -- rather, they are produced by imitation. Such programs can be translated into different languages and written down on paper. They can also be transmitted to other robots who read the instructions or who imitate the behavior and write their own programs.

This is precisely the sort of process that goes on in humans. At some distant point in history, biological evolution provided our ancestors with a capacity to imitate behavior. This meant that humans could observe the behavior of others and their brains would produce the neural wiring needed to produce the same behavior. A neural wiring pattern that produces behavior is essentially a list of instructions, which can be translated into other mediums -- written language, outward behavior, or computer code. A list of behavior producing instructions is the thing that replicates and spreads into the minds of others. A list of instructions is a meme.

## What can we do with Memetics?

Having a definition of a meme is one thing; doing something useful with it is another. How can we use our knowledge of memes? There are several applications for memetics. First, it can be used as an explanatory tool. Thinkers have been looking to aspects of human behavior and have been using memetics to offer explanations for why such behavior exists. Memetics can also be used to explain human creations such as technology, music, and literature. Memeticists look at an aspect of human creativity and then construct a memetic history that may have resulted in that aspect of creativity. Of course, constructing historical accounts of any sort of evolutionary process is a dangerous business. Evidence is fragmentary, and it is impossible to determine the truth from the "just so" stories.

Another approach is to deconstruct a human creation -- such as a piece of music -- and discover the components that brought the creation together. By doing this, memeticists may eventually come to understand why it is that certain memes manage to attach to each other for mutual survival in a memeplex. They may also discover what it is about certain groups of memes that make them such good replicators.

In addition to the above, memetics has the potential to enhance our study of psychology. In the future, psychologists may look to memetics to discover the origin of certain psychological conditions. Perhaps multiple personality disorder could be explained by the existence of two (or more) competing memplexes that each define a sense of self (Susan Blackmore (1999) calls such a memplex a 'selfplex'). The idea behind this thinking is that a human mind is basically a memetic construct. When a brain becomes inhabited by a suitable collection of memes, they form a mind and a selfplex develops. Anomalies such as psychological depression (non-physiological) or addiction might be explained by memetic viruses that influence the behavior of the selfplex.

Putting these possibilities aside, the ultimate goal of memetics should be its ability to *predict* behavior and the evolution of future memetic structures. Future memetic psychologists could use their knowledge of memetics to predict what will happen when people are exposed to certain combinations of memes. If they are successful at making such predictions, then it will be

possible to determine which combinations of memes will lead to the production of criminal behavior. Attempts could then be made to filter certain memes out of the memepool. Of course, this would open up a new debate on the wisdom of censorship and the purposive destruction of memes. Who, after all, would decide which memes to force into extinction?

## **Into the Future...**

The "meme" meme has successfully gotten itself entrenched in the memepool. It is spreading rapidly around the human species at the speed of light, and it will one day have infected everyone's mind. Its reproductive success is a testimony to its infectious power. There is something about the "meme" meme that makes it a good replicator. It fits in well with the other memes that inhabit our minds, and there is something about it that makes us want to communicate it to other people. We are enabling its survival.

Memes offer us a way to understand our psychology and the evolution of our thoughts, technology, artifacts, music, and art. They can be defined as small sets of instructions that produce behavior. When enough of these instructions get together in a brain, a mind develops. Such a mind can be understood and predicted by looking at its composite memes. With its explanatory power, and its potential to make behavioral predictions, memetics will become an essential addition to a psychologist's tool kit. As its success increases, memetics will take over where psychology has left off, and will become a driving force in the study of human behavior.

## **Memes were Assimilated from the Following Sources:**

Blackmore, Susan. (1999). *The Meme Machine*. Oxford University Press.

Dawkins, Richard. (1976). *The Selfish Gene*. Oxford: Oxford University Press.

Dawkins, Richard. (1986). *The Blind Watchmaker*. London: Longmans.

Dennett, Daniel C. (1991). *Consciousness Explained*. Penguin Books, 1993.

Dennett, Daniel C. (1995). *Darwin's Dangerous Idea*. Penguin Books, 1996.

Silby, Brent. (2000). "The Evolution of Technology: Exposing the Myth of Creative Design".  
[http://www.def-logic.com/articles/evolution\\_of\\_technology.html](http://www.def-logic.com/articles/evolution_of_technology.html)

# **EVOLUTION OF TECHNOLOGY: EXPOSING THE MYTH OF CREATIVE DESIGN**

**By**

**Brent Silby**

Department of Philosophy, University of Canterbury, New Zealand.

## **Evolution of Technology...**

Evolution in the biological world has been pretty much accepted. Those who once argued against it have settled down. Most people now accept that our existence is the result of an extremely long history of evolutionary process. This process is responsible for the diversity among species on this planet, and is also responsible for the development of conscious, creative human minds. In the past, evolutionary theories of biological life have been resisted for a number of reasons. One of these reasons is that it challenged our ancestor's view of themselves and humanity's place in the universe. If evolution is true, then it seems that we humans are not the result of purposive design -- rather, we are the result of an inconceivably long history of inheritance of random change. In the past, many thought this view was unacceptable. Even today, some people find the view unpalatable and insist that the human species is the result of intelligent design. Despite the resistance, however, scientific discovery has given us good reason to suppose that evolution by natural selection takes place. Fossil records, DNA analyses, and experiments with short generational span bacteria are examples of empirical data that can best be explained by evolution in the biological world (see the Lenski Lab <http://www.msu.edu/~lenski/> for experimental research on bacteria).

Does evolution remove humanity from its special place in the cosmos? It could be said that the truth of evolution does not alter the fact that humans, unlike other species, possess *creativity*. We have the ability to produce beautiful music, paint scenery, write poetry, and design artifacts. For most people, the common sense view of the human mind is that it is a source of *intellectual originality*. In addition to its artistic abilities, it can use its creativity to *design* technological devices that serve specific purposes. There are many examples of technology that has been designed to serve a specific function. Consider the watch (a well worn, but effective example). People needed to tell the time so they set about designing a timepiece. After many experiments and many design alterations our ancestors developed the first timepiece (it should also be noted that in doing so they inadvertently invented the concept of "running late", an unforgivable consequence).

In this paper, I will argue that the common sense view of the human mind is wrong. I will suggest that human creativity is a myth -- or at best, an illusion. To begin, I will outline the theory of evolution and will show that it can be applied to non-biological systems. I will then examine the so-called creative process, and will attempt to discover the mechanisms involved in designing technological artifacts. My conclusion will be that technological artifacts are not the result of creative thought -- rather, the creative process can more accurately be described as the execution of an evolutionary algorithm.

## 1. What is Evolution?

Let's quickly run through the theory of evolution. The beauty of evolution is its simplicity and its portability to a number of domains. Once you understand how the system works, you can see that it can be applied to several different fields of enquiry.

Charles Darwin is thought to be one of the first people to theorize about evolutionary process. His ideas were published in his famous "On The Origin of Species by Means of Natural Selection" in 1859 and have resulted in a drastic change in the way we think about our place in the universe. For such an influential theory, however, evolution is based on surprisingly few premises:

- 1) In ideal circumstances (limitless resources), populations will grow exponentially.
- 2) Resources are limited.
- 3) Because resources are limited, populations tend to remain stable.  
These first three premises were formulated by Thomas Malthus in his essay on the principle of population. The remaining premises are Darwin's additions.
- 4) Individuals within a population have unique characteristics.
- 5) An individual's characteristics are passed on to its offspring (inheritance).
- 6) Change can occur, and this sometimes results in offspring having slightly different characteristics to their parents'.
- 7) Given the fact that individuals have varied characteristics, it is reasonable to suppose that *some* individuals will have characteristics that give them a better success at acquiring resources and reproducing.
- 8) Given premise 5, it is reasonable to assume that characteristics that enhance an individual's survival and reproductive success will be passed on to subsequent generations -- this is *Natural Selection*.
- 9) Evolution is the result of this process.
- 10) Populations that possess the above characteristics will evolve.

(This list has been adapted from a list provided by Herbert, T. (2000) Historical Perspective, Darwin and Evolution).

You will see that there is no requirement for the population to be one of biological entities. The above system can be applied to a number of different population types. All you need for the system to work is a population of entities that are: 1) competing for resources, 2) have unique characteristics, 3) can reproduce, 4) produce offspring that inherit their characteristics, and 5)

whose offspring might be different. This is to say that the offspring may find themselves with characteristics that their parents did not have -- perhaps through some sort of random mutation, the details of which are not important here.

In the following sections, I will compare the biological world to the world of ideas. With the comparisons I draw, I will show that technology has developed according to the above evolutionary system.

## 2. How do you build a watch -- or a dog?

How do you *build* a piece of technology, say a wristwatch? The answer to this is obvious. You follow a list of instructions. These instructions will tell you what sort of hardware to purchase and how to put it together to produce a functional watch. If you want to go any further and ask "why" the watch does what it does, you simply have to appeal to its *designer*. The watch tells the time because that's what its designer intended it to do. Perhaps you purchased your instructions from a company like 'ACME' (the one that supplies the coyote with road runner catching equipment -- remember the old Warner Bros cartoons). You know that a lot of very clever people work at ACME, and so you know that if you follow the instructions correctly, you will build a fully functional watch. You don't need to understand how it works, because someone at ACME has already figured it out.

But what if you had no instructions? What if you had never seen a watch before? Now, ask the same question again. *How do you build* a wristwatch? The answer in this case is not so easy to find. Think about it, and you will probably find yourself seeking answers that rely on the existence of earlier examples of wristwatches -- or at least some knowledge of the passage of time. You might think to yourself that you would build a watch that is something like an old grandfather clock, just much smaller. Or perhaps you will design your watch based on the timing of the sun crossing the sky -- thereby utilizing your knowledge of the rotation of the Earth. But suppose these things do not exist. Suppose you have never considered the passage of time -- perhaps the Earth does not rotate -- and timepieces have not yet been invented. In such a scenario, would it be possible to design and construct a wristwatch? I think the answer is no!

Given your lack of knowledge about the passage of time, it seems that you would not even be able to conceive of the *need* for a wristwatch, much less design one from scratch. It seems difficult to believe that someone with no prior knowledge of watches could sit down and design a wristwatch, or any other sort of timepiece. But wait. Wristwatches exist, don't they? They didn't design themselves and they did not appear out of thin air, so surely *someone* must have designed them. If not, how could we explain their existence? Was it God? Or are we missing something? Let's turn to the biological world and ask an analogous question. How do you build an animal, say a dog? If you are a scientist in a laboratory, you might answer the question in the same way as before. You follow a list of instructions. In the case of animals, the list of instructions is known as a genome. The genome is an incredibly large collection of genes. Genes are chemical instructions, encoded in DNA (Deoxyribose Nucleic Acid), that direct the development of cells. A fertilized egg will contain genetic information from both its parents, for example a male and female dog. As the egg cells divide (first into 2 cells, then 4 cells, then 8 cells, and so on), they chemically read the instructions that are contained within their genetic blueprint. These

instructions tell them how to develop. Some will become skin cells; some will become brain cells, and so on and so forth. Eventually a fetus is formed, then a cute puppy, then a fully functional adult dog.

Now let's ask the further question: *why* does the dog behave and look the way it does? Unlike the wristwatch case, we cannot appeal to a designer because in the case of dogs the idea of a designer simply does not enter the picture. Dogs were not designed. Nevertheless, it is a good question and evolutionary biologists can provide us with an answer. The dog behaves and looks the way it does because of the genetic instructions that were used in its construction. But these genetic instructions were not provided by a designer. They are the result of an incredibly long history of subtle modification of earlier instruction sets. It is important to emphasize that these modifications were not guided by an intelligent process. They were the result of random change and transmission to offspring. A rundown of the process might go something like this:

*At some point in history, an ancestor of the contemporary dog existed. One day, one of these early dogs gave birth to offspring whose genetic instruction sets were ever so slightly different (perhaps because of some mutation due to something in the environment, or perhaps because the copying mechanism is not perfect). The difference may have been almost undetectable -- perhaps they had a slightly enhanced sense of smell. Now, because of their better sense of smell, these offspring were a little better at avoiding predators, finding food, and finding mates (very important for us mammals). Since genes are passed on to offspring, the change in the instruction set was reproduced in subsequent generations. Hence, dogs got better at using their noses.*

An important point that is often misunderstood here is that the development of slightly different characteristics *does not necessarily lead to the extinction of animals that possess the original characteristics*. They may continue to reproduce effectively for millions of years. Furthermore, descendants of the original animal may one day give birth to something that is slightly different in other ways, and it too will pass on its features to its offspring (if it manages to reproduce, of course). And thus we get a 'branching' effect in the history of life on Earth.

Genes are the key to this system. They provide the instructions for how to build an animal. They are reproduced in offspring, and sometimes they change. The genes that get themselves reproduced effectively -- perhaps by providing a creature with a better sense of smell -- are the ones that survive. Less successful genes, for example genes that reduce eyesight in dogs, will tend to have limited reproductive success and will tend to lose their representation in the *gene pool* (the total gene set on Earth).

The explanation works well. By iterating (repeating many times) a very simple process of subtle change and inheritance, we can account for the existence of extremely complex, well adapted dogs. Could a similar answer work for technology? Can a technological artifact such as a wristwatch be broken down into a collection of small components like genes? Is it possible to account for technology by grouping large numbers of these small components together? The answer to this question can be found by appealing to the study of *memetics*.

## 2.1. Selfish Memes

In his 1976 book "*The Selfish Gene*", Richard Dawkins introduced an entity known as a *meme*. Memes are entities that primarily inhabit human minds (but you can find them in other places as well). To express it simply, a meme is an idea. Some modern day examples of memes are musical phrases, jokes, trends, fashions, car designs, and poetry. Any thought or idea that has the capacity to replicate is a meme. A well used example of a meme is the first four notes of Beethoven's 5th symphony. Another example is the "Happy Birthday" song. These are ideas that inhabit our minds and have been very successful at replicating. Not only have these memes found their way into literally millions of minds, they have also managed to leave copies of themselves on paper, in books, on audiotape, on compact disks, and in computer hard-drives. At first glance the idea of a meme may seem trivially true. Of course ideas spread, what's the big deal? Well, the big deal is that memes behave in similar ways to genes, and in this way their behavior and development can be described in terms of evolution.

Like genes, memes are in competition with each other. While genes compete for representation in the gene pool, memes compete for representation in the *memepool* -- the huge collection of ideas that are currently circulating the world. Human minds have limited room, so only the best memes manage to implant themselves. Memes that are good at replicating tend to leave more copies of themselves in minds and in other mediums such as books. Memes that are not so good at replicating tend to die out. There is a gigantic history of extinct memes, but since they are extinct we do not know what they were. To know what they were and to mention them here would only be possible if they were still circulating in the memepool. But we can imagine what sorts of things they were. Ancient songs that were once sung and never written down are one example. Another example is the many stories that were once told but have since slipped into oblivion. A Story is a vast collection of memes (a memplex), which are subject to the same selection pressures as other memes. If they replicate through the story getting told and read by people, then they will survive. If they stop getting read, they die. Libraries are full of memetic fossils in the form of books that contain a multitude of ideas that are never looked at.

The memes that replicate the best are the ones that manage to fit into existing collections of memes. Music is a great example (see Dennett (1999) for an interesting examination of the evolution of music). Contemporary music conforms to a set of rules. If a song were produced that followed an entirely alien set of rules, it would not fit in with our existing set of ideas about music and would probably die out. But if a new song appears that follows all the rules, then it will fit in nicely and will copy itself from mind to mind and to CD's, MP3's, and hard-drives. New songs are collections of memes that have been formed through the subtle change of existing memes. This change may be the result of some sort of blending of ideas or random mutation perhaps resulting from imperfect copying. The result is a slightly different song that contains memes that are competing for representation in the memepool.

An important point to note is that memes replicate for their own sake and not for the sake of the minds they inhabit. If a meme appeared that resulted in people carrying out life threatening activities, it *could* still be successful if it managed to transmit itself to other minds before its host died. Some examples of such memes might be the ideas involved in parachuting, high speed motor racing, smoking cigarettes, or taking other dangerous drugs. The memes do not care about

the danger that people face when carrying out these activities. Memes are simply reproductive entities. If there is something about a meme that makes it good at getting copied into other minds or into other mediums, then it will continue to exist. If it mutates and produces a meme that is slightly better at replicating then the new meme will compete more effectively for representation in the memepool -- an event that could lead to the demise of an older meme.

Many thinkers have resisted the idea of memes on the grounds that it is not possible to determine exactly *what* the meme is. We know that it is an idea that replicates, but how can we *point* to a meme? How can we isolate a memetic unit? This is a valid question. After all, there seems to be big differences between memes such as the first four notes of Beethoven's fifth symphony, the phrase "You're damned if you do, and you're damned if you don't", and the behavior of shaking hands. How can each of these be considered to be single memes? A similar problem arises when scientists talk about genes. There is no real gene 'unit'. Genes are packets of information that are encoded in DNA. Distinct genes can vary markedly in length. The same goes for memes. Memes are best thought of as packets of information that can be encoded in a number of different mediums. They can be encoded in the complex neural architecture of the brain, and they can be encoded in magnetic patterns on a hard-drive. A memetic 'unit' can be described as a *self-contained information packet that reproduces*. So, the first *three* notes of Beethoven's fifth symphony do not constitute a meme because it is not a self-contained information packet that reproduces. The *fourth* note is required to complete the reproductive unit. Of course, the question of the rest of the symphony now arises. Is it made up of a multitude of four note units? The answer is no. The symphony *is* made up of a vast collection of memetic units, but they are all different lengths. The entire symphony constitutes a memeplex that has good replicative power, but many of the individual memes would not make it alone. Some do, however -- you often find small pieces of music reproduced in other work -- but for the most part, the memes are dependent on their counterparts in the memeplex for survival. It is possible that the memes that comprise the fifth symphony memeplex were common in Beethoven's day. He was undoubtedly influenced by the tunes and musical themes of his time, and these would have found their way into his work. Even though they would no longer survive on their own, they manage to survive by being a part of the large symphonic memeplex that continues to reproduce.

### 3. Does Memetics fit in with the Evolutionary Theory?

Now, let's consider evolution's premises (see section 1 above) and apply them to memetics. Is the evolution of ideas plausible?

The first premise states that in ideal circumstances, populations grow exponentially. In the case of memetics we can see that this is true. If you have an unlimited population of communicating minds -- ideal for memes -- then the population of memes *would* grow exponentially.

Communication would ensure that the memes are continually replicated and implanted in new minds. If an idea appears in one person's mind, and she transmits it to everyone she knows, who in turn transmit it to everyone they know, you get an exponential growth in replication of that meme -- hence the memetic population grows exponentially.

Premise two states that resources are limited. This is obvious in the case of memetics. There are a finite number of vehicles for memes, which include minds, written text, pictures, and objects like wristwatches.

In premise three, we state that limited resources lead to the stability of population. It is difficult to see that this is true of memetics because the memepool is continuing to grow. New people are being born and are continuing to create new ways of storing memes. Despite this, however, the memetic population is reasonably steady. Old ideas disappear and are replaced by new ideas -- there simply isn't the room for all of them to exist and keep replicating.

Premise four is very important. It states that individuals in a population have unique characteristics. The uniqueness of memes is obvious. Some memes are similar, but others are extremely different. Consider the difference (mentioned above) between the first four notes of Beethoven's fifth symphony and a behavior such as shaking hands when meeting people. These are distinct memes with unrelated, unique characteristics. One manifests itself as a sequence of sounds, either imagined or real, while the other manifests itself as an action.

From premise four we can go on to show the truth of premise five -- that an individual's characteristics are often passed on to its offspring. In the case of the hand shake, the meme reproduces when people observe the behavior and imitate it. Imitation keeps the characteristics of the behavior intact while copying the meme into the mind of the imitator. Whether it survives in the new mind is another question that depends upon the behavior being carried out again. Of course change can occur, as premise 6 goes on to state. This can be the result of memes merging with existing memes, or perhaps because of imperfect copying of the original meme -- sometimes stories change; sometimes behavior is not imitated accurately. Such changes give rise to new memes that compete to get themselves expressed and reproduced.

This brings us to premise 7, which states that some individuals will have characteristics that give them a better success at acquiring resources (minds, paper, hard-drives) and reproducing. Some memes enter minds and get stuck there -- never finding a way out -- until they degrade and disappear. Consider, for example, a joke that you might have heard several years ago. Its memetic structure copied itself into your mind, but for some reason it did not compete well with the memes that were already there. Consequently it never found an outlet. There was always something else that was more important to say -- more successful at reproducing. Eventually the joke's memetic structure degraded and faded away. Fortunately for the joke, many other people probably *did* utter it, and its propagation continued in the minds of others. But this is not always the case. There is an enormous history of memes that were once effective replicators but eventually died out because they could not compete with other memes. Old songs from the second century were probably very good survivors, until they were eventually replaced by memes whose characteristics enhanced their own reproductive success.

Now, given the fact that a meme's unique characteristics are passed on to its descendants, and given the fact that only the best survivors get to reproduce, we can assume that characteristics that enhance reproductive success will appear in subsequent generations. This process, as stated in premise 8, is known as *Natural Selection*.

The result of the process outlined here is memetic evolution.

Now that we have a definition of memetic behavior, and given the idea of what constitutes a memetic unit, we can turn our attention back to the development of the modern wristwatch.

## 4. A Brief History of Timepieces

Deep in our history, humans had very little understanding about the passage of time. We can suppose that their only real knowledge of time was that when it got dark, they went to sleep. This sort of behavior is innate and so does not fall under the category of memetics. One day, however, a human might have accidentally discovered a new use for watching the sun. This primitive person might have realized that lions and tigers do most of their hunting at dusk. If this happened, a new behavior would probably have emerged -- a behavior that would have led this intelligent person to seek shelter when the sun reached a certain position in the sky. Now humans, unlike most animals, have the ability to acquire new behavior through *imitation*. When members of the group noticed their friend seeking shelter in response to the changing position of the sun, some would have copied that behavior, thus giving rise to the transfer of the '*action at a specific time*' meme.

As the centuries passed, more time dependent behavior would have surfaced. The behavior that copied itself into the minds of many individuals tended to continue reproducing, while behavior that stopped occurring did not get imitated and eventually died out. Looking at the position of the sun to determine what action to take is a behavior that was very successfully copied into the minds of most of the human species. Eventually the 'look at the sun meme' mutated. Somehow it got merged with other ideas that were circulating, and gave rise to an idea that involved looking at the shadow of a stick for a more accurate impression of the passage of time. In time, the 'look at the shadow of a stick' meme was influenced by other memes that were appearing and developing at an exponential rate. Memes for constructing items out of stone and memes for using symbols to represent different times of the day, blended with the 'look at the shadow of a stick' meme to produce new memes that eventuated in the construction of the *sun-dial*.

Later in history, a gigantic collection of memes were at work producing mechanics and devices that were made from moving parts, like the water wheel and the windmill. Through some incredible combinations of memes, the mechanical memes became attached to the sun-dial memes and produced the first clock. A huge number of variations of the clock appeared prior to the appearance of the modern wristwatch. Ideas were adjusted as they were influenced by other ideas, and this process led to the refinement of the clock. Many early versions did not work very well and were not reproduced. Memes that led to the failure of early clocks did not appear in later versions and subsequently became extinct.

The modern wristwatch can be described as a collection of successful memes -- the ones that survived. When a watchmaker builds a new watch, he/she is not 'creating' or 'designing' something from scratch. He/She is pulling together a bunch of memes that have slowly evolved since the first humans noticed the sun's movement across the sky. The modern wristwatch owes its existence to the gradual evolution of memes -- an evolution that was driven, in part, by trial and error experiments on existing memetic constructs.

## 5. A Concluding set of Memes

I have attempted to show that the idea of creative design can be replaced with an evolutionary picture. The basic premises of evolutionary theory are not exclusive to the biological world, and can therefore be applied to other domains. In this paper, I have run through a description of memetics, and have shown that memes (or ideas) evolve in a similar fashion to genes. My aim has been to show that technological devices, such as the wristwatch, are a collection of memes that have evolved over the course of human history. If the assumptions I have made are correct, and if evolutionary theory is correct, then it seems that human creativity can be replaced with memetic evolution.

The human mind is full of memes that are competing for resources. The most effective memes are those that get themselves reproduced. So, in the case of the wristwatch, we can suggest that the designer's mind was buzzing with memes relating to previous versions of timepieces. Some of those memes were very good at getting themselves expressed, and in doing so, found themselves incorporated in the design of the wristwatch. The memes that were not so effective at getting expressed (for example the idea of attaching the watch to one's jacket via a chain) were not included in the new design and are now dying out.

Where did these wristwatch memes come from? Well, they originated deep in human history and looked much different than they do today. They jumped from mind to mind through the process of imitation. Sometimes they would change ever so slightly; perhaps by being influenced by other memes, or perhaps through imperfect imitation. Thousands of years of memetic transfer resulted in the memes that comprise the modern wristwatch. It was just a matter of time before they all accumulated in one mind, and the wristwatch was born.

### Memes were Assimilated from the Following Sources:

Blackmore, Susan. (1999). *The Meme Machine*. Oxford University Press.

Darwin, Charles. (1859). *On The Origin of Species by Means of Natural Selection*. London: Murray.

Dawkins, Richard. (1976). *The Selfish Gene*. Oxford: Oxford University Press.

Dawkins, Richard. (1986). *The Blind Watchmaker*. London: Longmans.

Dennett, Daniel C. (1991). *Consciousness Explained*. Penguin Books, 1993.

Dennett, Daniel C. (1995). *Darwin's Dangerous Idea*. Penguin Books, 1996.

Dennet, Daniel C. (1999). "The Evolution of Culture" in *Edge* online journal,  
<http://www.edge.org/documents/archive/edge52.html>

Herbert, Tom. (2000). *Historical Perspective, Darwin and Evolution*.  
[http://fig.cox.miami.edu/Faculty/Tom/bil160/03\\_darwin.html](http://fig.cox.miami.edu/Faculty/Tom/bil160/03_darwin.html)

Matthews, Robert. (1992). *Unravelling The Mind of God*. London, Great Britain: Virgin Books.

Kendal, Jeremy R. & Laland, Kevin N. (2000). "Mathematical Models for Memetics" in *Journal of Memetics - Evolutionary Models of Information Transmission*, Vol 4.  
[http://www.cpm.mmu.ac.uk/jom-emit/2000/vol4/kendal\\_jr&laland\\_kn.html](http://www.cpm.mmu.ac.uk/jom-emit/2000/vol4/kendal_jr&laland_kn.html)

Reveal, James L., F.L.S., Bottino, Paul J. and Delwiche, Charles F. (1997). *The Darwin-Wallace 1858 Evolution Paper*. <http://www.inform.umd.edu/PBIO/darwin/darwindex.html>

Reader, S. M. & K.N. Laland (1999). "Do animals have memes?" in *Journal of Memetics - Evolutionary Models of Information Transmission*, Vol 4. [http://www.cpm.mmu.ac.uk/jom-emit/1999/vol3/reader\\_sm&laland\\_kn.html](http://www.cpm.mmu.ac.uk/jom-emit/1999/vol3/reader_sm&laland_kn.html).

Silby, Brent. (2000). *Revealing the Language of Thought*.  
<http://www.def-logic.com/articles/RevealLanguageOfThought.html>

Silby, Mark. (2000). Personal Communication. Department of Genetics and Molecular Biology, University of Canterbury, New Zealand.

# A SHORT COMMENT FROM A BIOLOGIST ON WILLIAM BENZON'S ESSAY "CULTURE AS AN EVOLUTIONARY ARENA"

By  
**Hans-Cees Speel**

Graduate Research Assistant, Dept. of Policy Analysis, Delft University of Technology,  
Jaffalaan 5 PO Box 55015 2600 GA Delft, The Netherlands. E-mail hanscees@hanscees.com

This article is published in the *Journal of Social and Evolutionary Systems*, 20 (3): 309-322

## Introduction

Many biologists have written on issues of cultural evolution from the 'memetic' point of view. The starting-point for their theories is the definition of both gene and meme as replicators. If we indeed value new classifications and theories more when they are connected to other, generally accepted theories, then cultural evolutionary theories which include memetic views have potency. Evolutionary theories in biological sciences are accepted, and the memetic view is strongly analogous to the structure of those theories. However, memetics cannot grow without serious attention from non-biologists studying cultural evolution. Therefore, I was happy to see William Benzon's essay integrating the memetics view on cultural evolution with views from other disciplines. Benzon's article focuses on cultural evolution, drawing examples from biological evolution for comparison. In this short reply, I shall present ideas formed in biological theory with regard to the meme-gene analogy.

My reply concerns four issues. (1) I will negate the statement that biologists agree that classification should reflect phylogeny. (2) I will comment on the examples Benzon uses with respect to biological evolution. I will argue that the view that only species can be considered evolutionary systems is too limited. (3) I will expand on the use of the concept meme as a replicator, introducing concepts as interactor and selective retention system. (4) I will reply to his view on the genotype-phenotype concept. I will argue that his use of these concepts is incompatible with the replicator view as used in biology. (5) A final remark concerns the comparison Benzon makes between the complexity of cultural and biological evolutionary systems.

### 1. Biologists, classification and phylogeny

The first issue I would like to challenge is that all biologists agree on the view that taxonomic classification must reflect phylogeny. This means that classification trees should reflect descent. Hull's (1988b) book On the social and conceptual development of science shows that not all biologists hold this view. The views of different schools that do not hold this view vary from acknowledging evolution as important, but not a practical base of knowledge on which classification can be based, to the view that classification is a discipline with its own value regardless of evolutionary theories. Benzon refers only to those that do hold the view that taxonomic classification must reflect phylogeny, for instance Mayr, leaving out members of other schools of thought.

## **2. Biological evolutionary theory and species**

Benzon uses species as the sole example of biological evolutionary theory. However, Hull (1982, 1988a, b) states that although species might be the 'paradigm example' of evolution by natural selection, we are biased in this regard. Species are by definition sexually reproductive units, but most organisms reproduce asexually and cannot be classified as species per se. Examples are bacteria, viruses, and many plants and lower animals. Thus, only to speak of species as examples of evolutionary biology is not representative of empirical reality. Asexual reproduction is very important in biology. Borrowing is more common in the asexually than in sexually reproducing organisms. Therefore, it must not be overlooked in the discussion of the prevalence of borrowing in biological and cultural evolution.

In biology, viruses can transmit genes over lineage-boundaries, and in bacteria plasmids are exchanged between strains. Benzon mentions that some plant species may mix their gene-pools and that animals now also are suspect of this. Because borrowing is so extensive in 'lower' organisms biologists often cannot use the classification rules that were developed to classify species. Based on these observations I want to assert that cross-lineage borrowing (Campbell, 1965), or multi-parental transmission (Boulding, 1978; Boyd and Richerson, 1985) are not a difference in kind between cultural and biological evolution. As Hull (1982) states, the only difference in kind between biological and cultural evolution is intentionality, and currently I think he is right.

## **3. Evolution, replication, interaction and selective retention**

Benzon uses the concept meme from Dawkins (1976), and also states that physical memes are dispersed. By saying this, he seems to accept Dawkins' view of memes as replicators. Replicators are units of information that are copied, or in other words replicated. However, replication alone cannot describe evolution. Hull (1982, 1988a) has expanded the replication concept with processes of interaction. Together with interaction, replication can generate evolution by either natural or artificial selection, be it intentional or not. Selective retention systems (Campbell, 1974), sexually reproducing or not, show adaptation because there is interaction and thus selection. In cultural evolution this expansion of the interaction concept to the replication view does not always play a role, while in biological evolutionary views selection is almost always pre-supposed (Speel, 1996). Using this elaborated abstract view on evolution the following sentences of Benzon can be amended:

'I take it that the first requirement for understanding culture as an evolutionary domain is to find suitable parallels to the biological concepts of gene, phenotype, environment, and species. We need something like the genes to vary, be replicated, and passed on from generation to generation. We need something like the phenotype to adapt to the cultural environment. And we need something like the species to bear the collection of genes which creates the adapting phenotype'.

This becomes:

I take it that the first requirement for understanding culture as an evolutionary domain is to find suitable parallels to the biological concepts of replicator, interactor, environment, and retention system. We need replicators to vary, and to be passed on from generation to generation. We need interactors to compete in the cultural environment. And we need something like a retention system to hold the collection of memes and in which adaption can take place.

The beauty of using the concepts replication and interaction in biological evolutionary theory is that it allows us to omit the view that evolution is restricted to species, and to describe evolutionary mechanisms in general. It has led to an hierarchical view on evolution (Brandon, 1988) where both asexually and sexually reproducing organisms can be described in the same theoretical framework. In biological evolutionary theory, replication is used together with interaction to describe evolutionary processes. I suggest that we do the same in cultural evolutionary theory. The biological parallel of the replicator gene is the replicator meme which is passed on from generation to generation, but also from human to human, or from culture to culture. Of course, it is not the phenotype which adapts itself but the selective retention system. In theories concerning humans, such systems can be the mind or a society, but also an organization or a discipline in science. Accordingly, phenotypes can be represented as traits, ideas or physical artifacts which interact in selective events. In other words, a phenotype is the element that interacts with the selective environment. The next section elaborates on the genotype-phenotype distinction.

#### **4. The genotype-phenotype distinction**

Benzon proposes that psychological traits in the brain form the phenotype in memetic evolution, and that physical memes should be seen as the genotype. Several other authors have written about the phenotype issue (Plotkin and Odling-Smee, 1982) and about what the phenotype should consist of in memetic evolution (Heyes and Plotkin, 1989; Hull, 1982). I would like to use the replicator-interactor-lineage model to explain my point of view. In my terminology the memetic counterparts of genotype and phenotype are respectively, the "memo-type" and the "phemotype". Even though Benzon does use the meme concept by Dawkins, he denies that memes can be located in the brain (Benzon, pers. comm.). These two views are inconsistent. Dawkins defines memes as replicators. Following this line of thought, memes should be defined as any unit of culture that is copied, be it from human to human, from culture to culture, or otherwise. Benzon's physical memes do get dispersed over cultures and thus they comply with Dawkins definition. However, I think memes in the brain are replicators as well. In my view, it is clear that ideas, songs and norms reside in the brain somehow, as they are copied as units from human to human, but also from brains to books, or to physical air-waves. These observations of replications define them as memes that are located in the brain. Thus I argue that under the definition which makes physical memes replicators, brain-memes are also replicators. The view that brain memes are units of evolution is consistent with Dennet (1991) and Lynch (1996).

Benzon considers physical memes to be the memotype, and psychological traits to be the phemotype. I think that this division is not correct. In biological theory, genes are replicators that do not interact with the environment. Almost all replicators in biological theory are genotype. Traits or parts of the phenotype are the structures that are decisive in interaction. Thus in biological theory, interactors belong to the phenotype, and replicators to the genotype.

Interaction takes place with regard to the selective environment (Brandon, 1988). Unless there are severe reasons not to, the theoretical connections between replicators and genotype and interactors and phenotype in biology, should be followed when defining equivalent concepts in cultural evolution.

By definition, environments are considered to be selective if there is a kind of competition or weeding out. In compliance with Benzon's terminology, we could call a selective environment also an arena. In biological evolution, species are subject to selective processes in many arenas.

Sperm-cells compete for egg-cells, males compete for females, females and males compete for their genes to be reproduced, genes involved in "meiotic drive" can compete by intra-cellular traits for replication into sex-cells (Ridley, 1993). Thus many arenas can be distinguished in which many selection processes take place, possibly at the same time. I see no reason why cultural evolution cannot be described with multiple arenas as well. Physical and brain memes can be replicators and they can both interact in different arenas. However, when they interact they should be considered phenotype, and when they are replicated they form the memotype. When Benzon states that psychological traits in the brain are phenotype this means by definition that parts of those traits can serve as interactors. I am not quite sure what Benzon considers psychological traits to be, and if he means that those traits are interacting. But I do agree that a "trait" like an emotional attachment to norms is not copied directly and thus cannot be considered to be a meme.

In the preceding paragraphs I have taken the view that memes in the brain can be interactors but also replicators. This is because replicators and interactors are definitions of process. The very same norm, or dispersed idea, can act as a replicator in one case and as an interactor in another. Norms are dispersed among humans by replication, and sometimes replicated from culture to culture. But ones established in a culture, these same norms can prevent other norms from entering the culture of which they have become a part. In this example such norms act as replicators first, and then as interactors, which compete with new incoming norms. In summary, Benzon's proposal that psychological traits are phenotype, and physical memes are memotype seems to be inconsistent with the underlying definitions from biological evolutionary theory. It is the role cultural units play in the evolutionary process which determines if they should be considered one or the other.

## **5. Is culture a more complex phenomenon than life?**

The very last issue I want to address is a small one, but one I have seen many times before. Therefore I would like to comment on it. Benzon states in his last paragraphs:

If cultural taxonomy is more complex and various in structure than biological taxonomy, that is surely because culture is a more complex phenomenon than life. Memes and traits flow between paradigms more freely than genes and phenotypes flow between species. Abstract cultural space is richer than abstract biological space.

I do not know if cultural taxonomy is more complex than biological taxonomy and do not think that phenotypes can flow between species. However, the next sentence seems to be the common

belief of non-biologists: 'Culture is a more complex phenomenon than life'. Sometimes I hear statements as 'human processes are more complex than biological processes'. I can only hope these are not seriously meant. My critique follows several lines. With regard to method, the measurement tools must be indicated when we make comparisons. Can we compare the evolution of culture and biological evolution with the same measurement device? If we cannot, the statement above is useless. Second, what is this 'life' that is being compared? And should we regard the evolution of one species more complex than the evolution of a culture? Or than the evolution of multi-species systems, or the evolution of eco-systems? Is a handful of soil with millions of bacteria that react to micro-climates and belong to thousands or more strains less complex than the evolution of culture? I do not support the statement that culture is more complex than life, but would be interested if anyone would attempt to compare this in a useful and valid manner. My interpretation of Benzon's statement is that culture is a specific domain of research; that intentionality is more present in culture than in biological objects of research; and that language with words, concepts and logic found in human interaction is not present in the objects of biological evolutionary research.

## **Conclusion**

In conclusion, part of Benzon's article is an attempt to compare cultural and biological evolutionary processes. He states that we have a long way to go doing this and I think progress can be made. The integration of the concepts replicators, interactors and selective retention systems can be useful in the study of cultural evolution. If we want to make the effort worthwhile, it would be wise to involve both cultural and biological scientists and their insights. I fully agree with the last statement of Benzon:

We have plenty of good work yet ahead, and a measure of pleasure as well.

## **Acknowledgements**

I thank William Benzon for his valuable critique on earlier versions and Martine van der Ploeg for improving my English.

## **References**

- Boulding, K.E. (1978) Ecodynamics: a new theory of societal evolution. Sage publications, London.
- Boyd, R. and Richerson, P.J. (1985) Culture and the evolutionary process. The University of Chicago Press, Chicago.
- Brandon, R.N. (1988) The levels of selection: A hierarchy of interactors. In: Plotkin, H.C. (ed) The role of behavior in evolution, MIT Press.
- Campbell DT (1965) Variation and selective retention in socio-cultural evolution. In: Barringer, H.R.; Blanksten, G.I. and Mack, R.W. (ed) Social change in developing areas, a reinterpretation of evolutionary theory. Schenkman publishing company, Cambridge Massachusetts.

Campbell, D.T. (1974) Evolutionary epistemology. In: Schlipp, P.A. (ed) The philosophy of Karl Popper. The Library of Living Philosophers volume XIV.

Dawkins, R. (1976) The selfish gene. Oxford university press.

Dennett, D. (1991). Consciousness explained. Penguin Books.

Heyes, C.M. and Plotkin, H.C. (1989) Replicators and interactors in cultural evolution. In: Ruse, M. (ed) What the philosophy of biology is; essays dedicated to David Hull. Kluwer academic publishers Dordrecht.

Hull, D.L. (1982) The naked meme. In: Plotkin, H.C. (ed.) Learning Development and culture, essays in evolutionary epistemology. John Wiley & Sons.

Hull, D.L. (1988a) Interactors versus vehicles. In: Plotkin, H.C. (ed) The role of behavior in evolution. MIT Press.

Hull, D.L. (1988b) Science as a Process: An Evolutionary Account of the Social and Conceptual Development of Science. The University of Chicago Press.

Plotkin, H.C. and Odling-Smee, F.J. (1982) Learning in the context of a hierarchy of knowledge gaining processes. In: Plotkin, H.C. (ed.) Learning Development and culture, essays in evolutionary epistemology. John Wiley & Sons.

Lynch, A.(1996) Thought contagion. How belief spreads through society. Basic Books.

Ridley, M. (1993) Sex and the Evolution of Human Nature. Viking.

Speel, H.C.A.M. (1996) Memetics: On a conceptual framework for cultural evolution. In:

Heylighen, F. and Aerts, D. (ed) The Evolution of Complexity. Kluwer, Dordrecht, in press.

# **MEMES ARE ALSO INTERACTORS**

**By**

**H.C.A.M. Speel**

Research assistant Policy analysis.

Faculty of Technology, Policy and Management

Delft University of Technology

Tel. 00 31 15 2785776

Fax. 00 31 15 2783422

E-mail [hanscees@hanscees.com](mailto:hanscees@hanscees.com)

## **Summary**

Dawkins (1976) defined memes analogous to genes as replicators, and rightfully so. However, it is not generally mentioned and thus probably often overlooked that memes can also be interactors. Both interaction and replication is needed to explain natural selection (Hull, 1988b). Memes, but also genes are interactors if their direct characteristics count in selective events. Hull (1988a) mentions humans as interactors in memetic evolution, not memes. I argue (Speel, 1997) that if we can and do judge memes by their merits without necessary interference of the physical world, which implies a kind of phenotype or conceptual phenotype, memes should count as interactors. If we take the definitions of Hull and Dawkins seriously we cannot deny this conclusion.

## **Introduction**

Memes are replicators by definition, in analogy to genes. This is the common denominator in memetics, and it is recognized by most writers on the subject. Dawkins (1976) started the definition of memes in this way, and where memeticists differ in opinion on what should count as phenotype-genotype distinction (see Hull, 1988a for some opinions) and even on what memes are (see Benzon (1998) versus Speel (1998)), they do agree that memes are replicators.

A large part of memetics focuses mainly on the dissemination part of memetics that is directly related to replication. The meme as thought contagion is a well-known view that focuses mainly on memetic processes of dissemination. Aaron Lynch's *Thought Contagion* (1996) is perhaps the main work giving the meme such connotations. Of course memes may be rightfully seen as disseminating through human systems, but this view alone cannot account for natural or artificial selection involving memes.

Any theory accounting for adaptations by natural or artificial selection must include processes of interaction (processes such as competition and predation in biology) and processes that account for new variation and possibly the re-combination of old variation. I have argued before (Speel, 1996) that memetic theory such as proposed by Dawkins (1976) does not make it clear if meme-complexes such as religions have adaptations and thus adapt. Worse, it does not provide us with possible mechanisms for such adaptations. The same can be said for the work of Lynch (1988) on religion that also lacks explanations on how religions as meme-plexes (co-adapted meme-

complexes) somehow incorporate adaptations, by which some religions become more successfully disseminated than others.

This lack of focus on adaptations and mechanisms perhaps explains why memetic theory has only incorporated that memes are replicators, but not that memes are interactors.

In the following pages I shall argue that memes are by definition replicators, but that in 'internal' selection they can be interactors as well. This view is consistent with the view that RNA selection is possible where there are no translation steps such as from DNA to RNA to enzymes (Calvin, 1997). Hull's (1988a) view that genes can be interactors as well is also consistent with this view.

If memes are taken to be interactors, consequences follow for the discussion on what the phenotype must be in memetic evolution. I shall argue, contrary to Benzon (1998), Hull (1988a) and others, that memetic (sometimes vicarious) selection implies two kinds of phenotypes.

## Memes can be interactors

Replicators in biological evolutionary theory were introduced by Dawkins (1976) and have become part of the received view on natural selection. Hull has argued that if the theory of natural selection is to be described fully, interactors are the necessary counterpart of replicators. Hull (1980) applies the following definitions in biological evolutionary theory:

Replicators, or the 'structural entities that are replicated', pass on their structure directly from generation to generation. In principle replicators can only replicate themselves, but they can also produce entities that interact and this indirectly results in the replication of the replicators (by reproduction). These entities are known as:

Interactors: 'entities that directly interact as a cohesive whole with their environment in such a way that replication is differential'.

In the paradigm example of evolution by natural selection, a sexual reproducing species, the paradigm replicators are the genes and the paradigm interactors are the organisms. Hull argues that replication or interaction alone cannot account for evolution by natural selection, but that both processes are needed:

Selection becomes defined as 'a process in which the differential extinction and proliferation of interaction causes the differential perpetuation of the replicators that produced them'.

Lineage is the third term in the Hull scheme and refers in this case to the historical changing entity called the species. Notice that Hull thus takes selection to be more than just 'weeding out'. For him selection refers to evolution by selection on replicators that form lineages.

Interactors play an essential role in biological evolution theory as counterpart of replicators. Brandon (1988) sees organisms, genes, groups and taxa as interactors in hierarchical levels of interaction. It might be noted that different stadia in reproductive cycles often amount to interaction in different arenas: sperm-cells usually compete for a few or only one egg cell; genes in meiosis can compete for inclusion in the sexual cells (meiosis-distorters) and so on. Beyond

this, there is probably also interaction between DNA in chloroplasts and other cell-organelles (Eberhart, 1980). It should be noted that some (Dawkins, 1982) use to prefer the concept 'vehicle' in stead of 'interactor'.

The interactor-lineage-replicator scheme as defined above, can and is used in memetic evolutionary processes with only minor changes to the definitions (Hull, 1988b; Speel, 1997). The question then arises what the interactors should be in memetic evolution, since it is of course clear that the replicators are memes. Hull (1988b) has argued that humans in memetic processes are interactors, although he later mentions that humans are agents of some sort in memetic evolution (1988a). He consequently holds memes not to produce interactors since of course humans are not produced by memes (1988b).

Human individuals can indeed be interactors in memetic evolution, but I have argued before (Speel, 1997) that memes can also be interactors in 'internal' interaction or selective interaction (s-interaction). [Footnote: I prefer 'selective-interaction', or 's-interaction' to just 'interaction' because this circumvents the confusion with 'interaction' in social sciences that usually means something like 'individuals reacting in some way to each other'. S-interaction implies a result in differential perpetuation of replicators.] Internal selection is selection inside a mind when someone makes up his mind on whether A or B is better, true, more important, or something equal to such criteria [footnote: Note that this is not the same as a decision to remember or to notice something. It is quite possible to know something, and not to agree with it.]. External interaction is interaction where human individuals are the interactors. In external s-interaction human individuals do not question memes themselves in their minds, but they defend memes against other human individuals that act to discredit the same memes. When humans thus in practice act to defend a meme, the interaction is more directly between humans, and not directly between memes. While this internal and external difference of s-interaction might seem far-fetched it is needed to make memetic terminology suited to describe internal selection. If only humans can be interactors in memetic evolution, which Hull seems to propose, there is simply no way to describe whatever factors inside a brain, or making up a mind, are responsible for the selection of ideas. Such factors might be called installed complexes of memes (Dennett, 1991), or a self. I think memetic theory ought to describe the internal selection of proposals, theories etc. which we all witness in our own minds, but that we can also verify in other ways, for example by the study of verbal and other behavior.

## **Interactors are part of the phenotype**

However, if memes are replicators, can they also be interactors? Can the same entity that is a replicator be an interactor? Can an entity that is code, geno-or memotype, be part of the pheno- or phenotype?

First of all replication and s-interaction are definitions of process or in other words functional definitions (and not structural definitions). If I use a chair to sit on, it is a seat, when I use it to hit someone it might be called a weapon. In principle words like interactor and replicator are definitions of process. The role an entity plays or has played in a process at a specific time and place is decisive for such a definition.

Memes can thus not be excluded from the interaction roles by definition a priori. In fact genes and RNA are also said to be both interactor and replicator. Hull (1988a) acknowledges that genes can be interactors, and Calvin (1997) mentions research on RNA:

"Weismann's genotype-phenotype distinction in biology is not a necessary condition for a Darwinian process, as recent experiments on "RNA evolution" have shown (there isn't a body that lives and dies, carrying the genes along, but rather patterns directly exposed to environmental selection). Envelopes such as bodies (phenotypes) are an example of stratified stability; they nicely illustrate why strict one-trait-at-a-time adaptationist reasoning is insufficient."

Calvin thus argues that parts of RNA, at least in experiments, can be both interactors and replicators.

Parts of DNA are usually not interactors: either a body of an organism or enzymes are the interacting entities coded for by the DNA. However, in some instances the direct physical characteristics of DNA might be decisive in s-interaction: Eberhart (1980) mentions the hypothesis that the size of DNA molecules might be of importance in DNA selection in cell-organelles. DNA would be s-interactor if there was selection on the form of DNA molecules, the stability of them, for instance in a high-temperature environment and so on. In RNA evolution RNA could be interactor if the chemical groups of the RNA-molecules are involved themselves in processes that are decisive for their survival.

We have seen that replicators are also allowed to be interactors because these concepts are definitions of process. We have also seen that genes and RNA can fulfill both roles. The conclusion must be that there is no reason why memes cannot be both replicators and interactors. Notice that the general evolutionary terminology of geno- and phenotype can be caught in the following definitions:

- An entity is part of the phenotype if it is directly subject to selection.
- An entity is part of the genotype if it contains a kind of code, and is replicated
- For memetics we can simply substitute pheno and geno with phemo and memo, and we have the memetic definitions:
- An entity is part of the phemotype if it is directly subject to selection.
- An entity is part of the memotype if it contains a kind of code, and is replicated

This leaves us with a problem: why should human individuals be the only interactors in memetic evolution as Hull (1988a) holds? If theories are tested, or only small parts of them, should not the theories be the s-interactors? Hull does speak of conceptual phenotypes (Hull, 1988a), but of scientists as s-interactors. I think that if characteristics of a theory describing reality are the decisive things that make a theory rejected or not, the theory (or a part of it) is a s-interactor. If the academics proposing the theory would be the interactors, they would be candidates for weeding out. Of course scientists can be s-interactors: they might socially become isolated, not respected, ridiculed or even released from their duties, which can count as a kind of social weeding out. But the weeding out of scientists would be a very different matter from the weeding out of the theories they defend. Ultimately scientists and the theory they defend might be weeded

out together, but this need not be of course. I think we need to be able to differentiate between the two selective processes, and the difference between internal and external s-interaction makes this possible.

## S-interaction and vicarious selection

Memetic evolution largely goes on within people's heads. Most memes never become accepted by the mind they are at some moment being reviewed in and get forgotten or are remembered as not being useful, untrue, and so on. Memes, being descriptions of situations, solutions, problems, opinions and so on can play a role in the behavior of human individuals and organizations. In such a view, memetic evolution inside people's minds is vicarious selection (Campbell, 1974), 'producing future' on behalf of the individual or organization involved. It is true that the behavior individuals and organizations show is part of the phenotype of the individual human or organization. When descriptions of such behavior are vicariously pre-tested in the mind of someone, against a virtual model of the real world, this is a test too. It is virtual in the sense that the test stands for something else, a possible action in the real world, but it is a real test in the sense that something is rejected in someone's mind. Even if memes are internally tested without having an effect on human behaviour, it is still a selection- or weeding-out event. If memes are thus internally selected, they are phenotype (or conceptual phenotype), and thus interactors.

## Conclusion

Memes are usually defined as replicators, and rightfully so. However, memes can also be tested inside minds, which makes them interactors. Human individuals can also be interactors in memetic evolution if they 'defend' memes in their struggle to be considered true, useful and so on. If a power struggle in for instance politics results in one party dictating policies, the party (or the party members) was (or were) the interactor(s), not the memes. They too won, but through selection of the party (members).

If memetic theory is to include explanations of why specific memes win from other memes, the focus on only the dissemination and replication of memes must be widened. Interaction of memes and humans must be included amongst other things, to explain both adaptations of meme-plexes (co-adapted meme complexes), and adaptations of human and organizational behavior by means of vicarious memetic evolution.

## References

Benzon, W (1998) Culture as an Evolutionary Arena. *Journal of Social and Evolutionary Systems* (in press).

Brandon RN (1988) The levels of selection: A hierarchy of interactors. In: Plotkin HC (ed.) *The Role of Behavior in Evolution*, Mit Press.

Calvin, W.H. (1997) The Six Essentials? Minimal Requirements for the Darwinian Bootstrapping of Quality. *Journal of Memetics - Evolutionary Models of Information Transmission*, 1.

[http://www.cpm.mmu.ac.uk/jom-emit/1997/vol1/calvin\\_wh.html](http://www.cpm.mmu.ac.uk/jom-emit/1997/vol1/calvin_wh.html)

Dawkins, R. (1976) *The Selfish Gene*. Oxford University Press

Dawkins, R. (1982) *The Extended Phenotype: The long reach of the gene*. Paperback. Oxford University Press, New York

Dennett, D.C. (1991) *Consciousness explained*. Boston, MA: Little, Brown & Company.

Eberhard, W.G. (1980) Evolutionary consequences of intracellular organelle competition. Quarterly review of biology 55, 231-249

Hull, D.L. (1980) Individuality and selection. Ann. Rev. Ecol. Syst., 11, 311-332

Hull, D.L. (1988a) A period of development: a response. Biology and Philosophy 3, 241-263

Hull, D.L. (1988b) *Science as a Process. An evolutionary account of the Social and conceptual Development of Science*. The University of Chicago Press

Lynch, A. (1996) *Thought Contagion: How Belief Spreads through Society: The New Science of Memes*. Basic Books, New York.

Speel, H.C.A.M. (1996) Memetics: On a conceptual framework for cultural evolution. In:

Heylighen, F. and Aerts, D. (ed.) *The Evolution of Complexity*. Kluwer, Dordrecht (in press).

Speel, H-C. (1997) A Memetic Analysis of Policy Making. Journal of Memetics - Evolutionary Models of Information Transmission, 1.

[http://www.cpm.mmu.ac.uk/jom-emit/1997/vol1/speel\\_h-c.html](http://www.cpm.mmu.ac.uk/jom-emit/1997/vol1/speel_h-c.html)

Speel, H.C.A.M. (1998) A Short Comment from a Biologist on William Benzon's essay 'Culture as an Evolutionary Arena'. Journal of social and evolutionary systems (in press)

# **ON MEMETICS AND MEMES AS BRAIN-ENTITIES**

## **A commentary on Gatherer's paper: Why the 'Thought Contagion' Metaphor is Retarding the Progress of Memetics**

**By**

**Hans-Cees Speel**

Faculty of Systems Engineering, Policy Analysis and Management

Department of Policy Analysis

Jaffalaan 5, PO box 5015, 2600 GA Delft, The Netherlands

[hanss@sepa.tudelft.nl](mailto:hanss@sepa.tudelft.nl)

Derek Gatherer argues quite strongly against memes defined as brain entities and for a memetics that sees memes as cultural artefact-replicators, including texts, parts of texts and so on (Personal communication). He argues that only memes as artefacts can be observed directly and that memetics is retarded by the use of the brain- entity definition. In doing so he refers to me (Speel, 1997b) as a user of the orthodox Dawkins B definition or of a close relative of that definition. I will argue that his claim for memes as cultural entities is valuable, but his dismissal of the meme as a brain entity is not. I will also show that his reference to me is not entirely correct: my definition of a meme is meant to allow both memes as brain entities and memes as cultural artefacts, where both can be member of identical memetic lineages.

Before I will start with the points above I want to thank Gatherer for his thought-provoking article. Although I do not share all his conclusions, I do share the care for analysis of the definitions memetics uses coupled to the question how memetics as a scientific endeavour can advance (A while ago in (Speel, 1997a) I commented on a paper by Bill Benzon (1997) in which he reached conclusions similar to those reached by Gatherer). However, it was only after having read Gatherer's paper that I realised properly that the memes I study in public policy are in fact almost always texts and phrases, and not brain entities. Most of the memes that show up in empirical studies in JoM-EMIT are cultural entities or Dawkins A memes as well.

However, I can see no evidence that memes seen as brain replicators halter the memetic endeavour. I think that memetics must include more empirical study. Therefore a better involvement with scientists that study cultural memes would be positive. In this regard I agree with Gatherer, who implicitly claims that memetics needs to include more empirical study.

However, I do not think that disregarding the meme as brain entity will help us at all. The meme as a brain entity is needed for the simple reason that cultural entities, like words in texts that function with specific meanings, have no meaning without referral to these brain entities. What makes a vase a vase is that humans think it is a vase. A vase can only be replicated if people do this replicating (or their machines). In other words, the replication of a cultural replicator can often not be described properly without describing some brain entity form of the same replicator.

To take another example, if words in texts that we grant to be concepts are only replicated after giving them thought, we cannot overlook the part of replication where the word is a brain entity. The word in the text replicates into the brain, making the brain entity and the word direct

replicators in the same lineage. Therefore memes as brain entities must be included in memetic theory, even though empirically we will often only see the word in the text.

Moreover, memetic cultural entities cannot be explained or defined without the notion that they are replicated with the involvement of (human) brains, or have been in the past.

Derek's arguments against memes as brain entities seem to boil down to mainly one thing: memes as brain entities are hard to define properly. I agree with that concern but not with his conclusion. There are two arguments I want to put forward.

First of all, there are also problems with definitions of memes as cultural artefacts. As an example I can refer to the analogous entity gene that to this day is defined in different ways in different fields. It is well known that the gene as used by Dawkins (1976) comes from G. C. Williams and is not the same as the referral to a gene as a piece of DNA that codes for a protein for instance. More directly relevant, Susan Blackmore (1988) has very recently shown that there are 'problems' with the definition of behavioral memes as well.

Secondly, I think that many problems with the definitions of a meme as a brain entity can and will be solved. Just stating that there are problems, such as the question if beliefs are memes, does not qualify the conclusion that they cannot be solved let alone that the question should be abolished. I think the question on what kinds of beliefs are memes or not needs an answer. The same argument is valid for the question if memes as brain entities can be measured. Measuring what people believe or think might be difficult, but not impossible. Surely the measurements take place in indirect ways (by interview for instance), but that is perfectly valid. I would think that many measurements in physics are not direct either: how direct are sub-atomic particles measured? Examples like the ones with the Windsor knot just show that the relation between cultural entities and brain entities is a complicated one so that we must be careful with our definitions and analysis.

Another argument against memes as brain entities is that people do not have memes, while organisms do have genes. Gatherer refers to this notion as the meme-host duality that he claims is wrong. I find it hard to reply to this argument because I do not understand the sense of it at all. It might be my lack of understanding of the English language, but because Gatherer puts so much emphasis on this argument it does need a response. I fail to see in any way that saying that 'people have memes' is wrong. It is just a manner of speaking as far as I know. Just as there is a duality between genes having organisms and organisms having genes, there is such a duality between memes and people. People can have memes, like I have knowledge of the word democracy and its meanings, and I also adhere to the values of democracy. This does not mean that memes like a style in pottery that spreads cannot be studied. I fail to see any problem. Perhaps Gatherer tries to show that the spread of memes cannot always be measured by referring to human individuals alone. There are many ways in which memes have a social status unrelated to specific individuals. Official reports of governments and concepts in it are such memes that are difficult to measure with referral to human individuals alone. But this does not qualify his strong conclusions at all.

According to my arguments above it might become clear that I see memes as replicators that can form lineages from artefacts to brains and so on. If I see a cultural entity like a Windsor knot I

might want to learn how to tie it, and ask my mother or father to show how to do this. When I noticed the knot and realised there were specific knots and that a specific one was called the Windsor knot, the cultural entity was replicated to my brain. After I have learned it, I had mastered the complex movements needed to tie the knot, and the tying-meme had replicated from the body of my parents to mine. Not a simple replication, but a replication nonetheless. I may forget all the things needed to learn the tying-patterns, but I do still know how to tie it. If we look very hard, the knot itself became (and was already) a meme when I noticed it, but the tying pattern another one. Often the analysis of memes is difficult, but not impossible. Often a meme-lineage 'jumps' from brain to cultural artefact. This cannot be ignored and therefore memetic theory needs to incorporate memes as brain entities.

The definition (Speel, 1997b) of memes I previously used is: '*pieces of data that are a) copied from individual to individual without too much alteration, or B) that are interactors*'. I cannot blame Gatherer for thinking I probably dismiss memes as cultural artifacts, since my definition does not mention them at all. The empirical matter I discuss in my papers however makes clear that I use texts and concepts in texts as memes. I do think however that a cultural artifact is only a meme if it has been replicated by a brain in its past. I do not mean to say that this definition will do in any circumstance, but it will in my research.

In conclusion I think that the current discussion is important. I hope memeticists become more aware of problems of definitions, but I also hope that they do not let such issues stop them from empirical research. I have argued that Gatherer's arguments need attention, for instance the question of the status of beliefs in memetic theory. However, narrowing the meaning of meme to a cultural replicator only seems unnecessary and counterproductive to me.

## References

- Benzon, W (1996) Culture as an Evolutionary Arena. *Journal of Social and Evolutionary Systems*, **19**:321-362. <http://www.newsavanna.com/wlb/CE/Arena/Arena00.shtml>
- Blackmore, S (1998). Imitation and the definition of a meme. *Journal of Memetics - Evolutionary Models of Information Transmission*, **2**. [http://jom-emit.cfpm.org/1998/vol2/balckmore\\_s.html](http://jom-emit.cfpm.org/1998/vol2/balckmore_s.html)
- Dawkins, R. (1976). *The Selfish Gene*. Oxford: Oxford University Press.
- Gatherer, D. (1998). Why the Thought Contagion Metaphor is Retarding the Progress of Memetics. *Journal of Memetics - Evolutionary Models of Information Transmission*, **2**. [http://jom-emit.cfpm.org/1998/vol2/gatherer\\_d.html](http://jom-emit.cfpm.org/1998/vol2/gatherer_d.html)
- Speel, H.C.A.M. (1997a) A short comment from a biologist. *Journal of Social and Evolutionary Systems*, **20**:309-32.2 <http://www.sepa.tudelft.nl/webstaf/hanss/jses.htm>
- Speel, H-C. (1997b); A Memetic Analysis of Policy Making. *Journal of Memetics - Evolutionary Models of Information Transmission*, **1**. [http://jom-emit.cfpm.org/1997/vol1/speel\\_h-c.html](http://jom-emit.cfpm.org/1997/vol1/speel_h-c.html)

# A MEMETIC ANALYSIS OF POLICY MAKING\*

By

Hans-Cees Speel

Faculty of Systems Engineering, Policy Analysis and Management

Department of Policy Analysis

Jaffalaan 5, PO box 5015, 2600 GA Delft, The Netherlands

[hanss@sepa.tudelft.nl](mailto:hanss@sepa.tudelft.nl)

## Abstract

A memetic framework is presented for the analysis of policy making. This framework is based on three concepts from modern biological evolutionary theory: interaction, replication and lineage. The framework consists of a view where descriptions of actions and people endorsing them, compete to get these action proposals into policy plans. This framework enables the inclusion of rational, non-rational and other selection forces, having their effects when choices are made. Because it also deals with the historical legacy of the ideas used to formulate a policy, it expands on the concept of 'bounded rationality', limiting the role of rationality or reasoning in making choices even further. It is argued that the degree of replication, versus rational or intentional deliberation in memetic evolution can yield interesting hypotheses. 'Weaseling', a mechanism for memetic variation production, is introduced as a new difference between genetic and memetic evolution. Levels of retention are introduced as differences in success criteria for memes.

**Keywords:** Evolution; Policy Making; Meme; University; Organization; Selection;

## 1. Introduction

This paper aims at establishing a memetic framework of definitions with which policy making in public and private organizations can be described and studied. We argue that the framework is useful because it systematically uses the language of the evolutionary paradigm. In doing so, this paper connects memetic theory with theory about public policy, and so establishes a link between biological evolutionary theory and theory of public policy.

The framework establishes a view of public policy processes that puts the importance of 'conscious' and intentional reasoning into perspective. As Price rightly (1995) [23] points out 'Companies are not only systems created and controlled by those who manage them but... an organization can be seen as a product of .... the meme'. This view is taken up and elaborated upon by considering policy processes to be memetic selective processes<sup>9</sup>. Rational, non-rational and other selective forces are looked upon as possible explanations in policy making. It is argued that the memetic view presented establishes an elaboration on and is complementary to the widely accepted concept of 'bounded rationality'<sup>1</sup> (Simon, 1976) [26].

Much has been written on memetics. However, clarity of definitions and a clear reference to specific cases are not often part of what has been written. For instance Price (1995) [23], while

writing about 'organizational memetics', comes up with very few definitions, and when he does they are vaguely defined while their usefulness is sometimes doubtful (he defines meme as 'mental model' for instance). Other confusions arise when levels of retention are not distinguished (see section 2.4).

The authors feel that clarity of reasoning is dependent on clarity of definitions, and that considerable progress in memetic theory can be achieved by this. While assessing the plausibility of (hypo) theses and the review of different mechanisms are important to progress in theory, it depends on defining clear concepts. Hypotheses must be conceived first, and can only be meaningful if underlying concepts are clearly defined. One example in biological evolutionary might be the 'levels of selection' debate which was at least partly resolved by discerning replication from interaction (Greene, 1989) [12]. We argue that, in memetics, concepts are often still unnecessarily ambiguous and that thus the establishment of clear definitions is important. Therefore this paper aims at providing such definitions in the form of a framework upon which further research can be done, pointing out difficulties where they arise.

The definitions are based upon modern biological evolutionary theory according to Dawkins forthcoming a, b) [27]. Policy processes are viewed as memetic evolutionary processes. Memes are 'pieces of data that are copied from individual to individual without too much alteration'. In fact, memes can range from simple ideas (or concepts), to complicated ways of interpretation, moral standards, conclusions from previous debates, and so forth.

In processes of policy making different views on how issues should be addressed compete for supremacy. Common for theories on policy processes is to assume that individuals and groups, usually referred to as actors, interact on the basis of different views and convictions, with a policy plan as a possible result. Problems and solutions (or elements of both) are tied to specific actors. Actors as well as problems and solutions compete for recognition in policy plans. Such policy plans, built on concepts, facts, opinions, and other memes, always contain plans to change (an) organization. No matter how different actors think, eventually policy plans are produced, even if it should only be a 'steady as you go' policy. Therefore, the study of policy processes is a good area of knowledge to study questions on how memetic evolution works, and how it is to be described.

In this paper we first introduce the framework and its conceptual building blocks In the description of lineages a mechanism for variation production (weaseling) is introduced, along with social mechanisms which can prevent this (concept-locking mechanisms).

A case study at a Dutch University is continuously used to illustrate the framework. In this case study the question is raised as to what the 'democratic qualities of the administration of the university' are. The following conclusions argue that the view established by the framework is both an elaboration on current policy theory and a connection between biological theories and theories on public policy. After the references a list of definitions is presented.

## **2. The Framework**

### **2.1 The institutional system and the case study**

In this paper we will refer to the case study 'University of Nijmegen', or KUN (after the Dutch name 'Katholieke Universiteit Nijmegen') to illustrate the framework.

After three years of debate the University of Nijmegen in the Netherlands carried through (in 1993) a major change in its regulations. The regulations were changed to what members in the University Council regarded as being 'less democratic'. The University Executive Board, however, saw this change as an important and necessary step forward to the future where 'the competition will become more difficult'.

This three-year-debate included over 80 documents which we used as data. In the end it resulted in three changes in the regulations. Two major policy proposals were put forward, but only a few of the points in these proposals were finally included in the regulations.

The policy making process at the University of Nijmegen, and in fact all such processes, do not take place in a void, but in an administrative structure which can be described by means of institutions. Institutions are rules stipulating which specific actors have the power to make decisions, to write down policy proposals, and to influence for instance decision makers with their knowledge of, or power elsewhere in the organization, etc.

These rules can be laid down in laws and regulations. Nevertheless, such rules are often only part of the organizational culture: they are not written down at all.

Policy plans are formulated and approved of in such an institutional system. The institutions determine where some of the conceptual building blocks of the framework are situated in a policy-making process. Therefore we will briefly go into the institutional structure of the University of Nijmegen.

We will restrict ourselves to examples concerned with the central institutional structure relevant to the changes in the university regulations (in Dutch 'structuurreglement') made in 1993. The University of Nijmegen has three central organs: the 'KUN Foundation', the University Executive Board and the University Council (from now on referred to as Foundation, Board and Council). The Foundation is the legal owner of the land and buildings, appoints the members of the Board, and can set new regulations. The members of the Foundation are appointed by the convention of bishops of the province of the Dutch Roman Catholic Church. The Board is appointed by the Foundation, except for the Rector who is appointed by the Council of Deans (chairmen of faculty<sup>2</sup> boards). The Council is elected by and from students and personnel. Members of the Board are not appointed without 'hearing' the Council and the other members of the Board. Before changing the 'structure regulations'<sup>3</sup> (from now on referred to as 'regulations') the Foundation must 'hear' the Board and Council. We will expand on the history of the democratic institutions of Dutch universities later on in section 2.5.

The institutional system so far had to do with formal written rules. A good example of cultural rules was the behaviour of the Foundation. It did not impose new regulations, but left it to the Board and Council to make the decision to change the regulations if necessary, and if so, to formulate a proposal. After the Foundation posed the first questions whether and how the regulations should be changed, the Board and Council needed three years to formulate a proposal that gained support from both.

Besides the decision makers, there were also other relevant actors: for instance one of the two major policy proposals was formulated by a committee appointed by the Board, the other one was formulated by the civil servants of the 'Bureau of the University'.

## **2.2 Concepts for policy making**

In institutional systems, as described for our case study above, policy plans are formulated from time to time. To consistently describe this process, we have need to introduce a number of concepts. These are memes, replication and retention systems; interaction, arenas and their selective environment; and lineages.

Policy making always takes place in (a group of) arenas where it is decided what memes (ideas, actions and such) end in a policy plan. In these arenas ideas come and go and compete for a place in the minds of people and, in the end, for a place in the policy plan. We will now describe the concepts more precisely.

## **2.3 Memes, replication and retention systems**

The 'meme' is one of the most central concepts in evolutionary policy processes. It is defined as a replicator: 'a piece of data that is copied from individual to individual without too much alteration' (However, we will argue in the next section that a meme can also be an interactor.) In an evolutionary framework a meme is a replicator, a term that was first used for genes (Dawkins, 1976) [8].

In human cognitive processes it is quite common that ideas, normative criteria and other entities are copied from individual to individual. When we hear an argument why a specification should be taken, we might remember it and repeat it to others, or incorporate it in a proposal for a policy plan. Ideas thus transmitted from individual to individual, or from report to report, are called replicators. The acts in which this transmission takes place are called replications. A replicator contains data or, in other words, has a 'coded structure' which stands for something else. Such replicators are called 'memes' when the replication system is the brain<sup>4</sup>. Memes may be ideas, but also ways of thought, complete models of how (parts of) our world works (theories), examples or metaphors to explain things, and so on. Just as genes, memes are replicators, be it in a very different replication environment and process. Likewise, memes may contain data which becomes information when interpreted by an interpretation environment. While the interpretation system of genes is always of the same kind (Ribosomes and so on), interpretation systems of memes usually differ ('democracy' may mean different things to different people and in different contexts).

When replicators are not replicated, they are stored in retention systems or memory. As to genes this is almost always cellular DNA, but memes can be stored in many kinds of retention systems. Books, brains, tapes, and computer memories are possible meme storages.

In the case of the University of Nijmegen several 'new' memes could be identified. They first surfaced in the second proposal for the policy plan written by the committee 'Zeevalking' (Second report Commissie Zeevalking, 1991) [6]. The committee was appointed by the Board while Zeevalking was the name of the chairman of the committee. In this report external factors are described which 'have an influence on the university'. Examples of these factors are: competitiveness and decreasing government funds (page 4 and 5). These memes, as reasons to legitimize a change in structure, were replicated without too much alteration to other actors (Foundation and Board), and were still evoked at the end of the discussion by these actors. These factors and their description were no original ideas but, partly but literally, copied from another document (Van der Zwan, 1990) [33]. The last mentioned document was the result of a conference held in December 1990 by the Ministry of Education, on the same subject. In fact this conference was one of the reasons for the Foundation to ask the Board whether a change in the regulations was necessary.

This is an example of replication. A part of one document is copied to later documents and, in this case, to another institutional domain: arguments from a national discussion were copied into a discussion on policy making for a particular 'private' university. The link between these two different discussions was fairly obvious since the national document was distributed among the members of the Council at the Council meeting in January 1991. Furthermore, important organizing concepts from the document, such as 'integral management', were part of the official assignment to the committee. These organizing concepts also replicated successfully via the field of actors, as well as over time (they were mentioned throughout the whole process).

## 2.4 Selective interaction and arenas

Selective interaction, or 's-interaction'<sup>5</sup>, is the second central concept next to replication. S-interaction processes result by definition in a direct or indirect weeding out of replicators. This always takes place in arenas. Therefore, arenas are defined as the place where s-interaction takes place at a certain moment in time. Arenas can be prescribed by the institutional system. However, different assemblies at different stages in the same kind of arena produce different s-interaction events. At these different events a different composition of the arena is probable: not all members will be present at all times. Especially if the collective actors in the arena consist of many members.

S-interaction can be illustrated by a particular difference of opinion between the Council and the Board of the University of Nijmegen. This s-interaction took place between the Council and the Board at several assemblies of the Council where the presence of the Board is required. The Board and Council had agreed to implement what they referred to as 'integral management', without damaging the 'democratic quality of the administrative structure'. However, this agreement came down to nothing when it became clear what each of the actors meant by these concepts. Firstly, the Board interpreted 'integral management' as the integration of 'policy and control' (in Dutch: 'bestuur and beheer'). In the past there had been a separation between the

management of buildings, personnel and such like matters (control), and the management of the content of research and education (policy). This separation was particularly recognizable at the departmental or faculty<sup>6</sup> level, where a director was in charge of 'control', and a board and council were in charge of 'policy'. The Board, however, not only wanted to put Faculty Boards in charge of both of these management tasks but, secondly, also wanted to restrain the authority of the Faculty Council. They proposed to abolish the right of amendment and initiative, the authority to decide on the budget, and so on. In addition they proposed to do the same at university level. The authorities of councils to decide on matters, would be brought back to the right to 'agree', together with the possibility to appeal against decisions of boards they would disagree with.

The Council was furious. They had agreed with 'integral management', meaning only to integrate the management of 'control' and 'policy' at faculty level, but they had certainly not agreed with any decrease of authorities of any council. The members of the Council felt that the board had distorted the agreement to a high degree. Besides the different opinions about the concept of 'integral management', it also meant a difference of opinion on what 'democratic' implies. The Board apparently argued that restricting the rights of the council to a right to agree did not damage the democratic qualities of the administrative structure; whereas the Council argued that the 'right of amendment and initiative' were 'fundamental' to any democratic administration.

In the end, the Board reached a compromise with the Council. The right to agree was taken up in the regulations, but not exactly as meant by the board. The right of amendment was ingeniously added so that the Council could add and change elements of proposals made by the Board before deciding whether to agree to a proposal as a whole. The Board in its turn could declare a particular addition 'inadmissible' in which case the Council decides whether to agree to the proposal if not containing the addition. This construction was claimed to be an existent one, and thus a replicated meme, and referred to as the 'vote blocque' construction. Furthermore, the Board still had to justify its actions to the Council, to which it opposed strongly. The Council lost its right of initiative, and was granted the right to advise.

In this example memes compete for a place in a policy plan. The memes that compete are the descriptions of authorities of boards and councils. However, these memes are not only replicated, but they are also in competition or, to put it differently, in s-interaction. The impossibility of both proposals (the right of amendment and initiative versus the right to agree) ending in the regulations, makes this event an s-interaction process where competition takes place: the right of initiative was weeded out. The right to agree and the right of amendment were not weeded out, since a new article was introduced to combine these rights. The right of amendment did change of course, because an amendment could be declared 'inadmissible'. The mechanism by which the competition took place was negotiation, where both sides handed in some points of view to come to an agreement.

Although memes can be weeded out by selection because they are subject of discussion and negotiation, many memes are not. Analogical to neutral mutations and 'spandrels' (Gould, 1997) [11] in biological genetic evolution, many memes in s-interaction processes are weeded out by coincidence. Most memes are never discussed, but simply make it into approved proposals

because they are linked to memes that are discussed and selected for. They are not selected for, but there is (positive) selection of them.

When a meme (a structural entity) is not replicated but in the process of s-interaction, we will call it an interactor, corresponding with the definitions of interactor and replicator by Hull (1980; 1982; 1988a; b; see also Speel, 1996b) [14].

Memes can thus be both interactors and replicators. The name they are given depends on the process they are in at a particular moment in time. So the concepts of replication and s-interaction are definitions of process. Memes are the basic structural entities. When they are replicated or stored they are replicators. When they are in s-interaction, the same structural entity, or a translated form of it, is an interactor. A definition of process can easily be understood by taking a chair as an example. When you sit on it, it can be defined as a piece of furniture. When you hit someone with it, it is more appropriately defined as a weapon. A meme like "bald is beautiful" might be transmitted from a bald man to a man with a lavishly head of hair. In this transmission the meme is a replicator. But when the man with the beautiful hair thinks about it, he might really disagree. In the act of judging the meme as untrue, the meme is an interactor.

While this standpoint may seem strange, we believe it to be well founded in the debate on species as individuals (see Ruse (1989) [24] for a series of papers relevant to this topic). Therefore, ignoring it will generate serious philosophic problems with the foundations of evolutionary theory.

These definitions of process are derived from biological evolutionary theory. There genes are replicators, which are translated into interactors such as enzymes, organisms and so forth (Brandon, 1988) [1]. The implications of this classification cannot be dealt with here, but are fundamental<sup>7</sup> to the difference between 'Lamarckian' and 'Darwinian' evolution.

S-interaction processes take place in arenas at a certain point in time. An arena can be the mind of an individual, but also the minds in collective or group decision processes, that is where people debate on the adoption of methods<sup>8</sup>. In the example of the University of Nijmegen the arena was the assembly of the Council. We speak of an 'internal' arena when s-interaction takes place in one individual. When s-interaction involves more people we speak of an 'external' arena. Accordingly we can speak of internal and external s-interaction. An internal retention system refers to (parts of) a brain and an external retention system is any retention system outside the brain, i.e. books, tapes and so on.

Summarizing the definitions of central concepts we end up with:

An arena is a place where s-interaction takes place (at a certain time) according to certain rules (institutions).

An interactor is an entity that 'interacts in an arena where this action results in the weeding out of a replicator in a retention system'.

A retention system is 'a kind of memory where replicators are stored'.

In the struggle for survival of description of authorities in our case study mentioned above, the external retention system was a single policy plan, but the decision processes took place in multiple external arenas.

We may add that arenas are always coupled to retention systems: arenas serve as a kind of filter as to what memes get into their accompanying retention system. This is shown by the above-mentioned example. The decision to choose for either the right to agree or the right of amendment and initiative was eventually taken in an external arena. The retention system the surviving meme was added to was 'the regulations'.

## **2.5 The selective environment and levels of retention**

S-interaction in an arena results in the weeding out of memes. This weeding out can take place on the basis of different kinds of 'criteria'. The sum of the factors being decisive on what memes are weeded out is called the 'selective environment'. The selective environment is what causes weeding out (Brandon, 1988) [1].

The selective environment may consist of more or less formal criteria to weigh options. It might be decided that the authorities of councils should remain the same because 'changing them will decrease the democratic quality of the structure', where the criterion is, of course, 'keeping the democratic quality the same'. Dunn (1993) [10] describes a classification of different criteria people use in policy processes.

The selective environment may also be based on other kinds of factors, for instance the relative powers of decision makers. If a decision maker can decide what he wants, without taking into account the opinions of others, his power will suffice for the decisions to be taken. The Foundation, for example, has the power to set and change the regulations of the University of Nijmegen without taking into account the opinions of Board and Council. The memes endorsed by the Foundation would thus win without contestation on content. Of course there are supplementary causes to this explanation. While the Foundation might do what it wanted without listening to Council and Board, we would still need to explain why it made decisions in the specific way it did. It may be possible that memetic s-interaction has taken place somewhere in the Foundation as well.

Memes that, through selection processes in arenas, make it into retention systems, may have a different status. A meme in a policy plan can, for instance, describe an action to be taken, or just be a description of something that needs to be changed. The policy plan (regulations) finally set in our example of the University of Nijmegen, included the decision to implement 'integral management', intending to be a compromise between that of the Council and that of the Board. But the decision to change the regulations also included arguments explaining why a change was needed. The decision implies memes referring to actions to be taken, whereas the other memes were just a part of the arguments to rationalize the change. The status of both kinds of memes is clearly different. We will refer to this difference in status as the levels of retention.

We distinguish four levels, to be referred to a 'levels of retention', in the 'status' a meme can have in a retention system:

- to have knowledge of a meme;
- to judge a meme to be relevant for a discussion;
- to endorse a meme; and
- to translate a meme into action.

An actor may have knowledge of a meme. One of the proposed policy actions can be to 'abolish authorities of boards'. However, a member of a council who thinks that decisions get better if councils can amend them, may think this not to be a good idea. He may fear that decisions will be taken which turn out to be unworkable. Therefore he may judge it to be unwanted, and talk to his fellow council members, explaining this to be a bad idea.

Another council member may hear of the proposed action, but may judge it favourably. He then knows the meme, thinks it to be relevant for the discussion, and endorses it. He will probably also defend it when the truth or usefulness of the meme is doubted by others.

The fourth level is a more special level, since not all memes can be translated into action. Only prescriptive memes can be translated into action. When an individual endorses a meme that is a prescription he will act on it. Religions hold prescriptive memes of course, but so do policy plans. However, in policy plans it is usually the case that the individuals that have to execute the memes (policy implementation) are not the same as those who have the competence to approve on it.

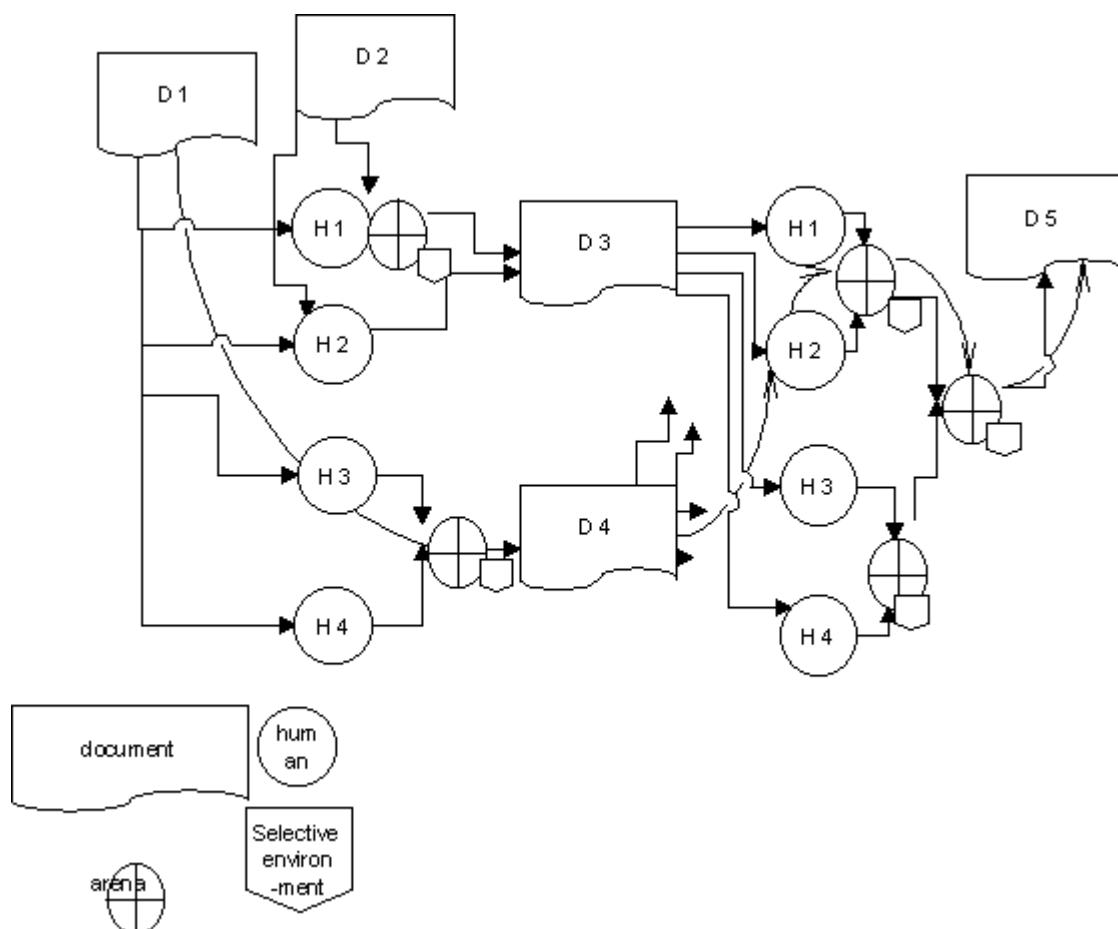
The first three retention levels can be applied to memes in all retention systems. The fourth retention level is a special case.

A failure to differentiate in different levels of retention leads to confusion. Especially since more levels are already being used in the memetic literature. Dawkins (1976) [8], when referring to religions as meme complexes, uses endorsement and applications of religious rules and beliefs as success criteria. Best (1997) [2], for instance uses replicative success (level one and two) as success criteria. Vaneechoutte (1993) [29] uses replicative success in the phrase 'successfully distributed - horizontally - into different brains', in the context of religions as meme complexes.

A failure to differentiate in different levels of retention can also obscure important insights. Many memes are very successful in spreading, exactly because they are not endorsed. There are, for instance, many discussions and papers about Lamarckian mechanisms in genetic evolution. In many discussions this meme is referred to in order to state the point that Lamarckian genetic evolution does not take place. The Weismannian barrier is often used to state the same point as well. Papers and books mention Lamarckian evolution with regard to the history of the field of biological evolutionary theory. However, in the context of cultural or memetic evolution Lamarckian evolution is mentioned to emphasise the difference between genetic and memetic evolution, and it is endorsed that cultural evolution is Lamarckian (see Gould, 1997) [11]. Taken together, Lamarckian evolution is a successful meme, but the factors for its success are mixed, involving endorsement and relevance to a discussion (because it is wrong) as the levels of retention.

In memetic theory a successful meme is often taken to be a meme that is 'true' (Darwinian theories are more successful than Lamarckian theories in genetic evolution) or useful, to others with what spreads well. Needless to say that these two different criteria for success will yield very different results in memetic analysis, both for the dynamics, the mechanisms, and the selective forces involved. It is probable that memes that spread because they are not true, or offensive for instance, spread because they evoke a reaction from people opposed to the message in the memes. These people will replicate the meme while stating it is untrue or offensive. Calling someone a 'Nazi' for instance on the internet will replicate that meme, but the emotional mechanisms involved will be different from when you argue that it is a shame that Princess Diana has deceased.

Diagram 1 shows a stylized outline of the framework where two collective actors interact in a policy-making process. Human H1 and H2 form one actor, and H3 and H4 the other one. In this process an old policy plan (D1 for document 1) is changed into a new one (D2). The memes in the documents are replicated to the internal retention systems. The lines stand for replication. Before external retention systems are produced (documents), s-interaction in arenas takes place, with its selective environments.



**Diagram 1.** A stylized representation of the framework. This diagram shows an imaginary case, with arenas, retention systems, actors and a lineage. Time runs from left to right.

The diagram is but a model of reality and, for the sake of clarity, not all details were depicted. For instance internal arenas can always be present with all humans in the diagram. Only the first arena on the left upper side is an internal arena and should be in the human. The second arena left below is an external arena. Also there are only single lines from humans to arenas. There could be a lot more lines, if the depicted external s-interaction is accompanied with internal s-interaction in the humans' internal arenas<sup>5</sup>. The third arena, in the upper middle, is followed by the last one without showing intermediate human replication. If the diagram should be exact, it would show humans after every external arena, since humans replicate memes.

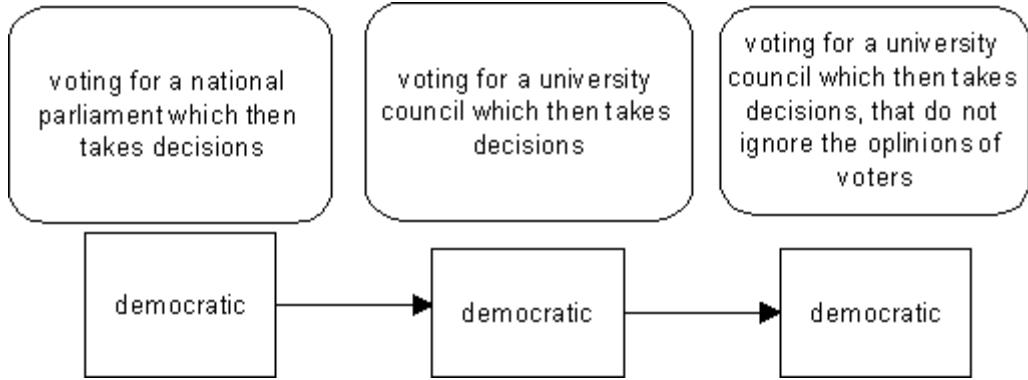
D2 stands for documents and other memes that are used, and which come from outside the organization. In the middle, two propositional documents are produced by actors consisting of H1 and H2 (say the board), and H2 and H3, (say a council). Collective actors reconsider their endorsed memes after reading each other's documents and in the final arena the final policy plan D5 is agreed upon. The line going from D1 to H3, D4, H2, D5 is an example of a lineage that runs through the process.

## 2.6 Lineage

In this framework we recognize lineages of memes (the line in diagram 1 going from the first to the last document). A lineage is a logical result of the concept replicator. Genes and memes, or in general replicators, are repeatedly copied and thereby, over time, form a lineage. A lineage is thus a line of replicators connected by heritage, i.e. a (changing) historical entity. In the framework lineages 'linger' through retention systems by replication, possibly changing.

Replicators in lineages can change or mutate in many ways in memetic or conceptual evolution, yielding variation (see Price (1995) [23] for a possible mechanism). Hull (1988a) [16], for instance, refers to 'weasel words' as words that change in meaning, whereas the actual word remains the same. To put it more precisely, the change in meaning takes place in a lineage, so we will refer to this as a weasel lineage. The discipline that studies such changes is called etymology. Take, for example, a word that is replicated inside a mind. Technically this is not a lineage, since memes are replicated from individual to individual. Therefore we shall call such a lineage an 'internal lineage'.

Suppose an individual hears of the word 'democracy'. He might take up this meme from others, with the explanation that it means 'voting for a parliament, which then takes decisions'. Suppose the parliament in the example he heard were a national parliament. We can call the description with the example of a national parliament the 'context' of the word, which gives the word its meaning. Suppose this individual then contacts a Dutch University and its administration in 1996. Here, there is then another kind of democracy, where an elected council takes decisions. The people involved tend to call this also democratic. The person then connects the word democracy he already knew from the national parliament to the new context of the council in the university. On top of that he might learn that the council, once elected, can ignore the opinions of the voters and that its behaviour, and therefore the council itself, is called 'undemocratic'. This situation is illustrated in diagram 2.



**Diagram 2.** A stylized representation of the weasel lineage 'democracy' that changes in meaning.

What we see here is an internal (in one brain, or mind of an individual) memetic lineage that has changed over time by influences from outside. The meaning changed owing to connecting new contexts to the same word. In biology we would say that the lineage has mutated. In this example, democracy is thus a 'weasel lineage'. If the change in meaning means that little of the original meaning resides (according to some standard), we can either say that the lineage has split or that it has ended. In the first approach the word democracy is regarded as a lineage that remains the same, but the meaning attached to it changes; it is linked with different other lineages of concepts. In the second approach the word democracy and its meaning are regarded as one meme, and the change equals a mutation or new variation: a new meme is born.

This weaseling is even more likely to happen when memes are replicated from individual to individual. When the word 'democracy' is replicated from a member of a national parliament to a member of a university council, it is bound to change in meaning. The individual, who takes up the meme, will attach his own meaning to the word, depending on what he uses it for, and what he already knows about it. Memes that are thus replicated will often change in meaning. However, together with a meme, its interpretation can be replicated simultaneously. So by communication on the meaning of a meme, individuals can 'level' their interpretations.

While it is shown that a meme can have a very different meaning for different individuals, it still counts as the same meme as long as it is replicated directly. The change in meaning is irrelevant to the definition of what counts as a meme. The only decisive criterion for a meme in the same lineage is heritage.

In the case of the University of Nijmegen such 'weaseling' can also be recognized. The Board in one of the proposals for a policy plan calls the 'structure' of administration democratic when there is 'a structured influence from representatives of the university community on the administration'. The Council interpreted a democratic structure as a structure in which the council has the 'right to submit amendments and propositions together with the right to decide on the budget'. This example shows that the same concept meant quite different things for different actors.

Once you think about it, it turns out that almost all words we use which are conceptual (whereas words like 'the' or 'it' are quite stable), and which have meanings in the sense mentioned above 'weasel about' quite a lot. In fact, when such weaseling is not wanted, it takes specific precautions to prevent its happening.

There are several stabilizing mechanisms in social processes to avoid weaseling of words or concepts. In science there are formal theories characterized by formal, and thus fixed, definitions.

In law there is jurisprudence which fixes the meanings of concepts by means of frequent interpretation of examples in cases. Jurisprudence acts as a 'lock' on definitions. Once a concept has been given a specific meaning by a judge, or court, other courts and judges are required to use the same interpretation. If courts come into conflict a higher court must resolve this and re-establish a uniform interpretation. We will refer to the stabilization of meaning as 'locking' a concept (or meme).

While weasel lineages are important for the creativity of language, they make a clear understanding in communication impossible where people need to concentrate on what they actually and precisely mean. In policy making this is of course very important, if we do not want to end up in endless misunderstanding.

Actors in policy situations that who to do something different from what they have agreed upon, can become very inventive with weaseling. In negotiation it is a common feature to make agreements which can be interpreted (connected to meaning) in many different ways, so that all partners in communication can agree. It is for instance an open question if the Board and Council would have agreed on 'democratic' and 'integral management' as starting points for a policy change if they could have foreseen that they would disagree later on the meaning of the words.

The examples of 'democratic' as an internal lineage and as a subject of debate and confusion within the debates at the University of Nijmegen is also relevant to the evolution of the 'democratic' institutions in Dutch universities in the past forty years. Before 1960 Dutch universities were governed by the Ministry of Education and by their own internal structures.

The internal structure was dualistic in that the university was governed by a Board of Governors (in Dutch: 'College van Curatoren'), but academic policy matters were decided upon by academics: Chairs (professors) at the basic level, the Faculty as intermediary and the Senate at the central level. The Faculty as well as the Senate were formed by all professors involved. In an act, seen as the first modern regulation of the Dutch universities in 1960, the University Education Act ('Wet op het Wetenschappelijk Onderwijs'; W.W.O), universities were given more autonomy (they were given an autonomous legal status), and the competencies of the Faculty were increased as opposed to the competences of Chairs. The Faculty became responsible for the coordination of exams and education. The universities were also financed differently in that they were no longer funded on the basis of expense account, but on the basis of centrally allocated budgets based on information given by the universities and connected to a formal planning system. The Chairs, however, were still appointed by the national government, thus limiting the autonomy of the universities. The special universities, such as the University of Nijmegen

mentioned above, were also included in the allocation of money, be it that they were only given 95 percent compared to state universities until 1970.

The basis of the 'democratic' institutions was first laid down in 1970 in the act 'Wet Universitaire Bestuurshervorming' (WUB; University Government Reform Act). It included the Council structure and also a list of explicit goals including 'democracy' at universities. This law was passed as a result of student and academic protest against a national policy plan, formulated because the university structure could not handle the growth of the universities in the sixties. The policy plan proposed a Presidium of three individuals having not only the competences formerly held by the governors, but also the academic. The plan catalyzed a strong protest movement in the opposite direction of the concentration of competences toward democratic structures.

The WUB was passed as an act in 1970 and would eventually last fourteen years, until 1984 (These data are based on Groot; 1988) [13]. The democratic structures, however, lasted until recently: in 1996 the 'Modernisation University Governmental Organisation Act' was passed ('Modernisering Universitaire Bestuursorganisatie' (MUB)) which reduces and changes the influence of councils.

A memetic empirical question here is where the 'democratic' memes in the WUB came from. Some say they were derived from the Dutch laws on the government of municipalities, but others deny this. This question might be answered by a study on who actually wrote the proposals for the WUB and what sources he or she used. Another more objective method would be to analyze the WUB, and proposals that preceded it, on specific word phrases to see if they fit words phrases in the law on municipalities.

However, it is obvious that since the WUB was introduced in 1970 the democratic memes (as well as their formal interpretations resulting in regulations of individual Dutch universities) have been around in the formal body of law in the Netherlands, and that they can be regarded as one lineage or perhaps a network of lineages.

It may be nice to observe that the University of Nijmegen was experimenting with democratic structures before the WUB was passed as an act, and that the changes in the regulations described above were also prior to the MUB. Since special universities in the Netherlands were allowed to have different regulations (and state universities were not, although this has changed somewhat since the MUB), they can change their regulations based on their own interpretations and goals. In other words locking mechanism of legal memes that have been present for state university regulations (they were to be approved of by the minister) were not operating as vigorously on the regulations of special universities. The results of both policy processes, the national one and the one at the University of Nijmegen, are also quite different. The board of the University of Nijmegen however, is already planning to change the regulations again toward the present situation in state universities.

So far we have been talking about conceptual lineages. However, to be complete we must mention that Hull (1988a) [16] discerns two kinds of lineages: social and conceptual ones. Conceptual-, or meme lineages are lineages of ideas, theories and the like as we have explained above.

A social lineage is a group of people that can lose and gain members, that has the same name or is bound together by a common interest, task, and etcetera. Social lineages are for instance actors in policy and in political theories. They are called lineages because they are bound together by a historical line that can be traced. While the membership of such a group does change, and possibly also the name and/or the ideology, we still see such groups as one and the same. Thus their identity is essentially based on historical lines.

Before completing this part of the paper, in order to avoid confusion, we need to mention that 'lineage' in biology can have two meanings. First of all 'lineage' may mean a replicated line of genes. In this case a gene stands for a complex of pieces of DNA that code for one protein which, in such a lineage, may change by mutation, including recombination.

The other meaning is that of a species (or a sexually reproducing 'lineage') which can also be called a lineage or a 'whole' that changes genetically over time. It is this type of lineage that is commonly said to evolve by natural selection. While a lineage of genes may mutate, a species as 'lineage' can adapt. In this paper we will refer only to the first meaning of the word lineage. The second meaning for systems that can adapt or learn should be referred to as a 'selective retention system' (Campbell, 1974) [5] or perhaps an evolver<sup>9</sup>.

### **3. Implications of the Framework**

The framework has implications for a) public policy theory and b) memetic theory.

#### **3.1 Implications for public policy theory**

There are five main implications for public policy theory.

Firstly, the framework describes the cognitive side of policy making according to the modern evolutionary paradigm. Therefore it can be used to 'translate' understandings from different disciplines of science that were separated before. The understanding acquired in different disciplines is more readily usable and this can yield new theoretical combinations. More specifically, the elaborated 'toolkit' of modern evolutionary concepts in biology can now be used in disciplines where it was not used before. This toolkit has many benefits, among which the dynamic characteristics, in contrast to, for instance, the approach currently used in institutional economics (Williamson, 1985) [32]. Furthermore it can describe well variation production or, to put it differently, innovations (which has enriched economic theory as well), and brings along population thought as an elaboration to platonic essentialism<sup>10</sup>.

Secondly, the framework focuses on cognitive processes in policy making. Many policy theories describe policy making as being mainly a political process, where actors interact to increase or keep their power or income, and to defend their ideology or points of view. Such theories refer to the political or power-driven rationality (often only to be established after actions have been taken) by which actions of actors can be described.

This framework can be used to emphasise that the cognitive processes by which choices are made are not only driven by political factors, but also by the development of the understanding

of the content of problems and solutions. Where many policy theories do acknowledge that this understanding is a factor, they do not have a model to describe how it develops. However, if we want policy processes to result in the solutions of problems, this understanding is crucial. In other words, the framework focuses on interpretative institutions, where most policy theories focus on interactive institutions (Campbell, 1995) [4]. Interactive institutions are about the interactive behaviour of actors with regard to the political power. Interpretative institutions are about the models with which problems, solutions and so on are described. So our framework can serve as an elaboration on theories on political interaction between actors.

Thirdly, rational, non-rational and other factors in selective environments can be viewed together in one framework. In the real social world both emotional factors and rational deliberations on what to do play a role. Emotional factors include group identity, non-rational preferences of individuals for specific actions over others. Furthermore, factors as the 'catchiness' of concepts, persuasion caused by an intimidating personality of an individual, for instance, and so on may play a role. From the specific memetic side of theory we can add that meme complexes may have become adapted to be endorsed and dispersed successfully, resulting in memes that are hard to resist.

Some actions can simply be popular at a time, while nobody can exactly tell why they should be taken. Some actions might be adopted because respected people say they should be taken, or because they are brought forward in a persuasive manner. Lynch (1996) [19] gives the example of memes that are selectively replicated among different social classes in society. Fashionable memes in a 'higher' class are often resistant to memes in lower classes. It may be that actors (or individuals) who see themselves as members of a higher class will resist memes that are proposed by actors in a lower class. This can be generalized by saying that members of a group with some identity might resist memes coming from other groups. Famous examples of course can be found in struggles between employers and employees, or between memes of environmental activists and entrepreneurs. Some actions that entrepreneurs can take to improve the environment are good for profit as well. Still such actions can be resisted, simply because they are proposed by environmental activists. Of course the same kind of processes play a part in policy making elsewhere, including public policy.

The concept of the selective environment enables us to group these factors into one concept. This may seem trivial, but many disciplines only consider either rational or non-rational factors without integrating them. In policy sciences rational deliberation is often seen as the basic explanation for decisions.

Fourthly, through the concept of lineage, the framework implies a (specific) historical view on deliberation. The models used to describe problems and solutions are historical entities, either made up when needed (new variation) or copied from already existing memetic sources. This puts the view that organizations are under the rational control of humans even more in perspective than the concept of 'bounded rationality' implies.

In a policy process descriptions of problems and solutions or, in other words, goals and actions can be copied from other institutional systems or from previous discussions in the same institutional system. For instance, the description of the problems of the University of Nijmegen

were copied from the national discussion on changes needed for public universities: a different institutional system. Another possibility is that descriptions are already available from past discussions in the same institutional system. When descriptions of problems and solutions, or conclusions from discussions (s-interaction) from the past are copied without deliberation whether they are correct, we might speak of 'conditioned', or rule based decision making. Memes from the past are copied when current events seem similar to previous events, or events experienced in other institutional systems. Thus the similarity of events leads to the appliance of previously developed memes.

As an example of theories that do not account for the cognitive development of concepts and models, we might refer to 'bounded rationality'. The concept 'bounded rationality', or satisficing behaviour, has become accepted when referring to decision-making. This concept implies that the selective environment for a decision (the problem to be solved) is already clear, but that there is a limit to the time and effort people will spend in finding the solution. Our memetic lineage concept elaborates on this in two ways:

The definition of the problem itself can be seen as part of, or the result of an s-interaction process. But more important is that the process by which solutions are searched for is also built on definitions, conclusions and concepts acquired in the past. These historically formed 'conceptual building blocks' are replicated in present time s-interaction processes. The descriptions of problems and solutions actually used, can be far from optimal. They may be ill-founded when conclusions reached in the past are incorrect. This may also include the complete frame in which problems and solutions are seen.

The search for the models in which problems and solutions can be described, as well as the application of such models in a new situation, takes time and effort. The restriction of the time and effort people are willing to invest<sup>11</sup> in finding the right models, must be distinguished from the restricted effort to find solutions when the problems are already clear. Only this last restriction is what is referred to in 'bounded rationality' (Simon, 1976) [26].

Thus rationality is not only bounded by the effort to find solutions, it is also bounded by the abundance of seemingly fitting models, concepts and conclusions already available. When models are available that 'fit' the problems to a certain extent, it still takes effort to 'tinker' (in the sense of Jacob, 1977) [18] these, so that they fit a new situation. The 'tools' of deliberation are tinkered from old memes to possibly new and better fitting variations of these memes.

While we can use the concept of bounded rationality to account for the cost of finding solutions, bounded rationality will expand in meaning if we include the cost for the tinkering of models to view solutions and problems with, and the solutions and problems themselves. Furthermore, bounded rationality of course does not describe mutating memetic lineages.

Fifthly, the framework can describe lineages penetrating different policy organizations, policy themes and retention systems. This description makes explicit what everybody knows: people, and also policy actors, only create very few new descriptions, and solutions. They are often copied from what is already known elsewhere. The memetic view contributes to the study of how this happens and, for instance, why some solutions are so popular at certain times.

Once fashionable opinions and plans are recognized, it can be asked whether these fashions make any sense. Why do all universities in the Netherlands follow each other when changing their rules? To us, this epidemic characteristic of meme dissemination is particularly fascinating. The grouping of explanatory factors in the selective environment enables us to come up with hypotheses on what factors were important in specific processes. The possibilities we have sketched as factors are irrationality, rationality (political and with regard to the description of problems and solutions) and conditional decision making.

### **3.2 Implications for memetic theory**

There are three main implications for memetic theory.

Firstly, if we view organizational processes in the memetic historical way as explained above, it becomes clear that even when organizations are managed and controlled by humans that behave with bounded rationality, their actions and deliberations are guided by already existing lineages of memes. The few choices bounded rational actors make, are thus partly under the control of memes already available. As Price (1995) [23] states: 'the organization can be seen as the product of the meme'.

This realization leads to the hypothesis that a considerable part of human memetic processes is replicative, as opposed to processes where deduction and deliberation yield new memetic variation. One can ask how much of human memetic evolution is replicative, and how much new memetic variation is involved? To investigate such a thesis we must, of course, first define what counts as new variation in practice.

In a policy context we can ask what processes and elements of institutional structures lead to more memetic variation? Price (1995) [23] for instance holds that a kind of isolation mechanism of meme complexes can yield new memetic variation. It would be a challenge to define and investigate such hypotheses.

Secondly, the issue whether memetic evolution is different from genetic evolution has generated some debate (see Gould, 1997 for instance) [11]. We agree with Hull (1982) [15] that many proposed differences are simply misunderstandings, among which the claims that memetic evolution is faster, Lamarckian or that only memetic evolution shows horizontal transmission of replicators. Instead we propose a new difference: Interpretation mechanisms in genetic evolution are almost always similar, contrasted by memetic interpretation mechanisms that almost always differ. This results in a different mechanism for variation production in memetic evolution: weaseling.

Finally, in examples used in memetic criteria, different levels of retention are used. We have argued that these levels should be distinguished properly to avoid confusion.

## **4. Conclusions**

Memetic theory has been applied in the philosophy of science (Hull, 1988a) [16], in models of consciousness (Dennett, 1991) [9], or for describing religions (Dawkins, 1976) [8]. If these

applications have been useful, we have shown that memetic theory can also be usefully applied for policy science. Our framework describes policy-making processes as memetic selective processes, using definitions derived from modern evolutionary theory.

We have presented some observations that are elaborations on current memetic theory. We have shown that memetic lineages of conceptual words may appear to replicate without too much alteration, but that their interpretation can vary at the same time. The mechanisms and forces that produce this variation or weaseling are probably important forces in memetic evolution in public policy since special social 'locking' structures exist when such weaseling is unwanted. We have also shown that memetic evolution theory contains different levels of retention. Without these levels in success-criteria memetic analysis will be less explanatory.

Apart from these observations the framework contributes to theoretical innovation by connecting different fields of theory: policy theory and evolutionary biology.

Furthermore, we have argued that the framework contributes to policy theory in two ways. The use of the concept 'selective environment' makes the framework suitable to view rational, irrational and other factors as competitive or complementary explanations (selection forces) for decisions taken. A solution may simply have a good ring to it or be popular in a certain social group. It may be taken in a rush, or pass unnoticed because it is linked up with more controversial issues (in which case it is neutral). Most other theories presuppose only one of these forces to be relevant.

Realizing that many forces can play a part in decisions is, of course, an implicit criticism of the view that decisions are taken by rational deliberation only. It results in the view that human planning and anticipation can have very little influence on human organizing processes. Thus memetic evolution is less different from genetic evolution as may intuitively be assumed, since one of the differences between biological and memetic evolution is often taken to be anticipation and intentionality.

## Acknowledgements

I am grateful to Jan Koolhaas, Tineke van der Ploeg, Inez Groothengel, Martin de Jong, Ronald Oosting and others for their help with the language and conceptual issues. Because I consider their help as an integral part of the conception of the framework I have used 'we' when referring to the author. Furthermore I thank the reviewers for their efforts which have improved this paper substantially.

## Glossary

### **Actors:**

Organs or agencies, individuals or groups which play a part in the policy-making process. Different parts are writing reports and interacting (including taking decisions).

**Arena:**

The place where interaction takes place at a certain time, possibly according to specific rules set by institutions.

**Bounded rationality:**

Actors are rational but their time, resources and knowledge to make and base decisions on is limited; so they stop deliberating on how to act at some point in time. Thus they are not perfectly rational. Actors that show bounded rationality are satisficing instead of optimizing. They stop searching for solutions when a solution they find or come up with meets their standards. They do not infinitely search for solutions comparing them in order to see what solutions are to be preferred (optimizing).

**Conditional decision making:**

(Also known as rule-based decision making (March, 1994) [20] the process of taking a certain decision, simply because it has been chosen in comparable situations in the past. The 'stimuli' of a certain problem leads to a reflex or routine of picking a certain predefined solution.

**Institutions:**

Rules according to which specific actors or individuals behave; in other words, stable patterns of behaviour.

**Institutional system:**

The rules of behaviour according to which actors act in arenas and by which some retention systems are defined. Formal and informal institutions include the definitions of some arenas and the part actors play there.

**Interactor:**

An entity that interacts in an arena, where this action results in differential perpetuation of memes into a retention system.

**Internal lineage:**

A lineage of memes which are replicated inside one mind.

**Levels of retention:**

The different status which memes in retention systems can have. Some imply actions that will be taken; others can be descriptions, arguments, and other memes.

**Locking a concept (or meme):**

The stabilization of meaning of a concept, or leveling the interpretation of it.

**Lineage:**

A line of replicators connected by heritage.

**Memes:**

Pieces of data that are a) copied from individual to individual without too much alteration, or B) that are interactors.

**Policy making process:**

- A process in which individuals, groups and/or actors interact on the basis of different views and convictions, with a policy plan as a result. Problems and solutions amongst other elements, often tied to specific actors, compete for recognition in policy plans. A policy plan can be a written document, containing plans to change (an) organization, or it can be an execution of changes made in a short period of time which are perceived to be related to one policy question or theme.
- A process in which an administration, or other actors and/or individuals think about, formulate, communicate and decide on plans to change things in society or in the organization itself.

**Rational:**

The use of deliberation or reasoning in making choices. As opposed to irrational choice (where affectionate and emotional factors are causes of decisions) and conditioned or rule based choice.

**Replication:**

The process where replicators are copied from one retention system to another.

**Replicator:**

A piece of data that is copied from retention system to retention system without too much alteration

**Retention system:**

A kind of memory where replicators are stored.

**S-interaction:**

the process that results in the weeding out of replicators, be it directly or not.

**Internal s-interaction:**

S-interaction where the arena is the mind of one individual.

**External s-interaction:**

S-interaction where the arena consists of more than one mind.

**Selective Environment:**

The sum of the factors, or in other words causes, decisive on what memes are weeded out in a selective event or in a number of selective events.

**Selective retention system:**

A system in which replicators, having coded for winning interactors that interacted on the basis of a common selective environment, are retained.

**Weasel lineages:**

Lineages which consist of replicated words that change in meaning, whereas the actual word remains the same.

**Weismannian barrier:**

The barrier between behaviour acquired in the lifetime of a biological organism and its genes. This barrier is referred to in order to show that the genes of an organism cannot take up such learned behavior. Genes and gene complexes only change by random variation and recombination. Natural selection takes care of deleting less 'fitting' variation, possibly resulting in adaptations.

**Notes**

## Title

Parts of this paper were previously published in a slightly different form as Speel, H.C. (1997) A Memetic Framework for the Analysis of Policy Formation. In: European Institute for Advanced Studies in Management. Organizing in a Multivoiced World. Social construction, Innovation and Organizational Change. Leuven Belgium, June 4-6.

## Note 1

Rationality, irrationality and bounded rationality are concepts that are accepted in public policy theories, but which have to be defined because they are used in many different ways. Rational here is used to refer to the use of deliberation or reasoning in making choices. Following Simon (1964) (25), choice is irrational when it is dominated by affective mechanisms such as emotions and instincts. Bounded rationality is used in the original meaning Simon (1974) (26) gave it: actors are rational but their time, resources and knowledge to make and base decisions on is limited, so they stop deliberating on what to do at some point in time. So this does not lead to perfectly rational based actions.

## Note 2

Dutch universities consist of a central level and faculties. The faculties are the organizational parts where research and education actually take place. American universities often refer to this as 'divisions'. The word thus refers to a part of the organization, and not to what Americans understand by faculty: all personnel that teach or do research. These Dutch faculties (until 1997) always have a board and a council.

## Note 3

The regulations of the University of Nijmegen are made up of three hierarchical levels of regulations: the articles of association, the structure regulations, and other regulations. These levels differ in the procedure by which they can be changed. The higher the regulation, the higher the actor needed to change it. In the case study the structural regulations were changed. In these regulations the administrative structure below the level of the Foundation is defined.

#### Note 4

To put it clearly, we would have to say that replicators are memes if their lineage has in the past been copied by brains. Here we assume that behavioural bodily patterns, also being memes, are also replicated by the brain.

#### Note 5

Hull (1980) (14) defines interaction as the process that results in the weeding out of replicators. However, in order to avoid confusion, we will refer to the same concept as selective interaction or s-interaction. Interaction in policy theory often means something like 'actors or individuals that take actions with regard to actions of others'. A policy scientist will call two individuals talking to each other interaction, while this is not what Hull is referring to when using interaction as a concept.

#### Note 6

See previous footnotes.

#### Note 7

The distinction between replicators and interactors is connected with the division of what Pattee (1977) (22) refers to as code and matter. Williams (1992) (31) speaks of the 'codical domain' when referring to code. Code can be replicated but, what's more important, can also be interpreted. The interpretation takes place by means of an interpretation system, and results in matter. The difference between matter and code is that code is in a 'rate-independent dynamical mode', whereas matter is in a 'rate-dependent dynamical mode'. In the example of DNA as code, rate-independence refers to the timelessness of the code. In transducing the code into matter, a time component is added, when the sequence of code is translated. Rate-dependent matter then has a form which interacts with other material in the cell where time is an important factor in the chemical processes of interaction. To code time is not important: the data are stored without time being an important factor to the code. The use of 'matter' here is, of course, no claim that code does not exist as matter as well. It simply states that the data in the code is not directly relevant to interaction processes. The data in the code becomes directly relevant to processes only after interpretation.

When a structural entity crosses this matter-code barrier, it is 'transduced' (Pattee, 1977) (22). In biology the interpretation system is the cell mechanism, which translates (while also transducing) specific parts of DNA at a certain time into enzymes. This theoretical theme is of course related to the genotype-phenotype distinction, since genotype is code (Calling phenotype only matter however would appear to leave behavioural parts of the phenotype out. However, it is clear that phenotype is no code.) Code has no meaning when separated from a translation system. Pattee (1977) (22) describes for instance that the biochemical mechanisms of the cell give DNA its meaning. The ability to discern between code and matter is for instance of importance to understand the Weismannian barrier. Without this barrier, the difference between 'Lamarckian' and 'Darwinian' evolution cannot be maintained.

In memetics we can also discern translations and transductions. They always take place in the brain. However, we do not have an accepted theory of how this takes place in the brain. For instance: are 'thoughts' code or matter? Are memories in the long-term and short-term memory code or matter? I have no answer to this. Does the claim that a memetic Weismannian barrier in the brain does not play a part in memetic evolution mean that there is no barrier, or that it is crossed?

The inability to describe what things in the brain are replicators or interactors poses difficulties. For instance, I do not know how 'thoughts' are connected to the material of the brain. This also means we don't know if interactors and replicators in the brain are separated as code and matter as they generally seem to be in genetic evolution. Replicators in biology are almost always code and interactors matter. But can interactors also be code, and are they in the brain? Calvin (1997) (3) refers to experiments on "RNA evolution" that do not involve a genotype-phenotype distinction. This implies that interactors can have codical characteristics while they function as matter.

We feel that in the end the part an entity plays in a process should be predominant in solving the question if an entity is a replicator or an interactor. This implies that some entities that are replicators can also be interactors (the RNA for instance). Interactors thus may have 'codical' characteristics.

This may result in the conclusion that a memetic interactor in the brain has features of code. Whether this is correct remains unknown until we have a proven transduction theory for memes in the brain.

Whatever the problems with memes as code or matter, it is clear that memes in policy plans can be translated in two kinds of interpretation-steps by humans. First a meaning is attached to a meme (as in the weaseling described in this paper), and second this meaning can be one of the factors that causes specific behaviour of humans. In organizational and policy theories the relevant concept describing the translation of described actions into actual actions is implementation of policy. Some policies, such as laws, have well -organized interpretation systems, such as courts. Other policies are never transduced into actions at all.

#### Note 8

When a group of humans is debating, s-interaction takes simultaneously place in multiple arenas: in the mind of one or more individuals (internal arenas), and possibly also in an external arena. In such a case, the question arises whether the humans or memes, or both, are the interactors. In all cases memes are of course the replicators.

The question what entities should be called interactors in an internal arena is easily to be answered. Inside an internal arena memes are always the interactors. In an external arena the answer is more difficult. Do not individuals s-interact just as memes?

The answer is that individuals are most of the time the interactors. Only when multiple individuals (or groups that speak with one voice) make a choice on the basis of a common view and thus criterion, is it that memes are the interactors in the external arena.

When a choice is made on the basis of one or more shared criteria, all individuals communicating in the debate should come to the same conclusion. We can safely state this to be an exceptional event. This is only the fact when the criterion and the interacting memes are interpreted in one way by all individuals. This can probably only happen in environments where strong 'locks' on weaseling are executed. It is far more likely that individuals 'think' they agree on interpretations, but that in fact they all make different deliberations, and yet do agree with the outcome.

When actors do not agree on interpretations, other factors will influence the outcome. For instance, the power of one group over others.

If memetic criteria are decisive in a decision process, s-interaction does not only take place in the external arena, but also in the minds of the individuals involved. Memes in internal arenas are interacting simultaneously with memes in external arenas. While it appears that a group is deciding, individuals decide themselves too, and 'level' (communicate and agree with) their decision with each other. Of course it is possible that not all members do agree, and that in the mind of disagreeing members a different s-interaction will have taken place.

#### Note 9

Consistently with the definitions of our framework a selective retention system is 'a system in which replicators that have coded for winning interactors that interacted on the basis of a common selective environment, are retained'.

Our framework could be used to focus on selective retention systems (Campbell, 1974) (4). We have however chosen not to do so in this paper, although we plan to do so in the future. The approach of the framework as presented is thus a selective one.

In evolutionary theory we can discern selective theories (Cziko, 1996) (7) and theories describing evolution by selection or, in other words, 'learning by trial and error'. Systems that behave in such a way are called 'selective retention systems'.

Theories of competition between firms resulting for instance in a monopoly, are selective theories, just as the description of different species competing for niches in ecological theory, without including the generation of adaptations of these species.

Evolution by selection-theories would of course include Darwin's evolution by natural selection, but also 'evolutionary learning' in economic theory as Vromen (1994) (30) mentions, or the 'trial and error learning' that Campbell (1974) (5) speaks of.

#### Note 10

Ernst Mayr (1982) (21) has pointed out that population thought is an important essence of Darwinian Theory. It is opposed to Platonic essentialism. In short this can be illustrated by

saying that essentialism describes entities according to their 'essence' for instance when we classify entities. For example describing a species in terms of characteristics can be called platonic. Darwinian Theory uses population thought which describes the difference in entities. A species would thus be described as having variation in a particular distribution. This is of course important because selection cannot be described without differences it can operate upon. The explanation of innovations thus needs population thought.

#### Note 11

Although we speak of 'are willing to', the choice of such models is often not consciously voluntary. People are often unaware of the models they use, and so have not chosen at all. They simply use the memes they are acquainted with.

### References

1. Brandon, R.N. (1988) The levels of selection: A hierarchy of interactors. In: Plotkin HC (ed) The role of behavior in evolution, Cambridge, MA: MIT Press.
2. Best, M.L. (forthcoming). Models for Interacting Populations of Memes: Competition and Niche Behavior. Journal of Memetics - Evolutionary Models of Information Transmission.
3. Calvin, W.H., (1997). The Six Essentials? Minimal Requirements for the Darwinian Bootstrapping of Quality. Journal of Memetics Evolutionary Models of Information Transmission, 1. [http://cfpm.org/jomemit/1997/vol1/calvin\\_wh.html](http://cfpm.org/jomemit/1997/vol1/calvin_wh.html)
4. Campbell, J. I. (1995). Mechanisms of Evolutionary Change in Economic Governance: interaction interpretation and bricolage. In: Netherlands Institute of Government: the role of institutions in the public sector. Papers for the NIG conference , november 9-10. Oosterbeek, the Netherlands.
5. Campbell, D.T. (1974). Evolutionary epistemology. In Schlippe PA (ED) The philosophy of Karl Popper, The Library of Living Philosophers volume XIV
6. Commissie Zeevalking (1991). Tweede rapport van de commissie Onderzoek - Universitaire Bestuurs - en Beheersstructuur, augustus. Katholieke Universiteit Nijmegen.
7. Cziko, G.(1996). Without Miracles. Universal selection theory and the second Darwinian revolution. Cambridge, MA: MIT Press
8. Dawkins, R. (1976). The selfish gene. Oxford: Oxford University Press
9. Dennett, D.C. (1991). Consciousness explained. Boston, MA: Little, Brown & Company.

10. Dunn, W.N. (1993). Policy reforms as arguments. In: Fisher, F. and J. Forrester (ED) *The Argumentative Turn in Policy Analysis and Planning*. Durham: Duke University Press, pp 254-290
11. Gould, S.J. (1997). Darwinian fundamentalism. *The New York Review*, June 12, 34-37
12. Grene, M. (1989). Interaction and evolution. In Ruse R (ed.), *What The Philosophy of Biology is. Essays dedicated to David Hull* p67-73. Dordrecht: Kluwer Academic Publishers.
13. Groot, T.L.C.M. (1988). Management van universiteiten. Een onderzoek naar de mogelijkheden voor doelmatig en doeltreffend universitair bestuur. Groningen: Wolters-Noordhoff.
14. Hull, D.L. (1980). Individuality and selection. *Annual Review of Ecological Systems*, 11, 311-332
15. Hull, D.L. (1982). The naked meme. In Plotkin HC (ED), *Learning Development and culture, essays in evolutionary epistemology*. New York: John Wiley & Sons.
16. Hull, D.L. (1988a). Science as a process. An evolutionary account of the Social and conceptual development of science. Chicago: The University of Chicago Press
17. Hull, D.L. (1988b). Interactors versus vehicles. In: Plotkin HC (ed), *The role of behavior in evolution*. Cambridge, MA: MIT Press
18. Jacob, F. (1977). Evolution and Tinkering. *Science*, 196 ,1161-1166
19. Lynch, A. (1996). Thought contagion. The new science of memes. How belief spreads through society. New York: Basic books
20. March, J. G. (1994). *A Primer on Decision Making. How decisions happen*. Toronto: The Free Press.
21. Mayr, E. (1982). *The growth of biological thought*. Cambridge, MA: Harvard University Press.
22. Pattee, H.P. (1977). Dynamic and linguistic modes of complex systems. *International Journal General Systems*, 3, 259-266.
23. Price, I. (1995) Organisational Memetics?: Organisational Learning as a Selection Process. *Management Learning*, 26: 299-318. <http://members.aol.com/ifprice/orgmem.html>
24. Ruse, M. (1989). *What The Philosophy of Biology is. Essays dedicated to David Hull*. Dordrecht: Kluwer Academic Publishers.
25. Simon, H.A. (1964). Rationality. In: Gould, J. and W.L. Kolb, (eds.), *A Dictionary of the Social Sciences*. New-York: Mac Millan Publishing Company. P. 573-574

26. Simon, H.A. (1976). *Administrative Behavior*. A study of decision-making processes in administrative organization. New York: Free Press.
27. Speel, H.C. (forthcoming a). A short comment from a biologist on William Benzon's essay 'Culture as an evolutionary arena'. *Journal of Social and Evolutionary Systems*.  
<http://www.sepa.tudelft.nl/webstaf/hanss/jses.htm>
28. Speel, H.C. (forthcoming b). Memetics: on a conceptual framework for cultural evolution. In: Heylighen, F., and D. Aerts (ED), *The evolution of complexity*. Kluwer, Dordrecht.  
<http://www.sepa.tudelft.nl/webstaf/hanss/hcesmem.htm>
29. Vaneechoutte, M. (1993). The memetic Basis of Religion. *Nature*, 365: 290.  
<http://www.sepa.tudelft.nl/webstaf/hanss/nature.htm>
30. Vromen, J.J. (1994). Evolution and efficiency. An inquiry into the foundations of 'New Institutional Economics'. Eburon Delft.
31. Williams, G.C. (1992). *Natural Selection: Domains, Levels, and Challenges*. Oxford University Press. New-York: Oxford University Press.
32. Williamson, O.E. (1985). *The economic institutions of capitalism*. New York: The Free Press.
33. Zwan, A. (1990) .Slotdocument: resultaten en voorstellen. Conferentie inzake de bestuursorganisatie van het hoger onderwijs. Kijkduin/Zeist 29/30.

# EXPLAINING CULTURE: A NATURALISTIC APPROACH

By  
Dan Sperber  
Oxford: Blackwell, 1996

How to Catch Insanity from Your Kids (Among Others); or, *Histoire naturelle de l'infame*

I don't know how many of my learned and refined readers have ever had occasion to pass through Roswell, New Mexico, located as it is in an especially bleak part of what is, on balance, a rather ugly state. While it perhaps does not deserve an actual visit, Roswell is certainly worthy of contemplation: like Lourdes, it has become famous for something which never happened. I refer, of course, to the infamous myth that a flying saucer crashed there in 1947, making the national news, only to be covered up by the US military. A shrine to this myth exists in the form of the grandiloquently named "International UFO Museum and Research Center"; but more diffuse homage is rendered to it in nearly every shop-window, down to wedding outfitters. The inside of the museum displays limitless credulity in its exhibits ("What I have seen refuted three times is true"), with tasteless and incompetent art for decoration, the latter all the more painful for its self-evident sincerity. To the skeptical but casual eye, the museum is a mis-guided monument to a particularly absurd fanaticism, to the strength of human gullibility, to a race of barely-thinking forked radishes' capacity for self-delusion. Having taken it thus in with a glance, that eye dismisses Roswell and its Museum, and moves on (probably on Route 285) to more pleasant prospects.

That eye moves on; but the truly philosophical eye will linger. The child of the Enlightenment, trained to uncover the hidden springs of conduct, penetrates the inner workings of the museum with ease. It is housed in an abandoned movie theater, obtained from the city at nominal cost. The list of founding sponsors of the museum, displayed as a discreet plaque in the lobby, contains no names distinguished among the believers in UFOs, but instead various members of the local business elite (who, as experience teaches, are the real rulers of any American small town), and their banks. This plaque is flanked by equally discreet awards from the local chamber of commerce. In sum: the museum was created as a tourist trap by the local rulers, to their direct benefit. Its existence and character are admirably explained by that blend of precepts which Marx and Engels took from the English and French historians of society and distilled into "historical materialism." This brew now meets with almost universal approval among the learned and sensible, as we have earlier remarked.

However satisfactory this potion may be in explaining deliberate products of human effort, like the Roswell Museum, it does not seem to work so well with other cultural phenomena, like the story of the Roswell UFO crash itself. For how are people to know what, in view of their stations in society, it would be advantageous for them to believe, *and* believe it sincerely? Have we all little Marxometers, pointing us in the direction of our socially preferred beliefs, of whose existence we are unaware, even when the connection to our social condition is so recondite that its elucidation requires hundreds of abstruse pages under the imprimatur of Verso or the Presses Universitaires de France? Do Gramsci's "organic intellectuals" act as external Marxometers? None of these conjectures is satisfactory. Nor is the obvious alternative: Agent Mulder, in his

more (or, perhaps, less) paranoid moments, may suspect the UFO phenomenon in its entirety of being fabricated by the "military-industrial-entertainment-research complex," but this is neither a general explanation for wide-spread beliefs, nor one which a loyal member of the complex aforesaid, such as your humble reviewer, will adopt unless no other alternative presents itself. What is needed is a way of establish consensus, of getting the same belief into the minds of the multitude, without Hidden Persuaders actually implanting it there like so many buttock-dwelling microchips.

Now, it is a fact of observation that belief is catching. Here is a reasonably old recognition of this: in the 1764, the public prosecutor of Geneva ordered Voltaire's *Philosophical Dictionary* to be banned, condemning it as "contagious poison." I do not pretend this rather unimaginative civil servant had obtained a new insight into the nature of belief, or even had come upon a novel turn of phrase; I merely mean to show that this insight is an old one, and being no historian, nor even diligent, I merely pulled old books off my shelf until I found something to the purpose. The heart of the matter is that people have long realized that ideas can spread as though they were agents of disease and infection; in a word, like viruses, using that term in its original sense.

There is today a great and rapidly growing body of knowledge concerning the agents of physical infection, and the temptation to exapt that knowledge to the sphere of culture, to declare beliefs, traditions, practices and the rest of their kith and kin "viruses of the mind" is more than a biologist may be reasonably expected to resist. There are different vectors; there are opportunistic infections; there are parasites altering the behavior of their hosts; there are all manner of subtle and bizarre adaptations employed to get the viruses from mind to mind, from speaking in tongues to multi-level marketing to the axiomatic method. Best of all, we have the theoretical apparatus of Darwinian evolutionary theory, which fairly cries out to be put to use explaining culture as the result of the differential reproductive success of randomly varying replicators.

Again: this is by no means a recent idea. The first published attempt to explain social and cultural change in a selectionist, that is to say a genuinely Darwinian, manner, was made in an 1880 essay by William James. (This essay is also notable for its discussion of the equivalents of founder effects and genetic drift in small populations, and of the dismal state of American politics in the gilded age.) The idea has been re-discovered at intervals ever since then: late in the 19th century, for instance, by the French sociologist Gabriel de Tarde (to whom Sperber, in the book under review, attributes its first formulation), and by the German psychologist Richard Semon, both of whom, alas, burdened it with unfortunate pieces of metaphysical baggage, and were on the losing side of academic quarrels, which combined to put an end to the notion among humanists and social scientists. Since then it has mostly been the province of biologists, such as the great demographer and pioneer of mathematical biology Alfred Lotka and the molecular biologist Jacques Monod, with some honorable exceptions among the other inhabitants of the commonwealth of letters: the philosopher Karl Popper, the geographer André Siegfried, the historian and philosopher of science Stephen Toulmin. The great breakthrough, however, was in the middle of the 1970s, when one of the biologists, Richard Dawkins, had the genius to give the idea one of its most compelling presentations; and the further genius to give the replicators a simple but neologistic name, "memes," which sticks in the memory, forces explanations (that is, opportunities for infection) and so helps propagation. Since then, even literary critics have

grasped the notion --- at least, those literary critics who can actually understand popular science books.

This is all very fine stuff; it makes a great mind-toy and is a wonderful stick to beat over the heads of the infamous. The problem with it --- and I say this with real regret --- is that it's very far from clear that most of the elements of culture are stable replicators.

Here at last we come to the contribution of Sperber. Having completed the initiatory field-work proper to a cultural anthropologist, he returned to France to effect a synthesis of anthropological knowledge and cognitive science. One product of this effort was an excellent book he wrote with the linguist Deirdre Wilson, *Relevance*. This, depending on how it is approached, is variously an assault of the self-proclaimed science of semiotics, a new understanding of how we understand natural languages, a hypothesis on hypothesis formation, or an outline of a solution to the frame problem of artificial intelligence (this last has been aired in these pages before). Another product was the reluctant conclusion that most of what passes for explanation in anthropology is no more explanatory than the scribblings of a newspaper editorialist; like them, it is really either merely description, or interpretation, or sometimes vacuous. Genuine explanations of culture would show how its phenomena arise as the compounded effects of the myriad acts of individual people. Any methodological individualist would have agreed to such a formula; Sperber seems to have been the first to rise to the challenge it presents.

To explain how it comes about that every sapient creature in the United States knows the Roswell story, and similar facts, Sperber saw that we need an "epidemiology of beliefs," or even of "representations" in general. So far he has the agreement of the whole line of cultural Darwinists, from James to Dawkins to Aaron Lynch. Now, however, he reaps the fruit of his anthropological and cognitive studies. Unlike the genetic apparatus, the mind never leaves its contents alone; to think, even to remember, is to change. Nor are these changes, as it were, equally in all directions. Rather the mind works upon its representations in the direction of what Sperber and Wilson call "maximum relevance" --- roughly, extracting the most new information from them for the least processing. (This concept is made clear and precise in their book.) The correctness of their relevance theory is not essential to Sperber's present argument; it is enough that representations change in a highly non-random fashion. The changes undergone by stories in oral transmission, for example, have been studied by folklorists and psychologists at least since Bartlett's classical work in the 1920s, and the tendencies are both well-established and pronounced. The stories become more dramatic; more reasonable (according to the lights of the tellers); more stereotyped (according to the stereotypes of the tellers); etc. (These effects have appeared in these pages before --- in connection with the Roswell story, no less.) But these are all changes that make the story easier to remember, easier to retell, easier to accept and easier to relate to the rest of one's life, i.e., they would increase the fitness of the story. But the decoupling of variation from selection has long been recognized as essential to selectionist theories; as James put it in his essay, it is necessary that they belong to "irrelevant cycles" of causes. Relevance thus constitutes one objection to the selectionist account of culture.

The same transformations block selectionist theories at another point as well, undermining the need for selection by providing another mechanism capable of inducing similar beliefs in a great many heads at once. Sperber imagines our representations (as instantiated in actual brains, not as

*abstracta*) as elements of the space of a dynamical system, with attractors and basins of attraction. He further imagines that these attractors are the representations of maximum relevance, and that the attractors and basins of attraction of two persons with broadly similar mental contents will also be very similar. But then, from very vaguely related starting-points, those two persons will arrive at very similar ideas about various subjects:

[C]onsider your views on President Clinton. They are likely to be very similar to the views of many, and to have been influenced by the views of some. However, it is unlikely that you formed your own views simply by copying, or by averaging other people's views. Rather, you used your own background knowledge and preferences to put into perspective information you were given about Clinton, and to arrive by a mixture of affective reactions and inferences at your present view. The fact that your views are similar to many other people's may be explained not at all by a copying process, and only partly by an influence process; it may crucially involve the convergence of your affective and cognitive processes with those of many people towards some psychologically attractive type of views in the vast range of possible views on Clinton. [p. 106] (The same mechanisms allow us, for instance, to correct typos, or even to read what we know should have been written in their place, without noticing that it isn't.)

It will not have escaped the attention of certain professionally-deformed readers that this verbal picture cries out for agent-based modeling. This would let us get at questions like when selection is more important than attraction, and the effects of introducing reliable, non-transformative storage --- writing and the kindred arts. There are fascinating issues in evolutionary game-theory as well, since the representational dynamics will to some extent depend on what representations are already held; though not so much as one might suppose. The reason for the limitation is connected to Sperber's account of religion.

Somewhere in the above-mentioned *Philosophical Dictionary* --- the passage eludes me just now --- Voltaire has fun with the effects of trying to believe an impossible or contradictory religious dogma. Sperber explains our ability to do so as follows. We can hold (says Sperber) "reflective" beliefs, "believed in virtue of second-order beliefs about them." There is no reason why these reflective beliefs must be comprehended or even comprehensible by those who hold them. Now, consider the following case. Young Bobby has in his belief box the two representations:

What Mom says is true.

Mom says that God is everywhere.

Bobby does not fully understand how somebody, be it God or anyone else, can be everywhere. However, his mother saying so gives him sufficient ground to exhibit all the behaviors symptomatic of belief: he will readily state that God is everywhere, will assent when the same statement is made by others, and may even refrain from sinning in places where (apparently) nobody can see him. That God is everywhere is for Bobby a reflective belief.... Here is a belief which, like most religious beliefs, does not lend itself to a final, clear interpretation, and which therefore will never become an intuitive belief. Part of the interest of religious beliefs for those who hold them comes precisely from the fact that you are never through interpreting them. While

the cognitive usefulness of religious and other mysterious beliefs may be limited... it is not too difficult to see how their very mysteriousness makes them "addictive." [p. 90]

Reflective beliefs are not limited to such statements as "God is everywhere," but include all of our knowledge which is not either direct perception, or arrived at from direct if not unconscious inference from perception: history, empirical science, most of mathematics. (But not all, apparently: see Stanislas Dehaene, *The Number Sense*.) Such "well-understood reflective beliefs ... include an explicit account of rational grounds to hold them. Their mutual consistency and their consistency with intuitive beliefs can be ascertained, and plays an important, though quite complex, role in their acceptance or rejection." By contrast, mysterious reflective beliefs are rationally held (when they *are* rational) solely upon the authority of those whom we learned them from. (Bobby may not be faulted for believing both his mother and his science teachers; but the one authority is valid, and the other a mere phantasm.) Thus, even if Sperber is correct, and the memeticists quite wrong, we do not lose a handy weapon with which to belabor the fanatical and the superstitious.

Sperber's book carries on from the "epidemiology of representations" in two directions. One is an explication of social institutions and social facts. Most if not all of these --- money, say, or marriage, or the existence of a research institute on Hyde Park Road in Santa Fe --- require, or even consist of, shared mental representations. (Money is whatever we all agree is money, be it stone circles, cowry shells, bales of tobacco, appropriately worked bit of soft metals, or the numbers stored in certain bit-fields of certain computers.) This is only going to work if the representations we hold internally are all sufficiently similar that no serious conflict as to our expectations arises: precisely what his conjectured mechanisms are designed to accomplish. This opens the way to fulfilling the demand made long ago by Friedrich Hayek, to "analyze social relations in terms of correspondence and non-correspondence, or compatibility and non-compatibility, of individual aims and desires" and expectations.

The other direction in which Sperber pursues the implications of his basic notion (that "culture is the precipitate of cognition and communication in a human population") is an explanation, grounded in evolutionary psychology and its view of the mind as an agglomeration of specialized modules, of how our representational attractors and basins of attraction form, and of how cultural diversity grows in this uniform mental soil. While this is, again, highly convincing, it defies adequate summary in less than the space (just above thirty pages) in which Sperber expounds it, and accordingly will be left to the reader.

Sperber concludes by reflecting on "What Is at Stake" in a naturalistic account of human culture; these are the dangers of success. (The inevitable charges of reductionism are brushed aside thus: "Let them [the critics] show that the so-called reductionist approaches are ill-conceived, or else let them articulate the moral reasons for their censorship" [p. 152].) The gravest of these is that it might distract the social sciences from their "fundamental role" of "enlightening citizens"; and another, more subtle one.

As in a mirror, we look for our image in the social sciences. When we do not recognize ourselves in the reflected image, we are disturbed. Cognitive psychology does not reflect an immediately recognizable image of ourselves; nor would an epidemiology of representations. Worse, what we

think of as essential and primary --- that is, our existence as conscious persons --- comes out, at best, as a changing pattern, socially projected on to a biological structure, itself precarious. [p. 155]

But Sperber views both of these possibilities as remote dangers, to be overcome by theoretical pluralism and the stubbornness which refuses to see the physical world as the crawling molecular chaos of the physicists.

Sperber's arguments carry conviction, at least at the hand-waving level at which the social sciences are accustomed to conducting themselves. Whether they can be carried further, to precision and even accuracy, is another question, technical and empirical. (It may even be possible to recover a Darwinian-Jamesian-Dawkinsian theory, complete with memes, by treating as equivalent all representations within the same basin of attraction, thus looking at the "slow modes," the infectious transmission that happens after the inner cognitive transformations have come to a halt.) Like the memeticists, Sperber raises the prospect of bringing to life the "specter of a natural science of the social": the question now is whether those dry bones can live.

ix + 175pp., endnotes, bibliography, index, 3 diagrams

Anthropology / Cognitive Science / Cultural Evolution / Sociology, Social Theory, etc.

Currently in print as a paperback, ISBN 0-631-20045-2, US\$37.75 [Buy from Powell's]

**AN OBJECTION TO THE MEMETIC APPROACH TO CULTURE**  
**In: Robert Aunger (ed.) *Darwinizing Culture: The Status of Memetics as a Science.***

Oxford University Press, 163-173.

**By  
Dan Sperber**

Memetics is one possible evolutionary approach to the study of culture. Boyd and Richerson's models (1985, Boyd this volume), or my epidemiology of representations (1985, 1996), are among other possible evolutionary approaches inspired in various ways by Darwin. Memetics however, is, by its very simplicity, particularly attractive.

The memetic approach is based on the claim that culture is made of memes. If one takes the notion of a meme in the strong sense intended by Richard Dawkins (1976, 1982), this is indeed an interesting and challenging claim. On the other hand, if one were to define "meme", as does the Oxford English Dictionary, as "an element of culture that may be considered to be passed on by non-genetic means," then the claim that culture is made of memes would be a mere rewording of a most common idea: anthropologists have always considered culture as that which is transmitted in a human group by non-genetic means.

Richard Dawkins defines "memes" as cultural replicators propagated through imitation, undergoing a process of selection, and standing to be selected not because they benefit their human carriers, but because they benefit themselves. Are non-biological replicators such as memes theoretically possible? Yes, surely. The very idea of non-biological replicators and the argument that the Darwinian model of selection is not limited to the strictly biological are already, by themselves, of theoretical interest. This would be so even if, actually, there were no memes. Anyhow, there are clear cases of actual memes, though much fewer than is often thought. Chain-letters, for instance, fit the definition. The very content of these letters, with threats to those who ignore them and promises to those who copy and send them, contributes to their being copied and sent again and again. Chain-letters don't benefit the people who copy them, they benefit their own propagation. Moreover, some chain-letters are doing better than others because of the greater effectiveness of their content in causing replication.

Once the general idea of a meme is understood - and especially if it understood fairly loosely -, it is all too easy to see human social life as teeming with memes. Aren't, for instance, religious ideas, with their threats of hell for unbelievers and promises of paradise for the proselytes, comparable to chain-letters, and in fact much more effective in benefiting their own propagation, come what may to their human carriers? More generally, aren't words, songs, fashions, political ideals, cooking recipes, ethnic prejudices, folktales, and just about everything cultural, items that get copied again and again, with the more successful items managing to invade more minds over longer periods of historical time, and to recruit those minds to further their own propagation? If this were so, if culture were made of memes in Dawkins's strong sense, then the study of culture could - and arguably should - be recast as a science of memes or "memetics". The Darwinian model of selection could be used, with proper adjustments, to explain the properties, the variety

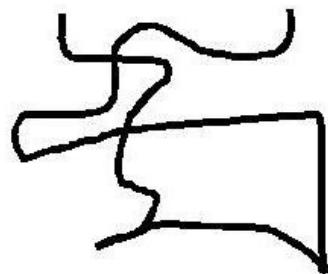
and the evolution of culture, just as it explains the properties, the variety, and the evolution of life.

The question is whether the claim that culture is made of memes is a true one. Several objections have been made to this claim. In his "foreword" to Susan Blackmore's *The Meme Machine* (1999), Richard Dawkins responds to the simplest and most serious objection: "that memes, if they exist at all, are transmitted with too low fidelity to perform a gene-like role in any realistically Darwinian selection process" (Dawkins 1999: x).<sup>1</sup> I want here to discuss Dawkins's responses, and, in so doing, develop a different fundamental objection to the meme model. This new objection is that most cultural items are "re-produced" in the sense that they are produced again and again - with, of course, a causal link between all these productions -, but are not reproduced in the sense of being copied from one another (see also Origgi & Sperber, forthcoming). Hence they are not memes, even when they are close "copies" of one another (in a loose sense of "copy", of course).

The objection of low fidelity had been envisaged and taken seriously by Dawkins himself. In *The Extended Phenotype*, he wrote:

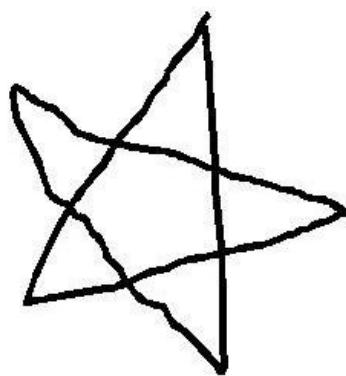
"The copying process is probably much less precise than in the case of genes: there may be a certain 'mutational' element in every copying event [...]. Memes may partially blend with each other in a way that genes do not. New 'mutations' may be 'directed' rather than random with respect to evolutionary trends. [...] there may be 'Lamarckian' causal arrows leading from phenotype to replicator, as well as the other way around. These differences may prove sufficient to render the analogy with genetic natural selection worthless or even positively misleading. My own feeling is that its main value may lie not so much in helping us to understand human culture as in sharpening our perception of genetic natural selection" (Dawkins 1982: 112).

Of course, what counts as "too low fidelity" for a given item is relative to the selection bias for that item (see Williams 1966). A greater selection bias allows for a higher mutation rate. On the other hand if, as Dawkins says, there is "a certain 'mutational' element in every copying event" (*loc. cit.*), then it is not easy to see how selection could work at all. It is to this problem that Dawkins now offers an ingenious solution. He uses for this a thought experiment of which I present a simpler but equally effective version (before discussing his version later on). Consider figure 1. A first individual is shown this figure for ten seconds and is asked, ten minutes later to reproduce it as exactly as possible. Then a second individual is shown for ten seconds the figure drawn by the first individual and presented with the same task. This is iterated with, say, nine participants. It is most likely that each drawing will differ from its model and that the more distant two drawings are in the chain, the more they are likely to differ. A judge given the ten drawings in a random order and asked to put them back in the order in which they were produced should perform, if not perfectly, at least much better than random. The "mutational elements" in every copying event are such that a drift is manifest, and no stable pattern is maintained.



**Figure 1**

Now imagine a similar experiment being performed, but this time with figure 2 as initial input. Again, each drawing produced by the successive participants is certain to differ from its model, since each participant will fail to reproduce the model in all its particulars. This time, however, the distance in the chain of two drawings on the one hand, and their degree of difference on the other hand should be two variables independent of one another (or nearly so). A judge asked to put the ten drawings in the order in which they were produced should be unable to do better than random. In spite of low fidelity of copying, a stable pattern is most likely to endure across versions, and individual variations are very unlikely to compromise this pattern.



**Figure 2**

What explains the difference between the two experiments? In the case of figure 1, people try and form a mental image of a drawing which they don't recognize in any way, and then try and reproduce this mental image on paper. In storing the information, in recalling it, and in

reproducing it, they are likely to introduce unintended variations that are either in random directions, or are in the direction of entropy, that is, plain loss of information. In the case of figure 2, people recognize the figure as a five-branched star drawn without lifting the pencil. They may well forget most of the other particulars of the drawing under their eyes, such as length of relatively straight segments, or angles. Still, they will produce another star of the same type.

Dawkins might describe the difference between the two types of tasks as follows. In tasks of the first type, what gets copied is the product, the drawing. There is no difference therefore between the "phenotype" and the "genotype", and phenotypic variations are also genotypic variations. In cases of the second type, what gets copied is the implicit instruction ("draw a five-branched star without lifting the pen"). These instructions are the true genotype, while the drawings are only phenotypes. Each participant in the experiment assumes that the preceding participant merely intended to follow the implicit instruction and that imperfections or idiosyncrasies were unintended and should be ignored. Individual variations in the productions of the phenotype do not matter. They are not genuine mutations. "The instructions," writes Dawkins "are self-normalizing. The code is "error-correcting" (p. xii).

Dawkins concludes the argument by stating: "I believe that these considerations greatly reduce, and probably remove altogether, the objection that memes are copied with insufficient high fidelity to be compared with genes. For me the quasi-genetic inheritance of language, and of religious and traditional customs, teaches the same lesson." (p. xii) In other terms, the stability of cultural patterns is proof that fidelity in copying is high in spite of individual variations. These variations are phenotypic, not genotypic, and Darwinian selection can take place without being jeopardized by too high a rate of mutation.

I, on the other hand, believe that what is here offered as an explanation is precisely what needs to be explained, what is offered as a solution is in fact the very problem to be solved. Saying that the instructions are "self-normalizing" amounts to resolving a problem by invoking a mystery. The type of thought experiment proposed by Dawkins is well worth analyzing so as to solve the mystery. The conclusions I draw from this thought experiment are, however, very different from Dawkins's. They point to yet another difficulty with the meme model.

Let me grant forthwith two points to Dawkins:

(a) Of course, one item A can be a replica (in the relevant sense) of another item B without being identical to B in every respect. From a memetic point of view, it is enough that A and B should share the properties the recurrence of which one is trying to explain.

(b) Of course, cultural items exhibit, over periods of time of various length (longer for folktales, shorter for modern dress fashions, for instance), the kind of stability found, on a much smaller scale, in Dawkins's thought experiment. That is, although there is much individual variation, items of the same type all remain in the vicinity of one another and instantiate a common pattern. The issue is whether the relative stability found in cultural transmission is proof of replication. Dawkins seems to think it is. In substance, he proposes a test to decide whether a causal chain that links the production of a series of items is a chain of replications. The test is as follows.

Present (or suppose you present) to an intelligent observer the items in the chain in a random order. If the observer finds it impossible to put back, at least approximately, the items in the order in which they were produced, then these items are replications in the relevant sense. Individual variations among these items are phenotypic and do not compromise the stability of the underlying genotype. Much of culture passes this test and is seen, then, as made of replicators.

To show that Dawkins's test is not as reliable as it may seem, let me first give an example of a causal chain that would meet the criterion, but could not be properly described as a case of memetic transmission. Consider the case of laughter. Laughter is a social behaviour that is typically triggered, in individual development, by the laughter of others, and that remains a highly contagious form of behaviour. Laughter is influenced in its intensity, style, and circumstances of arousal by cultural factors. Moreover, even within a cultural group, there are important individual variations. Now, imagine a series of registerings of causally linked individual laughers (linked either in the stabilization of laughing behaviour across generations, or in a much shorter causal chain of contagious laughter). If these registerings were presented in a random order, they could not, I take it, be rearranged in their causal order. Laughter passes Dawkins's test. Yet, surely, it is not a meme.

Why is laughter not a meme? Because it is not copied. A young child who starts laughing does not replicate the laughers she observes. Rather, there is a biological disposition to laughter that gets activated and fine-tuned through encounters with the laughter of others. Similarly, an individual pushed into convulsive laughter by the laughter of others is not imitating them. The motor program for laughing was already fully present in him, and what the laughter of others does is just activate it.

Let me generalize and define three minimal conditions for true replication. For B to be a replication of A,

- *B must be caused by A (together with background conditions)*
- *B must be similar in relevant respects to A*
- *The process that generates B must obtain the information that makes B similar to A from A.*

Another way to express this third condition is to say that B must inherit from A the properties that make it relevantly similar to A. Discussions of memes take implicitly for granted that the co-occurrence of causation and of similarity between cause and effect is sufficient evidence of inheritance. But this is not so. The cause may merely trigger the production of a similar effect, as we saw with the case of laughter. Even if conditions (1) and (2) are satisfied, condition (3) may not be.

Consider a theoretical example, with two cases to be compared. In both cases conditions (1) and (2) are satisfied, but condition (3) is satisfied only in the second case. *First case:* Ten sound-recorders with the same repertoire of melodies in each have been fixed so that they are activated by the sound of the last five bars of any melody in their repertoire, and then play this very melody. They are placed in such a manner and at such a distance of one another that the first one

activates the second, the second the third, etc. The first recorder plays melodies in random order at appropriate time intervals. *Second case*: Ten sound-recorders have been fixed and placed so that the second recorder records sound from the first, and then replays it, the third recorder records sound from the second and then replays it, and so on. Only the first recorder has a ready repertoire of melodies, and it plays them in random order at appropriate time intervals. In both cases, an observer listening to these devices playing, each in turn, one melody after another, and unable to inspect them otherwise, would have some reasons to think she was witnessing a series of replications. In fact, this would be true in the second case, but not in the first, where only triggering takes place and no copying at all. This illustrates the point that, in the case of a causal chain that satisfies conditions (1) and (2), further evidence about the causal processes involved must be available before one is in a position to argue that condition (3) is also satisfied, and that one is dealing, therefore, with a true chain of replications.

Let us go back, now, to our thought experiment. In the first task (memorizing and reproducing figure 1), participants rely on general perceptual, memory and motor abilities. In other term, they rely on the general human ability to imitate, an ability which is taken by memeticists to be extremely powerful. In this case, however, it fails. In the second task (memorizing and reproducing figure 2), the stimulus is *recognized*. That is, it triggers the activation of pre-existing knowledge. The stimulus is categorized as a token of a general type: a five-branched star drawn without lifting the pencil. Properties of the actual stimulus that are irrelevant to this categorization are just ignored. When asked, after ten minutes, to reproduce the stimulus, participants just produce another token of a five-branched star without, in most case, even trying to remember what the original figure exactly looked like. Their ability to perform well in this second task is not an ability to perceive and copy. It is an ability to recognize and re-produce, using, for this, knowledge of the five-branched-star type that they already possessed before encountering the token. It is not, then, that people are better at imitating figure 1 than at imitating figure 2. They are indeed bad at imitating figure 1, and they are not imitating figure 2 but merely producing a new token of the same recognizable type.

Dawkins's original thought experiment involved a comparison of two tasks: reproducing a drawing of a Chinese junk, or making an origami Chinese junk after having been taught, by demonstration, how to make one. Unlike my simpler version, the two final products - the drawing or the origami - are recognized by the participants. In the drawing version, however, participants are unable to recognize the series of strokes that would yield the full drawing, whereas in the origami version the successive foldings are individually demonstrated. Thus, the two tasks are different, not just in the type of item to be copied (a drawing vs. an origami) but also in the fact that participants observe only the product in the first task, and the process of production in the second task. If participants were just shown a finished origami junk, they would, presumably, do even worse in reproducing it than in reproducing a drawing of a junk.

The crucial difference between the two tasks is that the second involves demonstration and the other not. From the demonstration, or so Dawkins assumes, participants can and do infer implicit instructions (e.g. "take a square sheet of paper and fold all four corners exactly into the middle"). These instructions are not a description of what the person making the origami is actually doing (the four corner are never folded exactly into the middle, for instance) but a description of what the person is aiming at, is intending to do. Inferring instruction involves much more than the

ability to perceive and describe actual movements; it involves the ability to attribute goals and intentions.

Contrary to what Dawkins writes, the instructions are not "self-normalizing". It is the process of attribution of intentions that normalizes the implicit instructions that participants infer from what they observe. When you see the person folding the four corners of a square sheet of paper into four different points in the vicinity of the middle, you assume that she was aiming at the middle rather than at these four odd points. Such intentions to realize regular geometrical pattern are familiar - in particular in the context of origami - and readily attributed. You recognize, in other terms, the behaviour as an imperfect realization of an intention of a familiar and regular type rather than as the perfect realization of an intention of an unfamiliar and irregular type. The instructions that you infer are, then, informed in part by what you actually observe, and in part by what you already know of human intentions, and of the type of instructions typically used in origami.

The instructions are not being "copied" in any useful sense of them term from one participant to the next. Certainly, instructions cannot be *imitated*, since only what can be perceived can be imitated. When they are given implicitly, instructions must be inferred. When they are given verbally, instructions must be comprehended, a process that involves a mix of decoding and inference (Sperber and Wilson 1995). The inference involved in either case draws on domain-specific competencies having to do with the attribution of intentions and with knowledge of the role of regular geometric forms in the formation of human intentions generally, and in paper folding in particular. Thus the normalization of the instructions results precisely from the fact that something other than copying is taking place. It results from the fact that the information provided by the stimulus is complemented with information already available in the system.

In the real world, and in particular in the cultural world, triggering and copying can and do combine in various degrees. What gets being triggered by cultural stimuli are acquisition mechanisms and competencies that are more or less domain-specific. These mechanisms are themselves in part genetically, in part culturally inherited.

Let us briefly consider the example of the acquisition of language. In acquiring a language, a child internalizes a grammar and a lexicon on the basis of linguistic interactions. Nowhere in these interactions, nowhere in the linguistic data is the child presented with the grammar present to be copied. Rather the grammar must be inferred from these data. As Noam Chomsky has long argued and as has become, if not universally, at least generally accepted today, this requires a genetically determined preparedness to interpret the data in a domain-specific way and to generalize from it to the grammar of the language, going well beyond the information given. Imitation in some sense may well play a role - though not a sufficient one - in the acquisition of the phonology of words, but not in the acquisition of their meaning. Meaning is not something that can be observed and copied. It can only be inferred. Language learners converge on similar meanings on the basis of weak evidence provided by words used in an endless diversity of contexts and with various degree of literalness or figurativeness. Acquisition of meaning in such conditions is a feat that would be wholly mysterious if it were not highly constrained by domain-specific competencies having to do with conceptual domains on the one hand and with the attribution of communicative intentions to speakers on the other hand. Thus the similarities

between the grammar and lexicons internalized by different members of the same linguistic community owe little to copying and a lot to pre-existing linguistic, communicative, and conceptual evolved dispositions.

The respective role of copying and that of pre-existing dispositions to construe evidence in domain-specific structured ways may vary with different cultural competencies. Learning to tap-dance involves more copying than learning to walk. Learning poetry involves more copying than learning philosophy. For memetics to be a reasonable research program, it should be the case that copying, and differential success in causing the multiplication of copies, overwhelmingly plays the major role in shaping all or at least most of the contents of culture. Evolved domain-specific psychological dispositions, if there are any, should be at most a relatively minor factor that could be considered part of background conditions. There is nothing obvious about such a view. While the view may have some popularity among unconcerned lay people, no psychologist believes that cultural learning is essentially a matter of imitation (this is true even of psychologist who attribute an important role to imitation, e.g. Meltzoff & Gopnik 1993, Tomasello & al. 1993). In fact, such an idea goes against all major recent developments in developmental psychology and in evolutionary psychology (see Hirschfeld & Gelman 1994). This, together with the problem raised in this article, puts a special burden on memeticists.

Memeticists have to give empirical evidence to support the claim that, in the micro-processes of cultural transmission, elements of culture inherit all or nearly all their relevant properties from other elements of culture that they replicate (i.e. satisfy condition 3 above). If they succeeded in doing so they would have shown that developmental psychologists, evolutionary psychologists and cognitive anthropologists who argue that acquisition of cultural knowledge and know-how is made possible and partly shaped by evolved domain-specific competencies are missing a much simpler explanation of cultural learning: imitation does it all (or nearly so)! If, as I believe, this is not even remotely the case, what remains of the memetic program? The idea of a meme is a theoretically interesting one. It may still have, or suggest, some empirical applications. The Darwinian model of selection is illuminating, and in several ways, for thinking about culture. Imitation, even if not ubiquitous, is of course well worth investigating. The grand project of memetics, on the other hand, is misguided.

## Note

1. Dawkins adds: "The difference between high fidelity-genes and low-fidelity memes is assumed to follow from the fact that genes, but not memes, are digital." The objection that memes are transmitted with too low fidelity can be made without this further claim, which I find vague and unconvincing.

## References

Boyd, Robert & Peter J. Richerson (1985). *Culture and the evolutionary process*. Chicago: The University of Chicago Press.

Dawkins, Richard (1976). *The Selfish gene*. Oxford: Oxford University Press.

- Dawkins, Richard (1982). *The extended phenotype*. Oxford: Oxford University Press.
- Dawkins, Richard (1999) "Foreword" to Susan Blackmore *The Meme Machine*. Oxford: Oxford University Press.
- Hirschfeld, Lawrence & Gelman, Susan eds. (1994). *Mapping the Mind: Domain specificity in cognition and culture*, New York: Cambridge University Press.
- Origgi, Gloria & Sperber, Dan (forthcoming). Evolution, communication, and the proper function of language. In Peter Carruthers and Andrew Chamberlain eds., *Evolution and the Human Mind: Language, Modularity and Social Cognition*. Cambridge: Cambridge University Press.
- Meltzoff, A. and Gopnik, A. (1993). The role of imitation in understanding persons and developing a theory of mind. N S. Baron-Cohen et al. eds., *Understanding other minds*. Oxford: Oxford University Press.
- Sperber, Dan (1985). Anthropology and psychology: towards an epidemiology of representations. *Man (N.S.)* 20, 73-89.
- Sperber, Dan (1996). *Explaining culture: A naturalistic approach*. Oxford: Blackwell.
- Sperber, Dan and Wilson, Deirdre (1995). *Relevance: Communication and cognition, Second Edition*. Oxford: Blackwell.
- Tomasello, M., Kruger, A. and Ratner. H. (1993). Cultural learning. *Behavioral and Brain Sciences* 16: 495 - 552.
- Williams, George C. (1966). *Adaptation and natural selection*. Princeton: Princeton University Press.

# THE MEMETIC ORIGIN OF LANGUAGE: MODERN HUMANS AS MUSICAL PRIMATES

By

**Mario Vaneechoutte**

Department of Clinical Chemistry, Microbiology & Immunology,  
Blok A, University Hospital, B 9000 Ghent, Belgium  
E-mail: [Mario.Vaneechoutte@rug.ac.be](mailto:Mario.Vaneechoutte@rug.ac.be)

**John R. Skoyles**

6 Denning Road, Hampstead,  
London NW3 1SU. UK  
E-mail: [Skoyles@globalnet.co.uk](mailto:Skoyles@globalnet.co.uk)

## Abstract

Song (musicality, singing capacity), we argue, underlies both the evolutionary origin of human language and its development during early childhood. Specifically, we propose that language acquisition depends upon a Music Acquiring Device (MAD) which has been doubled into a Language Acquiring Device (LAD) through memetic evolution. Thus, in opposition to the currently most prominent language origin hypotheses (Pinker, S. 1994. *The Language Instinct*, W. Morrow, N.Y.; Deacon, T.W. 1997. *The Symbolic Species*, W.W. Norton, N.Y.), we contend that language itself was not the underlying selective force which led to better speaking individuals through natural selection. Instead we suggest that language emerged from the combination of (i) natural selection for increasingly better mental representation abilities during animal evolution (thinking, mental syntax) and (ii) natural selection during recent human evolution for the human ability to sing, and finally (iii) memetic selection that only recently (within the last 100,000 years) reused these priorly evolved abilities to create language. Thus, speech - the use of symbolic sounds linked grammatically - is suggested to be largely a cultural phenomenon, linked to the Upper Paleolithic revolution. The ability to sing provided the physical apparatus and neural respirational control that is now used by speech. The ability to acquire song became the means by which children are able to link animal mental syntax with syntax of spoken language. Several studies strongly indicate that this is achieved by children through a melody-based recognition of intonation, pitch, and melody sequencing and phrasing. Language, we thus conjecture, owes its existence not to innate language learning competencies, but to innate music-associated ones, which - unlike the competencies hypothesized for language - can be straightforwardly explained to have evolved by natural selection.

The question on the origin of language then becomes the question on the origin of song in modern humans or early *Homo sapiens*. At present our ability to sing is unexplained. We hypothesize that song capacity evolved as a means to establish and maintain pair- and group-bonding. Indeed, several convergent examples exist (tropical song birds, whales and porpoises, wolves, gibbons) where song was naturally selected with regard to its capacities for reinforcing social bonds. Anthropologists find song has this function also amongst all human societies.

In conclusion, the ability to sing not only may explain how we came to speak, but may also be a partial answer to some of the very specific sexual and social characteristics so typical for our species and so essential in understanding our recent evolution.

**Keywords:** origin, human, language, natural selection, cultural evolution, music, intonation, rhythm, song, children

## 1. Introduction

### 1.1. Memetics and the origin of language

A major topic of memetics is the transmission of information by words. Thus better knowledge about the origins of language could throw light on many of the issues that are presently debated in memetics. Understanding what language is about is also important because it can be put that language is the only essential difference between human and animal existence, a difference which enables to explain most or all of the features characteristic for human psychology and human behaviour (e.g. [92]), which in turn explains why some memes - which we define broadly as bits of behaviourally transmissible information [Note 1] - are spread more successfully than others [Note 2].

Finally, memetic selection tends to get ignored by theories seeking to explain the origins of language. Most prominent theories instead argue for a gene- rather than meme-based origin. Some, for example, conjecture that language arose from Darwinian, adaptationist selection processes by which better speakers had greater reproductive success (Pinker [67], Smith and Szathmary [83]). Also Deacon [21] relies on genes (genetic assimilation of phenotypic characteristics (Baldwinian evolution)) and long term evolution to explain how language could arise, although memes (i.e. the use of symbols) already play an important role in this Baldwinian evolution. To the opposite, the memetic selection hypothesis defended here assumes that all preadaptations for language production and language understanding were naturally selected for other reasons than language, wherefrom language emerged and evolved rapidly and only recently by a process of cultural evolution. Thus, we do not reject natural selection - indeed, our approach largely depends upon it - but we try to understand how and when cultural/memetic selection comes into play and eventually takes over. The emergence of language in a human community by interaction between humans and symbols is not specifically addressed here, but will be the issue of a forthcoming paper [82].

### 1.2. The evolution of information and language

Informationally, 'life' can be considered as a giant chemical process which took off some 4 billion years ago - with the origin of the first self-replicating cell (also see [Note 2]). For instance, the product of one enzyme can be used as the substrate for other enzymes, cells interact by means of hormones, humors and neurotransmitters and multicellular colonies do so by means of pheromones and scent molecules. And across these organisational levels, the interactions between enzymes enable cells to interact by means of hormones, humours and neurotransmitters, while the cellular interactions enable multicellular colonies to interact by means of pheromones and scent molecules.

When brain (and eye and ear) possessing animals arose, a new manner of information transmission was introduced into biology. Now, living creatures could exchange information in a nonchemical manner through sound and sight, which increased the speed and the flexibility of informing each other and of influencing each other's behaviours or of adapting one's behaviour to the cues provided by others, i.e. behavioural instead of chemical interaction. One of the characteristics of such information is that it is 'inherited' in a nongenetic manner [35, 93]. One easily imagines how it is impossible to inherit rules for complex social behaviour genetically while on the other hand such morals, habits are readily 'learned' (rather: assimilated, absorbed) by young animals, e.g. through (emotional) punishment and reward (also see [Note 1]).

Dawkins [20] has called these 'culturally inherited' bits of information, 'memes'. Their development went hand-in-hand with parallel evolution of semantic abilities: these signals have to be turned by the brain into information that can be processed by cells that interact chemically (that is, by neurons and their information transmission by neurotransmitters). For example, an alarm cry (auditory cue) has to get linked/recoded by semantic abilities into the neurological (biochemical) concept or category or mental representation of 'dangerous situation'; this then triggers the various neurotransmitter- and/or hormone-transmitted cellular interactions that occur in fright and fear.

Importantly, the alarm cry is already here functioning as a kind of symbol, since similar sounds or signs may have a completely different meaning depending on the situation or the species. The full progress to symbolism in communication - spoken language as used by humans - requires that further encoding takes place: not only are symbols (words or signs) linked to mental images (linguistic semantics), but also word order, affixes, and other morphological modifications/operations/processes enable the communication of relationships between representations (linguistic syntax). Of course, the possibilities for individual thinking and for information transmission between individuals as a result of symbolic language are increased exponentially again.

### **1.3. The major questions about language**

Two questions exist about language development and its origins.

First, 'How is it that the complex phenomenon of spoken language arose in the animal world?'. Second: 'How is it possible that children acquire language with such ease irrespective of the language they learn?'

We argue that none of the presently available hypotheses [21, 67] provide final answers to the questions about the origin of language both during phylogeny of the human kind and during ontogeny of the individual human being. First, we briefly review the present dominant approaches.

The present prominent hypothesis on the phylogenetic origin of language is the natural selection approach. This claims that speech - and its associated characteristics like voice, and specialized brain regions and the ability to comprehend syntax of spoken language - was selected by gradual natural selection of genetic changes, made possible by the selective advantage of speech itself [66, 67, 83] or was made possible by genetic assimilation [21].

The present prominent hypothesis on the origin of language during individual ontogeny [16, 17] - a hypothesis taken for granted by e.g. Pinker [67] and Smith and Szathmary [83] - is that children have a language acquiring device (LAD), that uses an innate Universal Grammar. Syntax, according to this view, is acquired by a child by setting a few parameters of the 'innate grammar' according to those in their parent's language.

In this paper, we argue, in opposition to these approaches, first that song production and song interpretation capacities were the essential, naturally selected, preadaptations that enabled language, which readily evolved in a cultural (memetic) manner. In other words, speech preadaptations were naturally selected but only in regard to singing and not in regard to the later use they came to have in language. Second, linked to this 'song being the preadaptation for speech' approach, we argue that children learn spoken language by means of innate melody recognition capacity (Music Acquisition Device or MAD). If genetic evolution contributed to our abilities to learn language, it was in an indirect manner by providing us with abilities to sing. Thus, language learning devices can in fact be considered as memetically adapted song learning ones.

## **2. Language components**

### **2.1. Four competencies**

The capacity to produce and comprehend spoken information presupposes several cognitive abilities.

#### **2.1.1. Mental semantics**

The mind must be able to make mental images (virtual representations) that represent externally perceivable objects, agents and situations.

Furthermore, the mind must be able to semantically link sounds (or visual signals in animal behaviour, writing or sign-languages) to these mental representations to enable their communication, i.e., to make sense of them or to convey meaning. These links are symbolic - that is arbitrary or established by convention. This is illustrated by the fact that many words exist in different languages for the same concept or thing. Accordingly, many completely different writing conventions exist.

#### **2.1.2. Mental syntax**

The mind must be able to establish the relations and interactions between these representations, whereby some are active or originators (agents, subjects), while others undergo a change in situation or receive actions (patients, objects).

#### **2.1.3. Vocal flexibility**

Spoken language depends on the vocal dexterity to produce a wide range of consonants, vowels and intonations.

#### **2.1.4. Linguistic syntax**

The mind must be able to process the word order, syntactic 'roles' such as verb, subject and morphological modifications (syntactic affixes and internal phoneme changes) so that it can produce and comprehend them when communicating with others.

### **2.2. Evolution of mental semantics and mental syntax (the first two competencies)**

#### **2.2.1. Mental representation and semantic ability**

The ability to form mental representations, mental images of the environment and to categorize these objects does not need explanation in the context of the origin of language. It is clear that categorization and generalisation is a selectively advantageous trait to any heterotrophic, multicellular, mobile, brained organism (i.e., to most animals), since this enables an animal to reduce the reaction time upon perception. If animals, for example, could not generalize the concept of a certain species of tree, they would be forced to investigate each tree as to whether its fruits were edible. It is easily understood how better and better mental representation or categorization capacities were continuously naturally selected for throughout animal evolution. We define semantic skill as the ability to link visual or auditory stimuli to a mental representation. This skill is nothing new as animals readily assign meaning to auditory or visual signals. For example, a dog observing a bell ringing will quickly learn to link it to the subsequent appearance of food - Pavlovian conditioning. The dog's brain finds no difficulties in linking an arbitrary sound to an internal representation of a real or possible event. Such semantic abilities can be expected to have been selected particularly for auditory and visual signalling to aid communication between members of a species. Thus, social animals seem well equipped to acquire and use culturally inherited semantic meanings.

The semantic ability to link auditory sounds and visual signs to meaning, it should be noted, lies at the heart of culture, and of memes (see also 1.2). Culture, indeed, can be broadly defined as the exchange of information by auditory and visual behavioural cues. Thus, semantic ability enables the existence of behavioural memes (cry x means danger, sound y means anger, melody z means affection ...) that survive as replicated bits of information with high heritability within learnt culture. This can be seen in the rules for social behaviour that get culturally, behaviourally, memetically inherited among social animals from generation to generation. The composition of memes in an animal culture is often more stable than the genetic composition of the population. These behaviours are the forerunners of the symbolic memes.

It should be noticed that linking arbitrary sounds (words) to specific meanings in essence requires no novel skills: also dogs can learn it. From this it follows that we do not need to explain any quantitative differences in semantics. All we need to explain is why humans are so good at this.

#### **2.2.2. Mental syntax**

Also we need not explain the existence of mental syntax in the context of language development: the expansion in thinking and intelligence by means of increased mental syntactical abilities is

observed throughout vertebrate and invertebrate taxa. Animals recognize different agents and their interactions and the causal links between ongoing processes connecting them. We have now plenty of examples of mental representation possibility and of generalization, categorization and causal reasoning, i.e. thinking in animals [36, 37, 40, 53]. There is now strong indication that chimps even succeed in forming mental representations of the knowledge present in the mind of another subject [Note 3].

Mental syntax offers a strong selective advantage because it enables animals to predict possible outcomes of current situations (aided by memory of past events and their outcomes) and it helps them to make the best choices between different possible actions.

### **2.2.3. Conclusion**

We conclude that mental representation and semantics (linking of observable behaviour to a mental representation or conveying meaning to an observation) on the one hand and mental syntax (recognition of causal links) on the other hand are two abilities that were naturally selected for long before humans appeared and long before the rise of spoken language. Before we explain how vocal flexibility and linguistic syntax arose, we will briefly summarize why we need other hypotheses than those proposed by Chomsky [16, 17], Pinker [67] and - to a large extent - Deacon [21].

## **3. Problems with Chomsky, Pinker and Deacon**

### **3.1. Noam Chomsky: Is there a universal grammar?**

Chomsky [16, 17] argues that syntactical skills are novel to human communication, arising in each variety of human language from parameters set in an innate Universal Grammar. However, current knowledge does not provide evidence of something like innate Universal Grammar. A peculiar feature of human language is the high degree of diversity in each of its characteristics. For example, in configurative languages - like Indo-European languages, subject (S), object (O) and verb (V) can mathematically be ordered in 6 manners. Although SOV (45%) and SVO (42%) are predominant [1], five of the six possibilities are used (there are no examples known of OSV), by itself a strong indication that almost any conceivable order can be used. Some languages (like Dutch) also use mixtures of SOV and SVO, dependent on the hierarchical position of the sentence. Even more curiously, there are nonconfigurative languages, like Guugu Yimidhirr from N.E. Australia (see also below).

Aitchison has reviewed the difficulties in finding underlying universal rules in grammar of spoken languages [1].

However, to Chomskyans all this diversity is an illusion and under the surface all of these languages are dialects of 'earthspeak' - as Pinker [67] puts it. The syntactical differences are due to different parameter settings given to an innate Universal Grammar. There is a problem with this answer: the notion of Universal Grammar, frankly, is philosophical speculation. We refer to Botha [11], Harris [39], Tomasello [89], Allot [3], Bates & Goodman [6] for useful sources underpinning our skepticism. This nonfactual status of the Universal Grammar hypothesis has

been ignored since Chomskyans have been remarkably successful in promoting the idea that Universal Grammar is to language what molecules are to chemistry, gravitation to astronomy or DNA to biology - an established fact.

Not only is 'universal grammar' not universal, it does not even concern many aspects of grammar. Really 'odd' languages exist like Nootka and Mohawk, (two Native American languages), Lisu from Burma and Mam from West-Guatemala. The latter for example has a rich vocabulary for the action of laying, depending on the position of laying (on belly, on back, on side), depending upon whether a human or an animal is laying, upon telling whether one lays sick or drunk, etc.). Some languages have no gender classes, some two, other three and Sothere even has six. Furthermore, languages use a limited subset from 757 phones (observed in a total of 317 languages) very differently. They do so with a varying number of consonants and vowels from as low as 11 in the case of the Polynesian Mura to 148 in the African !xu or !Kung. Most average between 20 and 35 [52]. This is not explained by Chomskyan theory.

It appears that 'Universal Grammar' is not a scientific fact but a program which for theoretical reasons assumes that language has a universal core. It is an assumption which Chomskyans have constantly failed to establish. Chomskyan linguistics, we further note, is ignored by many linguists. Rather than being the only approach to grammar - as has been suggested by Pinker [67], it is only one of several. Further, the people who we might expect to make use of it - that is those seeking to computerise speech, completely ignore Chomskyan theory. Also many cognitive psychologists reject it as it fails conspicuously to fit the process by which syntax is acquired by children [6].

The noncredibility of 'Universal Grammar' leaves us with a hard problem: how could natural selection have created the diversity we find amongst human languages? Diversity does not offer the user of any language any advantage. (The only people, we might note, that gain from it are linguists who can make careers based upon studying obscure languages).

In the case of phonetics this is particularly problematic since it is known that infants before nine months are prepared to hear the phone contrasts present in all languages [27, 86, 90], an ability which is lost once they are familiar with their own language [4]. What advantage could exist for such an ability?

This is an unacknowledged but puzzling anomaly in the evolution of phonology and one can wonder why there should exist such a variety when only a small subset of phones (roughly one in twenty to one in forty) are used in any particular language.

In our opinion, this diversity creates work for phoneticians, but it makes no sense evolutionary, except in one circumstance: that linguistic evolution was not responsible for selecting the processes responsible for phone differences but instead co-opted existing diversity to the - phonetically more limited - needs of speech. In this view, speech evolution limited itself to developing means to use preexisting information processing sensitivities in the temporal/parietal and motor cortices.

In fact we know this is the situation in phonetics: animals as different to us from chinchillas [46] to quails [45] can hear phones. The auditory cortex of monkeys is as able as that of humans to hear the auditory features which characterise phones [85]. Neurons in the homologous areas to Wernicke's area process phonetic parameters such as fundamental frequencies, voice onset times and place of articulation (for instance, [85]). Phonetics appears to be a case where an important component of speech was not a direct product of natural selection but one that came about from a reuse of processes that had been already evolved much earlier for other reasons.

In conclusion, although we argue that something like universal mental syntax exists (see 2.1.2), we fail to see how spoken syntax could rely on a universal linguistic grammar.

### **3.2. Steven Pinker: Is it plausible to assume that natural selection for better speech capacities could have happened?**

#### **3.2.1 Introduction**

Can language be understood as 'an organ', a 'language instinct' [67], that was developed by gradual natural selection in which better speakers had more reproductive success, resulting in the selective survival of genes encoding for such better language abilities?

Pinker [67] and Pinker & Bloom [66] have suggested that the Chomskyan innate Universal Grammar arose by natural selection. There are many problems with this proposal. Bickerton [9], for example, in spite of being committed to the idea that an innate Universal Grammar arose by natural selection, felt the problems of this happening were so great that it could only be explained by a single and extraordinary macromutation, which is clearly unacceptable to any evolutionary biologist.

The following quote summarizes how Pinker and Bloom [66] propose that natural selection could have played a role in the development of language:

*"Furthermore, in a group of communicators competing for attention and sympathies there is a premium on the ability to engage, interest, and persuade listeners. This in turn encourages the development of discourse and rhetorical skills and the pragmatically-relevant grammatical devices that support them. Symons' [87] observation that tribal chiefs are often both gifted orators and highly polygynous is a splendid prod to any imagination that cannot conceive of how linguistic skills could make a Darwinian difference."*

Below are some of our objections.

#### **3.2.2. Language genes**

Natural selection for language only works if it can be genetically inherited. However, thus far no language genes have been reported in spite of intensive searches that have been made for inherited disorders of language. The case of the family known by the initials KE with an inherited disorder demonstrates the failure of this search, paradoxically by the enthusiasm with which this example has been misreported as an inheritable language disorder. From early on in

life members of this family suffer a devastating neurological dysfunction that requires many of them to communicate by sign-language. The disorder affects the coordination of orofacial musculature, both for nonlinguistic uses and for speaking. The syntax problems of these people are reported as their primary problem by those seeking a gene specific for language [32, 33, 67]. Still, recent clinical reports upon this family stress that this claim is false since virtually all aspects of their expressive language - from syntax to articulation - is found impaired [30].

However, let us suppose - for the sake of argument - that specific language abilities are genetically encoded. Would such genes increase the reproductive success of better speaking individuals? There are several problems, some of which are addressed below.

### 3.2.3. Individual vs group fitness

First, the advantages offered by speech, like more successful hunting, would have benefited all individuals in a hunter-gathering band of early humans, even those with less well developed language capacities.

Indeed, better communication possibilities favour the group (or the species) as a whole and as such it seems implausible that natural selection, which works on differential reproductive success of specific genes, could have worked at all. Group selection means here that reproduction of all genes present in a group is influenced in a similar manner by newly developed behaviours. Accordingly, Allott [3] notes:

*"However, in the case of humans there can also be cultural selection, behavioural selection at the group level, where the patterns of behaviour adopted are not tied to individual genetic differences."*

Although Ridley [70] has convincingly argued that in nature group selection usually is a much weaker selective force compared to natural selection, in cases where memetic information transmission plays a role (like in imitating/learning behaviour) we might understand easily how group selection could play its role in evolution. Changes in individual behaviour - regardless whether these changes are genetically encoded or not - can be taken over by other members of a group. In case this behaviour happens to confer some selective advantage, every member of the group can quickly profit of this, regardless the genetic make-up of the individual. Certain groups as a whole then may be favoured since they acquired some behaviour. Therefore an initial coincidental link will exist between certain behaviours (memes) and the collection of genes which happen to be present among the group members with this behaviour. So, all individuals and genes will be favoured, and this will obscure natural selection for the gene which possibly led to the successful behaviour. Moreover, many complex social behaviours do not have a genetic basis but can originate as coincidental inventions of some individual (see the example of Japanese macaques below and our remarks on language genes (3.2.2)).

Taken to its extremes, Pinker's claim that the complexity of language arose as a gradually selected feature compares to stating that our tool making refinements were a consequence of genetic selection. As such, people who had a mutant gene which gave them the possibility to keep a fire burning were reproductively more successful than those without the gene. A later

mutant enabled some to make fire by firestones and his/her offspring was reproductively more successful than people who did not possess this genetically encoded capacity, because it is obviously more advantageous to be able to make your fire yourself whenever you want to. Thus the gene for making fire with firestones spread in the population and outcompeted the 'keep the fire burning' gene. Of course, the genetic mutants which could make matches were better off and outcompeted the firestone firemakers through better reproductive success. But alas, present day mutants which use lighters are competing out the match using genetic mutants.

The silliness of this argument is obvious. Still, this is largely what is being claimed by the gradual natural selection approach of Pinker about the origin of language. Moreover, it is even far more difficult to explain selection of language this way than it is to explain tool use (see 3.2.7).

### **3.2.4. Can language be explained by assuming better reproductive success of better speakers due to an increase of social status?**

Even in the case of genetic encoding of cultural phenomena like language, and even in case where individuals gain a higher social status which results from their socially highly valued and (for the sake of the argument) genetically encoded cultural behaviour (but see 3.2.5), this new skill which is first owned by an individual with a mutant gene or a novel recombination of genes, must remain hidden to other members of the species. Because of mimicking capacity, other members will readily copy the art such that the eventual higher social rank brought by the new trick and which might lead to more successful reproduction, is readily lost. The mutant parent even must hide this skill for its own offspring, otherwise both mutant and wild type offspring will take advantage of it, and again no natural selection will be possible.

The example of the Japanese macaques who readily adopted washing sand from sweet potatoes as it was first done by one member of the group is well known (see [Note 1] for remarks on imitation). Whether or not this single group member had some gene for this behaviour (which we heavily doubt), the gene could not lead to higher social status as a result of the behavioural change it introduced, since several group members readily behaved the same way.

### **3.2.5. Does better speech production really bring along a higher social status?**

Pinker [67] claims that people with better speech capacities have more chances to acquire a higher social rank, becoming a tribe leader or politician, and from this it is inferred that they will have higher reproductive success leading to spread of genes for better speech.

However, there are several pitfalls in this line of reasoning. First, one should not confuse current macrosocial politics - where indeed leadership often has to do more with ones' public image, which indeed partially depends on ones' linguistic capacities - with the original small tribe policies. Under these original conditions, being the leader is often the mere consequence of having a father who was the previous leader, regardless one's (linguistic) skills. Second, one can question whether it were especially better speech capacities which led to high social rank and thus reproductive success. Being a successful hunter, a good parent, an efficient food gatherer, a socially enjoyable person (which depends not necessarily on speech capacities), a very

aggressive and physically strong male or a good singer or a sexually attractive partner, are all other and probably more important reasons of why an individual could be reproductively successful. Physical attractiveness may even be a more important reason for reproductive success than social rank in humans (see 3.2.6) and later on we will argue in favour of the attractiveness of male singers - above male orators - on females (see 3.2.5). Whatever, it appears that natural selection will be too weak a force to explain language by the social status it might provide, since speech happens to be only one of many possible other factors which determine social status.

Anthropological studies moreover show that in the small hunter-gatherer bands, the 'big man' is not distinguishable from the other members of the band [25, 26]. To quote Richard Lee upon the hunter-gatherer !Kung: 'None is arrogant, overbearing, boastful, or aloof. In !Kung terms these traits absolutely disqualify a person as a leader and may even engender forms of ostracism... Another trait emphatically not found among traditional camp leaders is a desire for wealth or acquisitiveness... Whatever their personal influence over group decisions, they never translate this into more wealth or more leisure time than other group members have' [47]. The kind of social organisation (tribes and kingdoms) which Pinker has in mind and where speech eventually might have increased social status, eventually (but doubtfully, see 3.2.6) leading to a minor reproductive advantage, does only exist since about 10 000 years, well after the origin of language [25, 26].

### **3.2.6. Does higher social status ensure reproductive success?**

There is strong evidence - for example from analogy with social animals - that reproductive success is indeed closely linked to social rank.

However, just in case of humans - where this link between social status and reproductive success is needed most to supply natural selection as an explanation for language - it may not strictly be applied.

Humans, living in fission-fusion societies with strong pair bonding and with prolonged periods of absence of the males, appear to be a special case. It has been suggested that females indeed do prefer partners for life with high social rank, thus ensuring material advantages for raising offspring, but that they try to choose physically - genetically attractive partners for sexual reproduction. Strong evidence for adulterous behaviour of females - at least in original human tribes - comes from the many highly complicated adaptations of both female and male reproductive behaviour at the level of oocytes and spermatozoids. For instance, it has been shown that males produce killer spermatozoids - able to kill spermatozoids from other males - and that these are produced especially when there may be suspicion of adulterous behaviour of females (for instance after long absence of the male). On the other hand, it appears that the female body can regulate which sperm of different partners is preferentially taken up, for instance by - subconscious - regulation of orgasmic experience [5].

In conclusion, the high social rank of a human male not necessarily confers absolute ensurance for better reproductive success.

### **3.2.7. The development of language by natural selection is even more difficult to explain than other cultural characteristics like tool making**

For language the problem for a natural selection explanation is even more difficult to overcome than it is for other cultural traits like tool making. Not only producers must be selected, but at the same time very different mutations - those mutations which enable understanding of what producers say - have to be selected gradually and naturally. We have previously pointed to the same bottleneck in the explanation of behavioural mate recognition systems [93] and this problem for language has been formulated also by Geschwind [31]. The arguments by Pinker & Bloom [66] to resolve this paradox, are far from convincing, and finally they have to rely on the Baldwin effect (see also Deacon [21]). It appears an odd supposition to state that better story-tellers will gain high social status, when one has to explain how 'better story understanding' genes have to be selected independently at the same time.

### **3.2.8. Summary**

In summary, there is as yet no convincing evidence that language genes exist (see 3.2.2). Second, in cases where behaviours can be inherited by learning and mimicking, natural selection for individual genes can be a weaker selective force than group selection, whereby group selection favours all genes in a social group indifferently. Natural selection can explain the increase of general abilities: better vision, higher intelligence, better singing capacities. When it comes to explain how specific, directly observable and mimickable abilities - like speech and like making tools - can be selected, group selection becomes important enough to counteract or overwhelm natural selection, because new findings of individuals will be taken over by others, whatever their genes (see 3.2.3). For the same reason of mimickability, it is unlikely that mimickable behaviours will lead to higher social rank (3.2.4).

Third, trying to explain how language could be naturally selected by assuming that better speech entails a higher social status which in turn leads to reproductive success, can be criticized by showing that speech was (and is) only one of several factors in determining social status (see 3.2.5), and that it is uncertain that social status of humans is a guarantee for reproductive success (see 3.2.6). Fourth, it should be noticed that there is not really a link between being the 'leader' of a hunter-gatherer band and social status or reproductive success (see 3.2.5). The example of Symons [87], adopted by Pinker [67] on the reproductive success of tribal chiefs (often both gifted orators and highly polygynous) then is not really applicable to the humans which first developed language. Finally, it is difficult to see how natural selection for better speech could work, when realizing the difficulty of selection for better speech understanding to occur simultaneously (see 3.2.7).

## **3.3. Terrence W. Deacon: Did language and brain co-evolve, leading to both a larger brain and better speech capacities?**

The hypothesis of Deacon [21] is well summarized by the following quote:

*"Considering the incredible extent of vocal abilities in modern humans as compared to any other mammal, and the intimate relationship between syntax and speech, it should not surprise us that*

*vocal speech was in continual development for a significant fraction of human prehistory. The pace of evolutionary change would hardly suggest that such an unprecedented, well-integrated, and highly efficient medium could have arisen without a long exposure to the influence of natural selection. But if the use of speech is as much as 2 million years old, then it would have been evolving through most of its prehistory in the context of a somewhat limited vocal capacity. It is during this period that most predispositions for language processing would have arisen via Baldwinian evolution. This has very significant implications for the sorts of speech adaptation that are present in modern humans."* (Page 358-359).

Here Deacon [21], like Pinker [67], relies on long term (2 million years) gradual evolution through selective advantage offered by the use of symbols, and he relies on Baldwinian evolution. It should be noted however that Deacon [21] clearly dismisses the notion of Universal Grammar [16, 17].

For the moment, it will suffice to say that we claim that our large brains did not expand to enable language and that language did not cause brain expansion. For example, microcephalics [77] and individuals with only half the brain of normal humans and so with brain masses within the upper limit of nonhuman primates - can learn normal speech [81]. It might help, but you do not need a large human brain to be able to speak. Furthermore, the archaeological evidence indicates a late origin for language [58]. For further comments on Deacon, see 6.2.

### **3.4. Conclusion**

We have summarized some possible criticisms on the most renown hypotheses on the origin of language and we have indicated why we have difficulties in accepting the existence of some kind of Chomskyan 'universal grammar' (except for some basic mental syntax which we share with animals (see 2.2.2)) and why a genetic explanation, adaptationist [67] or assimilative [21] seems implausible to us. There are several other criticisms possible [3, 82, and 89].

Still, this denial of a direct role of natural selection in the origin of speech and the arguments in favour of a cultural evolution process to understand the origin of language do not supply us with the concrete genetic preadaptations we need to understand how both production of symbolic sounds (vocal flexibility) and giving structural value to words in sentences (linguistic syntax) have been achieved.

We will try to answer the first question on vocal flexibility largely by evolutionary considerations about the phylogenetic origin of language (section 4), while the second question will be approached in an attempt to understand how infants acquire language (section 5). It is now finally time to readdress one of the most ancient explanations for the origin of language: our musicality or singing capacity, which is essential in explaining both the phylogenetic and developmental origin of language. Both the origin of language and its development in children, we argue, can be best understood by recognising that we are musical or singing primates in the first place.

## 4. Phylogenetic origins of spoken language

### 4.1. Protolanguages

Are there cues to protolanguage in close relatives of ours? Burling [13] states: "Since our surviving primate communication system remains sharply distinct from language, it is implausible that it could have served as the base from which language evolved. We are more likely to find hints about language origins by studying how primates use their minds than by studying how they communicate." The same conclusion was reached by Jonker [42]. This is not in contradiction with our claim for some universal mental syntax among higher animals (see 2.1.2).

However, we should mention that opinions differ:

*"The analysis of the so-called long calls in chimpanzees and bonobos make it likely that the group-living great apes preserved the ability to create syntactically different calls, which would be developed by requirements of social life. A call repertoire emerged in these species which contained a large number of call variants at group level available for each group member via social learning. This type of animal call is different from ordinary animal communication; it shows some features of human language."* [91].

There is also some controversy with regard to the speech capacities of our ancestors. Is speech already present in *H. erectus*? Since when can *H. sapiens* (which originated about half a million years ago) speak. Could *H. sapiens neanderthalensis* speak?

The archaeological evidence indicates that planning and other complex activities date back at the earliest perhaps 60,000 years ago. Noble and Davidson [63] argue that increasing tool use capacities, the occurrence of cultural artifacts (paintings, statues), and burial practices follow from the mental activity enabled by language. Such behavioural evidence of language starts with the Upper Paleolithic around 40,000 years ago. Maybe significantly, it was at this time that anatomically modern humans started to replace Neanderthals which only became extinct between 40,000 to 32,000 years ago. Others also argue in favour of a late origin of vocal language [58].

Here, we adopt the point of view that spoken, symbolic language is quite different from primate languages and that it originated only recently.

### 4.2. The musical primate

#### 4.2.1. Humans have unique adaptations for singing.

The idea that the origin of speech lies in our ability to sing can be traced back to at least Jean Jacques Rousseau, in the seventeenth century [73]. It was suggested by the famous linguist Wilhelm von Humboldt in the nineteenth century [94] and by Otto Jespersen early in this one [41]. However, this approach to language has been ignored in more modern times. Indicative is that the word 'music' lacks in the index of the recent books of Pinker [67] and Deacon [21]. In

recent times, music has received serious attention by some linguists [48], but this was done within the Chomskyan paradigm and did not address the origin of language.

Just like song birds possess highly sophisticated syringes, there are very characteristic morphological changes of the human glottis and larynx, unequalled in any mammalian species [75]. Aitchison [1] remarks: "*Our language has more in common with the singing and calling of birds, than with the vocal signals of apes.*"

The resemblance to bird song was noticed already by Charles Darwin [19]:

*"(Language) is certainly not a true instinct [Note 4], for every language has to be learnt. It differs, however, widely from all ordinary arts, for man has an instinctive tendency to speak, as we see in the babble of our young children; whilst no child has an instinctive tendency to brew, bake, or write. ... The sounds uttered by birds offer in several respects the nearest analogy to language, for all the members of the same species utter the same instinctive cries expressive of their emotions; and all the kinds which sing, exert their power instinctively; but the actual song, and even the call-notes, are learnt from their parents or foster-parents. These sounds, ..., are no more innate than language is in man."*

Provine [68] has shown that a unique overlooked feature of human speech is our ability to integrate respiration and vocalisation. We, as humans, breathe in a way unique among the primates - since only we can neurally modulate sequences of tonal vocalisations upon our expirations. Other primates can vocalise but they are limited to only one vocalization per expiration. For example, both humans and chimpanzees laugh: however, chimpanzees do so by an 'ah', 'ah', 'ah' sequence of repeated inspirations and expirations. In contrast, we do a modulating 'ha, ha, ha...' or 'ho, ho, ho ...' upon a single out-breath - this modulation often going on continuous for 16 laughter syllables [68, pp. 40-41]. Moreover, we can subtly tune our series of vocalisations upon a single continuous out-breath. Only amongst birds - not other primates - are there species that possess comparable respiratory-control ability. This underlies the curious fact that while some birds can imitate human speech, the much more closely related chimpanzee or any mammal cannot.

The neural control that allows song was, we suggest, a profound revolution: the 'one breath one-vocalisation' rule stops chimpanzees not only from laughing like humans but also from being able to control the expiration needed to speak. This, as Provine [68] notes, is the reason why attempts to teach spoken language to chimpanzees have failed in spite of them being able to learn sign and token-based languages and even to understand spoken speech [76, p. 40-41].

Neural control of respiration allows many more kinds of vocalizations: over 700 vowels, diphthongs and consonantal phones were found in a sample based upon only one-twentieth of all the world's languages [52]. Moreover, such control allows the concatenation of very complex sequences. Thus, vocalisations upon a single out-breath combine into words, and these in turn combine into clauses, phrases and sentences. Neural control also allows modulations to be superimposed upon these vocalisations, such as intonation (linguistic, pragmatic and emotional), and this can be upon a wide variety of speech types like whisper, song, chant, scream, motherese, 'Donald-Duck speech' and ventriloquism.

The tonal modulation of song is not only enabled by neural control but also by anatomical specialisation of the vocal tract for producing a wide variety of pitches and timbres. The peculiarity of our vocal tract is usually attributed to enabling speech, although it is sometimes also considered as a mere consequence of postural changes between the head and thorax that accompanied the upright stance and human-style bipedal locomotion (see also the postscript). However, the anatomical characteristics of the vocal tract are more closely linked to our capacity to sing than to our capacity to speak. People cannot sing without fully using all their vocal tract. However, people can speak without using large parts of the vocal tract (for instance in buccal speech, more familiarly known as Donald-Duck speech). Although normal speech contains a range of vowels and consonants that fully exploit the vocal tract, sufficient variety amongst the world's languages exists to suggest that intelligible speech only needs a subset of possibilities, exploiting only part of the vocal tract's pronunciation potential.

Without the neural control that enables song, speech could not exist. But which came first? We argue that we can speak because we can sing, and not that we can sing because we can speak, also for parsimonious reasons: the capacity to speak requires in addition to respiratory control also syntax, phonology and the capacity to use and learn a vocabulary of words (see also the remarks in 3.2.7), while singing requires none of these (songs can exist without words). Second, in the development of speech by children, melody - in terms of interest in and production of intonation and rhythm - comes before other aspects such as phonology, syntax and vocabulary (see section 5).

#### **4.2.2. Musical primates and song birds: examples of convergent evolution?**

##### **4.2.2.1. Introduction**

The exact reason for the origin of singing behaviour is beyond the scope of this paper, but it is clear that the ability to sing has been naturally selected on many separate occasions - e.g. birds, whales and gibbons. Where this has happened, there have often been highly complex adaptations both anatomical and neural. The major idea here is that the complex changes which were necessary to develop an organ which eventually could be used for symbolic language production were selected for singing and not for speech. Convergent evolution to what may have happened to modern humans can be observed in song birds. Also song birds developed highly complex adaptations, anatomical and neural, as a result of natural selection for better song capacities [Note 5].

##### **4.2.2.2. Music and mating**

Song production and song preference play an important role in mating in song birds. Possibly music had a similar role originally in human mating - and it still has to some extent. Below are some of the several possible examples of the central role of music in courting behaviour. In several cultures males indeed bring serenades for their beloved. Also, male singers and musicians in general exert strong physical attractiveness on females (some females even have orgasmic experiences during concerts). Much poetry and love texts sound silly when proclaimed, but are quite acceptable and even touching and convincing when sung. Adolescents meet through singing, listening to music and dancing.

Moreover, sexual selection of the ability to sing is more plausible than sexual selection of the ability to speak. Sexual selection requires only an inherited preference for singers of distinctive emotional melodies rather than good story telling - something that requires that language itself is first well understood.

However, it might be objected that this fails to explain why females would also sing and speak. It should be noted that, while it is true that in many song birds only males sing, females inherit genetically the abilities to sing - something that can be shown since female singing can be triggered by hormonal treatment. Therefore, it is evolutionarily possible that a small genetic change triggered hormonal changes so that singing by females became possible, after it had first been sexually selected for in males. From considering some tropical song birds, we might understand how song capacity of females might have been selected for, eventually after it arose in males by sexual selection first.

#### **4.2.2.3. Music and bonding**

Indeed, the situation whereby male song birds exclusively sing happens to be true only for temperate regions. In some species of tropical song birds, females, as much as males, can engage in singing. Moreover, unlike in temperate areas, where male song links to the defense of territory and attracting potential mates, in these tropical species male and female singing links to bond formation and bond maintenance. This becomes apparent from the following quote [88]:

*"In the tropics, although there are many species of birds the song of which is doubtless just as territorial in function as is usual in the temperate regions, the ornithologist is also struck by the number of examples where song appears much less aggressive in intent and where its function is apparently as a social signal, for maintaining pair and family bonds and as part of the sexual display, rather than a territorial one. Moreover, it is perhaps significant that most of the outstanding vocal imitators are found among tropical or subtropical species."*

Thorpe & North [88] give the example of a pair of birds which communicated via a 15 note antiphonal duet. However when one bird died the survivor resumed the performance of the whole - something it had never done previously! They note of another case of duetting, reported elsewhere, that '*when the partners were absent, the remaining bird would use the sounds normally reserved for his partner, with the result that the partner would return as quickly as possible, as if called by name*'. This strongly suggests that we witness here a real case where song is used meaningfully in social communication as a bonder. On top of that, the vocal tract of these birds has attained such sophistication, that it enables them to imitate human speech.

Music has bonding function in close relatives of ours as well. As noted above, male and female Siamang sing (the male bitonally and without melody; the female monotonously) to establish and maintain pair-bonding and the social recognition of their territory [38].

The cue of the use of song as a bond strengthening means of communication, rather than song being a trait which has evolved by sexual selection alone, itself leads to some intriguing remarks with regard to the special 'sociological' case humans are among primates (and animals/mammals in general). We know, by comparing the social nature of humans with other apes, that we too

have evolved an unique capacity to bond with each other. Indeed, it is also in the depth and complexity of our bonding that humans differ (apart from language) from other primates. From these observations and considerations we are tempted to conclude that musicality not only can explain how symbolic language evolved, but also that song, as a means to aid bond formation, can help to explain how the characteristic sexual-social relationships between humans became possible (see also 4.3.2).

### **4.3. Present and past social music**

#### **4.3.1. Various observations**

What evidence exists for the key role of music in the lives of humans? Below we give a very limited excerpt of the functions and possibilities in human social life. All human cultures possess lullabies and use them to sing children to sleep. The music business is among the world's major industries. Going to war is so much more fun with a drum band marching along. Dancing to music can give people mystical trance experiences. Music brings up deep emotions such as hope, pleasure, comfort or sadness, and probably no other 'art' can do this as profoundly as music. From observations of currently existing 'premodern' societies, it is clear that music (and its counterpart, dance) must have played an even more important, pivotal role in early human societies. Music has a role, not only in rituals, but also in many practical activities. For instance, Australian aborigines memorize the look of landscapes in songs. Although the music making of early humans has left no physical remains, it must have been a major part of their lives, as it still largely is an essential part of our lives.

#### **4.3.2. Music and group identity**

There is the observation that rituals, dance and song enhance group identity. With respect to territorial behaviour, it should be noticed that singing is indeed used for that purpose in close relatives of ours: "*In addition to the well-known territorial bird songs, some monkey species and all species of lesser apes have territorial songs.*" [91].

From what we know about ourselves as apes, increasing group identity could have put strong evolutionary pressure on singing behaviour. To understand this we must digress upon what has recently been found about our uniqueness as social apes. Humans, chimpanzees and presumably our earliest shared ancestors mix a life-style of belonging to a group, while separating into smaller parties during much of the lives. This is called an atomistic or fission-fusion social existence [72]. We, however, do so in a way that is unique because the bonds are robust and long-term and allow for long periods of separation. Biological parents in all human societies form bonds with each other (though not necessarily monogamous ones). People form life-long attachments with friends and distant kin. We, moreover, usually form a life-long attachment with our 'identity group' from the level of our extended family to that of our nation and religion.

Early humans faced the paradoxical problem of relying for survival both on a group and on the recurrent need to split-up. Anthropologists and historians identify the mechanism by which people create and sustain the required social attachments with rituals and group activities involving synchronised song and dance [54]. The need for sustained social bonds may have

further selected (after possibly initial sexual selection and selection for stronger pair bonding (see 4.2.2.3)), for dance and song competence [Note 6].

Modern remnants of this ancient function of music might be the supporters' songs of sports teams, songs of any kind of club (e.g. students), war music and the national hymns, closely linked to the notions of territory and group identity. Indeed, music, singing and dancing still plays the central role in social life of all extant original bands. Ceremonies, rituals, and many other group activities (work-gangs, parties, festivals) all exploit the strong emotions which come with the ensemble of vocalisations and movement. Just think of the emotional bonding, the sense of belonging, the experience of 'together we are invincible' that accompanies marching songs, football stadium chants, National Anthems, camp-fire songs, hymns, corals, etc.

Increasing group identity exists in a nonmusical form in the collective intoning and synchronisation of bodily movements in religious prayers, petitions, supplications, orisons and worship. In modern societies, such synchronisation offers people a temporary sense of belongingness. In most cultures, they form an important part of rituals, ceremonies and other shared enjoyments which result in the affective togetherness that creates and sustains a society's collective existence [10, 54, 78].

#### **4.4. Conclusion**

Whatever the role of early singing was (territorial marking, courting, pair bond maintenance, enhancing group identity) it is clear that singing, musicality and dance had an important role to play in human social interactions, and that consequently musicality is plausibly selected for by good old natural selection. The development of a complex organ like the *Homo sapiens* vocal tract then can be understood to have been developed by natural selection more easily than in case we have to hypothesize that this natural selection occurred on the basis for selection of better speech [21, 67]. Only later on, these vocal abilities were used for speaking, and this view coincides with the proposition of Gould & Lewontin [34] that language is a spandrel or an exaptation: language was possible because of a preadaptation which developed for other reasons. While singing is an innate capacity, an instinct, speaking is a possibility emerging from singing and increased mental representational capacities. We could better speak of the song instinct than of the language instinct.

Comparing the role of song in some tropical song birds and in the siamang, one is tempted to state that song co-evolved with pair bonding, and thereby also helps to explain how the intriguing social and sexual characteristics of human life evolved.

### **5. Melody and language learning by children**

#### **5.1. Introduction: the semantics of spoken syntax**

Do humans have a language acquiring device as Chomsky has proposed?

Most students of language easily accept that semantics is about linking mental representations of objects and concepts to the symbolic lexicon that happens to be used by a language. However,

when it comes to syntax, most linguists seem to assume that there is only linguistic syntax. Above we have argued that all higher animals possess some universal mental, thinking syntax (see 2.2.2), while on the other hand it is tremendously difficult to discover any universality among the amazing diversity of spoken syntaxes (see 3.1).

The problem with spoken syntax therefore boils down to the same problem of linking lexicon to mental representation: semantic meaning must be given to spoken syntactic entities by linking them to the mental syntax. We think this approach has been overlooked by most students of language. Then one must wonder how this can be achieved, since spoken syntax can be any kind, while mental syntax can be supposed to be largely alike among humans - and basically even among higher animals.

## 5.2. Music and language development

One of the big mysteries in speech acquisition is how children identify words. While the words on this page are divided by spaces, spoken words are not. Before you can identify words you have somehow to identify where and when they start and end. Failure to solve this hard problem holds back artificial speech recognition. The earliest voice recognition programs required that people spoke words slowly and in isolation. We suggest, backed up by a growing research, that infants solve this problem by listening to the rhythmicity and to the melody of stresses and tones in speech.

Even before children are born, their brains are familiar with the sounds that will surround them after birth. Newborns prefer the voice of their mother over that of strangers [23]. If a mother repeats a short story twice a day for the last six and half weeks of her pregnancy, her newborn child will prefer hearing it to one she did not [24]. The womb is an acoustic filter that preserves the intonations of a mother's speech. Thus, the brain is learning to hear speech as a melody from long before birth.

This is supported by other work upon newborns. It has been shown that newborns can discriminate the rhythm of multisyllabic stressed words suggesting they are already sensitive to the word-rhythm [74]. Moreover, newborns already prefer infant-directed prosody stressing speech (motherese) over adult-directed speech [18]. Complementary to this, mothers expand the intonation contours of their speech to their child as soon as it is born [28]. Such motherese compared to adult-directed speech has emphasized prosody, namely higher overall pitch, wider pitch excursions, broader pitch range, increased rhythmicity, slower tempo, longer word durations and increased amplitude. Newborns moreover can distinguish their own language from a foreign one, something which must be due to the unique, prosodic cues of a language [55]. This suggests they are increasingly able to focus upon the unique intonation aspects of their 'mother' tongue.

Children's own vocalisations, it should be noted, also start to be affected by these intonations:

*"A cross-cultural investigation of the influence of target-language in babbling was carried out. 1047 vowels produced by twenty 10-month-old infants from Parisian French, London English, Hong Kong Cantonese and Algiers Arabic language backgrounds were recorded in the cities of*

*origin and spectrally analysed. ... Statistical analyses provide evidence of differences between infants across language backgrounds. These differences parallel those found in adult speech in the corresponding languages." [22].*

There is also the observation of the tremendous similarity of pronunciation within a slang. We all know the phenomenon that one can easily recognize the region where one comes from. Many people never succeed in speaking properly the standard language because of an ineradicable accent, which indicates the thorough imprinting which occurs: we do not only acquire lexicon, we mimic intonation almost exactly from our environment [Note 7].

The previous paragraphs lead us to suggest that some auditory equivalents to Rizzolatti-cells (see Note 1) must exist. Rizzolatti-cells and equivalents may be an important cue to understanding mimicking, to link the behaviour of genetically encoded cells to copyable observable (visual, auditory) behaviour of animals.

Intonation provides cues to how words are structured in sentences [59]. Words are not said uniformly but are intonation phrased. Spotting this intonation structure facilitates children to grasp how words are syntactically put together. Children use the intonational cues that tend to identify word beginnings [62]. These cues vary with language: stress for example in English, syllable in French and mora in Japanese. Children in all these languages develop a sensitivity to the intonational beat provided by these cues that mark off word separation.

Let us take a famous Chomskyan example which relates to inversion of the word order of statements in order to turn these into a question. Children with English speaking parents readily adopt that 'The man is here.' becomes a question by reversal of noun and verb: 'Is the man here?' But how does one turn the slightly more complex sentence: 'The man, who is tall, is here.' into a question? One might expect a child, who has just mastered the simple example to place the first 'is' in front of the sentence, to say: 'Is the man who tall is here?' But children never make this mistake. Do we need Chomskyan theory, borrowed from mathematics and logic? Linguists developed rather complicated theories (like X-bar theory) whereby humans use 'null' elements to cope with this and related problems (see Smith and Szathmary [83] for a brief explanation).

What if we adopted the answer that children simply hear which of the two verbs is the main verb. Say the complex sentence to yourself and listen how the intonation on the second 'is' is different from the first 'is'. Now, try to reverse intonations. It requires a little exercise to do so, since it is experienced as a very 'unnatural' (we should actually say 'uncultural') thing to do, which by itself provides circumstantial evidence on the importance and the strict use of intonation. Children just hear which verb is the one which goes along with the man, because of the intonation of the main verb. Remark that the pitch of the main verb in the complex sentence is exactly the same pitch the main verb carries in the simpler sentence AND in the question. Once this has been acquired, children can generalize this principle to any similar sentence they meet. The intonation recognition capacity is one which stems from our innate musicality (naturally selected recently for other reasons than language itself), while the generalization capacity is part of the mental syntax capacity which we have inherited from animals (naturally selected for still other reasons). Bringing the two together one can have something like syntactic symbolic language.

Thus, children start off experiencing language as a kind of music. Parents and others respond to this sensitivity by making their language to them more musical - motherese. The rhythms of speech, which are heightened by motherese, provide the child with a means to use their sense of rhythm to spot the words and sentence structure. Memetic ontology thus replicates memetic phylogeny. In other words: music is both the answer to the phylogenetic and to the developmental origin of language.

Children, before acquiring the language spoken around them can distinguish phonetic categories of foreign languages they have not heard [27, 86, 90], only to lose this ability at around ten months [4]. One wonders why children should have this ability, in case language was naturally selected for, since this would require only the evolution of recognition of a limited phonological set. While explaining this from a 'natural selection for language' point of view is a real conundrum, it becomes triviality when adopting an innate sensitivity for melodizing.

Also, there are the numerous reports on the application of Music Intonation Therapy [2] to treat language disorders, as is exemplified by the quotes below:

"In order to develop a useful communication system, a 3-year-old, non-verbal autistic boy was treated for 1 year with a Simultaneous Communication method involving signed and verbal language. As this procedure proved not useful in this case, an adaptation of Melodic Intonation Therapy (signing plus an intoned rather than spoken verbal stimulus) was tried. With this experimental language treatment, the patient produced trained, imitative and, finally, spontaneous intoned verbalizations which generalized to a variety of situations." [56]

"We examined mechanisms of recovery from aphasia in seven nonfluent aphasic patients, who were successfully treated with melodic intonation therapy (MIT) after a lengthy absence of spontaneous recovery." [7].

*"In patients with brain lesion, a pre-verbal, emotionally-focused tonal language almost invariably is capable of reaching the still healthy sections of the person. Hence, it is possible for music therapy to both establish contact with the seemingly non-responsive patient and re-stimulate the person's fundamental communication competencies and experience at the emotional, social and cognitive levels."* [43].

Furthermore, it has been shown that music is not only important for developing linguistic skills, but also serves as a memory aid [65] and plays a role in the development of motoric skills [12]. Strong suggestions for the existence of a music acquiring device comparable to the hypothetical Chomskyan language acquiring device have been made by others. We claim that this MAD **is** our LAD:

*"Full-term infants' performance in detection of melodic alterations appeared to be influenced by perceptual experience from 6 months to 1 year of age, and an experiment with infants born prematurely supported the hypothesis that experience affects music processing in infancy. These findings suggest parallel developmental tendencies in the perception of music and speech that may reflect general acquisition of perceptual abilities for processing of complex auditory patterns."* [51].

*"This indicates the existence of a partly innate and partly acquired competence to judge what is acceptable and what is not, within the tradition of Western popular or classical music. This seems to indicate the existence of some deep structure of tonality, comparable with Chomsky's deep language structure. Asians who have not been much exposed to this kind of music find the task very difficult."* (Kalmus & Fry [44], reporting on experiments whereby subjects were asked to evaluate some characteristics of Western classic music).

The last sentence from the previous quote is again a strong indication for the importance of the tonality of the language and the music of a child's culture in molding its innate recognition capacities. Depending on the culture, one's experience of what sounds acceptable and what is not, is completely different (by itself again an indication against a universal spoken grammar and natural selection of language). This is nicely illustrated by the fact that (Western) MIT therapy has to be adopted when it is used in an Asian country. When applying MIT for use with Japanese patients, the authors report that basic changes were necessary, because of the completely different 'pitch' of Japanese language [79].

Furthermore, Simmons & Baltaxe [80], studying adolescent autistics with linguistic impairments, suggested that:

*"... Perception of prosodic features may be crucial for decoding and encoding linguistic signals. Autistic children may be lacking in this ability."*

## 6. Final conclusions

### 6.1. Summary

We argue that a combined genetic and memetic explanation is needed to understand what language is about and how it developed originally and develops with almost every new human. According to the point of view presented here, symbolic, spoken language emerges from the (coincidental) combination of complex representational capacity with intonation recognition/reproduction capacity (which itself develops in close connection with singing capacity). As such, it is claimed that it is not language itself which has been naturally selected for. Language is considered as a cultural phenomenon very well comparable to bird song culture, only more sophisticated (variable, flexible, more symbolic, syntactic) just because of the more sophisticated mental representation capacities of higher apes. In summary, birds did not develop symbolic language to the extent that humans did, because of more limited mental representation capacities, chimpanzees did not because of lack of singing capacities. Humans simply happened to combine both characteristics. How language then can develop by memetic evolution, might partially be answered by work presently being done with interacting robot agents [84], and is the subject of further work [82].

Once the preference for sound variety has been selected for, something which may happen for various reasons and which has occurred independently in different animal taxa, individuals which can produce any kind of primitive song may be reproductively more successful through sexual selection. Moreover, the group of singing and dancing individuals as a whole, whatever the genetic make-up of the individuals, may become more successful because of the increased

group identity awareness which makes its members cooperate more efficiently or which may make the members lose their individuality to some degree, resulting e.g. in more fierce, aggressive behaviour with regard to non tribe members. Indeed, another typical characteristic of humans is our long tradition of warfare and genocide [25].

## 6.2. How does this approach compare to the hypotheses put forward by Chomsky [16, 17], Pinker [67] and Deacon [21]?

With respect to the development of language in children, one can agree with Chomsky that humans have special abilities to adopt language and syntax very spontaneously early in childhood and this can be called an innate language acquiring device. Still, it probably might best be understood as an innate music acquiring device, which enables to link any possible syntax of spoken language - the one used by the adults which happen to raise the child or by other children which happen to grow up with the child - to the universal mental syntax, of which we share the general basic possibilities for categorization and for generalization of causal rules with animals.

We do not agree with the Chomskyan suggestion, taken for granted by Pinker [67], but thoroughly criticized by e.g. Allott [3], Deacon [21] and Tomasello [89], that there is such a thing as universal linguistic grammar.

Furthermore, the explanation of the origin of language in evolution and during individual development, as proposed here, has nothing in common with the adaptationist explanations of Pinker (see 3.2). Not only Allott [3] and Tomasello [89] point to different shortcomings of this kind of reasoning, but also Deacon [21] has clearly indicated several flaws. Several other criticisms are possible [82]. What Pinker [67] calls a 'boring conclusion', is simply a completely erroneous conclusion.

We can largely agree with Deacon [21] that we are a symbolic species, and his evolutionary reasoning is much more relevant than that of Pinker. However, Deacon [21], like Pinker [67], relies on long term (2 million years) gradual evolution through selective advantage offered by the use of symbols, while instead proposing Baldwinian evolution (evolution by genetic assimilation of behavioural characteristics).

Both our approach and - to a certain degree (because of the pivotal role of symbolic gestures and sounds) - that of Deacon could be called 'memetic'. The difference is that in Deacon's approach gestures and symbolic sounds come into play already 2 million years ago (at the stage of *Homo habilis*) and reshape the brain by genetic assimilation. In our approach natural selection for better general mental abilities and, only recently (possibly with the advent of *Homo sapiens sapiens*), natural selection for musicality explains the reshaping of the brain and the vocal tract and we claim that it is from the combination of increased intelligence and vocal flexibility that language emerges as a cultural process, while we dismiss natural selection or Baldwinian evolution guided by the advantages brought along by the linguistic capacity - as is proposed by either Pinker [67] or Deacon [21].

Once humans combined mental capacity and musicality, we rely on genetically encoded flexibility of the brain to explain how symbolic sounds - memes - could develop and restructure

brain mapping in a nongenetically inheritable manner. In other words, genes provide general capacities like brain flexibility, vocal dexterity, intonation recognition and reproduction capacity, while memes - through interaction with the developing brain - strongly influence the rewiring of the neuronal connections which make up a brain.

Although we date the influence of symbolic sounds much later than Deacon [21], we claim that once they originate, further changes occur in an almost purely memetic manner. The example below of the differences between literate and illiterate persons indicates how influential the means of communication are with respect to our mental abilities.

Our musical language origin theory coincides best with "the idea that removal of vocal limitations released untapped linguistic abilities which has been a major theme of a number of language origin theories (most notably argued by Philip Lieberman, in a number of influential books and articles)[49, 50]" (quote from Deacon [21], page 354). Deacon however considers this as an oversimplification and states that: "... *the development of skilled vocal ability was almost certainly a protracted process in hominid evolution, not a sudden shift.*" ([21], page 354), whereupon we disagree, backed up by the archeological record (see 4.1). Our hypothesis provides strong support for the insights of Lieberman [49, 50] (see also the postscript).

### **6.3. Why has the musical origin of language hypothesis been overlooked?**

There is a further intriguing question, in case our hypothesis - which we will defend also on grounds of a more linguistic and neurolinguistic approach [82] - turns out to be a major key in understanding the origin of language. Indeed, one keeps wondering why this obvious, straightforward, and with hindsight even trivial approach to explaining the origin of language has been overlooked by linguists during the last decades. This is even more astounding, first because some of the earliest theories posed that musicality had to lay at the origin of language [19, 73, 94] - even Darwin [19] pointed to the resemblance and second because the importance of rhythm, intonation, melody, etc. in everyday life, in language therapy and in child language (as briefly reviewed above) is so overwhelming, and is well studied.

Several explanations can be thought of. First, there is of course the adaptationist paradigm which keeps us thinking in terms of function, usefulness, and which makes us overlook that usefulness is a post hoc consideration which can only serve as an explanation once the necessary events leading to the existence of some characteristics have taken place. Natural selection can explain why something still exists, but not how it came into being. The necessary variation is not a matter of natural selection, it is a matter of contingency, coincidence, mutation, recombination, symbiosis, evolution of characteristics for other reasons than the ones for which they eventually are useful now (preadaptation, exaptation).

Second, and closely linked to the previous considerations, there is the fact that we all are impressed by the explanatory power of natural selection of genetic characteristics in general, which makes us forget that natural selection is just a special case of selection (see Note 2). Therefore, there is a tradition of trying to explain everything with genes only.

Third, with respect to language, another important bias may exist. It appears that most linguists depart from their considerations from the present form of language, which needs a sophisticated grammar because much of communication is in the form of written code, which lacks the intonation characteristic of spoken language. E.g., writing down a joke may be experienced as an insult instead of as the tongue in cheek remark it was meant to be. In oral communication this will in most instances be clear, because of the facial expression and the intonation. Using written code, we need question marks, exclamation marks or ":-)" (the smile-sign as used in e-mail discussions) to indicate that what we write is meant as a question, an important remark or a joke. Written code, lacking intonation and eye-contact, compensates grammatically for the absence of a shared context with the listener, and finally influences more and more the way we speak, as becomes clear from studies comparing cognitive linguistic capacities between literates and illiterates.

Illiterates - when compared with literates of the same background - have been found to show cognitive difficulties in nonreading tasks such as phoneme awareness [8, 61], repeating nonwords (phoneme sequences that do not pronounce a familiar word) [60], memorising pairs of phonologically related words compared to semantically related ones, and difficulties in generating words which start with a common phoneme sound or which are the names of animals or furniture [69]. Several other studies lead to the suggestion that learning to read and write might not only challenge how people process oral language but also does change the organisation of people's brains [14, 95]. This was already suggested upon nonpsychological and nonneurological grounds [64].

However, most linguists start from the current situation (a literate world) and extrapolate and/or impose our way of thinking, living, interacting, communicating to the illiterate societies in which the original humans lived at the time language originated (see 3.2.5 for a comparable bias), thereby forgetting how different we are because of the completely different memes which populate our brains and because of the fact that the environment we have to cope with is incomparably different to the natural environments in which language first evolved.

#### **6.4. Could large vocabularies alone be sufficient for the development of syntax?**

It is important to quote here recent work of Bates & Goodman [6], which indicates that syntax abilities parallel very tightly vocabulary size over a wide variety of ages. Thus, though children may vary widely with respect to the size of the vocabulary at a certain age (some children acquire words more easily than others), the degree of grammatical competence they acquire is strictly linked to the lexical stage at which they are. This means that two children - one 3-year-old and one 5-year-old, but each with a vocabulary of 200 words, will have both the same stage of syntax.

Bates & Goodman [6] point out the implications of this for language in chimps. Chomskyans make it a slogan that 'animals cannot learn grammar' and hence that 'grammar is unique to the human species'. Bates points out, however, that chimps taught language in fact attain the level of syntactical competence you would expect from human children with the same size of vocabulary. Bates and Goodman [6] state that, if chimps lack syntax, it is not because they lack a human competence for syntax, but because their vocabularies are too limited.

This becomes apparent from the following quote:

*"These differences between grammar and vocabulary are usually interpreted to reflect a qualitative difference in the language-learning abilities of non-human primates (that is, they have lexical abilities, but they lack a 'grammar acquisition device'). That may well be the case; after all, they are not human. However, the data that we have presented here suggest another interpretation: Because the animals studied to date apparently find it difficult to produce more than 200-300 words, symbols or signs, we should not be surprised to find that they also have very restricted abilities in expressive grammar. Consider the developmental relationship between grammar and vocabulary size that we have observed in human children. From these figures, it is clear that children with vocabularies under 300 words have very restricted grammatical abilities: some combinations, a few function words in the right places, the occasional bound morpheme, but little evidence for productive control over morphology or syntax. Viewed in this light, the difference between child and chimpanzee may lie not in the emergence of a separate grammar 'module', but in the absolute level that they are able to attain in either of these domains. Chimpanzees do not attain the 'critical mass' that is necessary for grammar in normal children; instead, they appear to be arrested at a point in lexical development when grammar is still at a very simple level in the human child. Hence, the putative dissociation between lexical and grammatical abilities in nonhuman primates may be an illusion".*

From these considerations, it appears that to explain the rise of syntax, the problem is not how to explain any 'syntax' module arose peculiar to humans, but the problem is to explain why large vocabularies arose. If you can explain that, you can explain the rise of syntax. The solution to the problem of how a large vocabulary could arise, follows from what we suggest: humans are originally musical primates. Once humans gained the neurological abilities to control vocalisation needed to sing, they gained the abilities to create vast vocabularies of words. Although a large vocabulary on its own may be sufficient for syntactical ability to develop, as Bates & Goodman [6] suggest, we think that it helps when you have a MAD, a well developed intonation recognition/reproduction device, at your disposal. Musical ability may explain the rise of a large vocabulary and at the same time may be an extra gain to create and acquire linguistic grammar.

A 'musical origin of language' theory enables to bring together the ideas of Deacon [21], Lieberman [49, 50] and Bates & Goodman [6] (among many others). One could say that at some point, quantity (increased intelligence/mental syntax capacity, increased vocal flexibility, increased vocabulary) may change into (or emerge as) quality (linguistic syntactic ability). The basic difference between humans and animals then can be explained almost exclusively by the usage of symbolic/syntactic language. Of course, the explosive cultural evolution which became possible - once symbolic information processors like modern human brains arose - at first sight justifies the claim that at least one qualitative difference must distinguish humans from animals. It should be kept in mind that a minor additional trick sometimes can make a large difference. Moreover, one of us has previously briefly argued that the widely spread human need to claim human uniqueness can itself be explained from the need for continued self confirmation, which again follows from adding symbolic memes to the emotional - animal - being we are in the first place [92].

## **6.5. Musicality may also explain other typically human characteristics**

Finally, it should be emphasized again that song as a powerful means for pair bonding, as it appears to function in some animal species, can very easily explain another intriguing and far reaching characteristic of (modern?) humans. Human musicality can explain how the typically strong human pair bonding could have evolved. As such, song could explain not only speech, but also could help to understand the typically sexual and social behaviour of humans.

## **Postscript**

After this manuscript was accepted for publication, Lieberman [Lieberman, D.E. 1998. Sphenoid shortening and the evolution of modern human cranial shape. *Nature* 393: 158-162] argued to consider *Homo sapiens sapiens* (modern man) as a separate species from '*H. sapiens neanderthalensis*', because of clear facial differences with other hominids, incl. Neanderthals. Lieberman suggests that these changes may be related to the ability of speech. These considerations coincide with the claims - embraced in this article (see 4.1) - for a late origin of language, while the essential facial morphological characteristics of modern man may have been selected for by singing ability, enabling speech, but not for speech.

## **Acknowledgements**

MV is indebted to the FWO Flanders for an appointment as a research director.

## **Notes**

### **Note 1**

In essence, the original memes (as used among animals) can be defined as behaviours which can be mimicked. Dawkins [20] referred to bird songs as memes. However, one reviewer remarked that only humans can imitate in an observable manner. If only this kind of conscious imitation counts for memes, than only humans produce memes and washing sweet potatoes by Japanese macaques (see 3.2.4) would not be caused by imitation and thus not be memetic. One may object that there is strong evidence for unconscious imitation underlying learning in animals, as becomes apparent from the work of Rizzolatti et al. [71]:

*"In area F5 of monkey premotor cortex there are neurons that discharge both when the monkey performs an action and when he observes similar actions made by another monkey or by the experimenter. We report here some of the properties of these 'mirror' neurons and we propose that their activity 'represents' the observed action. We posit, then, that this motor representation is at the basis of the understanding of motor events."*

Finally, it should be noted that conscious imitation itself might be a secondary consequence of the development of language, which makes possible reflexive awareness. If one could show that conscious imitation is a consequence of reflexive awareness (i.e., consciousness), this kind of imitation could be considered itself largely as explained once one has explained language.

## **Note 2**

It is essential here to reflect on the definitions of selection and natural selection. Selection is a general principle: whenever there is variation on a theme, selection by the environment will occur, since none, one, more or all variations (configurations) may fit for existence in this environment. Natural selection is a special case which follows from the fact that selection takes place among variants on the theme of self-replicating systems, i.e. cells. The survival of the information processor (the cellular enzymatic machinery) is intrinsically linked to the information itself and vice versa. While differential survival of the information processors (the cells and the multicellular colonies) determines the reproductive success of the information molecules (the genes), the (genetic) information in turn determines the survival rate and reproductive success of the information replicators.

We could speak of a closed semantic circle (present in a metabolically open system).

However, in cultural-memetic selection, the information processors (animals, humans, copy machines, presses, computers) can die or stop functioning while the instantiations of information (memes, habits, knowledge) continue to flourish, and vice versa some instantiations of information can be lost or gained - for different reasons - without influencing the survival and/or activity of the information processors. As such, selection of behavioural/memetic/cultural information is basically different from the 'special case' of natural selection, although the general principles of evolution (change over time) and selection can be applied.

## **Note 3**

Consider the following experiments:

A chimpanzee named Panzee first saw a keeper hide food in one of two locked boxes. When a second keeper entered, Panzee learned to point the second keeper in which of two cages the food was hidden in order to obtain the food. The next experiment however seems definite proof of the fact that the chimp knows which knowledge is in the mind of the attendants and which knowledge it should add to get the food: keeper 1 hides the food, locks the box and gives the key to keeper 2, while leaving. After keeper 1 left, keeper 2 hides the key and leaves. Keeper 1 then returns without knowing where the key is hidden. If the chimp had learned by trial and error alone, she would still point to the box where the food was hidden. Instead, on her first try, she pointed to where the key was hidden. The chimp showed she could fathom the working of another mind: she knew that keeper 1 did not know where the key was.  
(After Mills [57])

## **Note 4**

This leads to the remark that 'The Language Instinct' as the title of a book claiming a Darwinian approach to the problem of the nature and the origin of language, would have been disapproved by Darwin [19] himself.

### **Note 5**

There might be some other resemblance between the song capacities of song birds and humans, although this is not really essential to the hypothesis put forward here. The front limbs in birds have been specially adapted for repetitive motor behaviour, flight, and Calvin [15] has proposed that special motoric capacities in humans, through e.g. natural selection for better throwing capacities, led to increased brain capacities in humans. Analogously, song birds are among the most intelligent birds. However, Calvin [15] and/or others seem to claim that these motor capacities by themselves are sufficient explanation for the linguistic capacities of humans, while it is argued here that these were only preadaptations which enabled singing, which itself then forms the essential preadaptation to speech. Thus, one could propose that for birds the flying capacity was a useful preadaptation for the possibility of song capacity, like for humans, specific motoric capacities - needed for e.g. throwing - prepared for the possibility of singing.

### **Note 6**

We focus on song here, because the aim of this paper is linking it to speech. However it is clear that song and dance go together. Many societies are known not to distinguish song from dance [78]. In most circumstances where singing and dancing have not been professionalised and so are done by all members of a group, when people sing they dance (or make other collective bodily movements), and when they dance they sing. Dance does not require vocal control but it can be suggested that the processes which modulate vocalisation are not restricted purely to the vocal tract but extend to incorporate other aspects of the body. Indeed, research indicates a close linkage between speech and gestures [29]. We suggest that part of the evolution of vocal modulation included the ability to incorporate with vocalisation other patterns of movement.

### **Note 7**

With respect to the 'environment', it should be noted in passing that children learn more readily from other children than from their parents and that they are more profoundly influenced by the habits (including language) of other children than by the habits of their parents (personal observations). A possible reason may be that they need to adopt the behaviours and habits of their play mates to get accepted in this social group.

## **References**

- [1] Aitchison, J. 1997. *The Seeds of Speech, Language Origin and Evolution*. Cambridge University Press.
- [2] Albert, M.L., R.W. Sparks, and N.A. Helm. 1973. Melodic intonation therapy for aphasia. *Archives of Neurology* 29: 130-131.
- [3] Allott, R. 1997. Pinker's language instinct: gradualistic natural selection is not a good enough explanation. <http://www.percep.demon.co.uk/pinker.htm>

- [4] Aslin, R.N., D.B. Pisoni, B.L. Hennessy, and A.J. Perey. 1981. Discrimination of voice onset time by human infants: new findings and implications for the effects of early experience. *Child Development* 52: 1135-1145.
- [5] Baker, R.R. 1996. *Human Sperm Competition: Copulation, Masturbation and Infertility*. Chapman & Hall, London.
- [6] Bates, E., and J. Goodman. 1997. On the inseparability of grammar and the lexicon. *Language and Cognitive Processes* 12: 507-586. (The paragraph quoted about chimpanzee language is pp. 544-545.)
- [7] Belin, P. P. Van Eeckhout, M. Zilbovicius, P. Remy, C. Francois, S. Guillaume, F. Chain, G. Rancurel, and Y. Samson. 1996. Recovery from nonfluent aphasia after melodic intonation therapy: a PET study. *Neurology* 47: 1504-1511.
- [8] Bertelson, P., B. de Gelder, L.B. Tfouni, and J. Morais. 1989. Metaphonological abilities of adult illiterates: New evidence of heterogeneity. *European Journal of Cognitive Psychology* 1: 239-250.
- [9] Bickerton, D. 1990. *Language and Species*. University of Chicago Press, Chicago.
- [10] Blacking, J. 1974. *How Musical is Man?* University of Washington Press, Seattle.
- [11] Botha, R.P. 1989. *Challenging Chomsky*. Blackwell, Oxford.
- [12] Brown, J., C. Sherrill, and B. Gench. 1981. Music may have importance to the development of general motoric skills rather than to language alone: effects of an integrated physical education/music program in changing early childhood perceptual-motor performance. *Perception and Motor Skills* 53: 151-154.
- [13] Burling, R. 1993. Primate calls, human language, and nonverbal communication. *Current Anthropology* 34: 25-53.
- [14] Castro-Caldas, A., K.M. Petersson, A. Reis, S. Stone-Elander, and M. Ingvar. In press. The illiterate brain: learning in childhood determines the functional organization of the adult brain.
- [15] Calvin, H. 1983. A stone's throw and its launch window: timing, precision, and its implications for language and hominid brains. *Journal of Theoretical Biology* 104: 121-135.
- [16] Chomsky. N. 1957. *Syntactic Structures*. Mouton, The Hague.
- [17] Chomsky, N. 1988. *Language and the Problems of Knowledge*. MIT, Cambridge.
- [18] Cooper, R.P., and R.N. Aslin. 1990. Preference for infant-directed speech in the first month after birth. *Child Development* 61: 1584-1595.

- [19] Darwin, C. 1871. *The Descent of Man*. Murray, London.
- [20] Dawkins, R. 1976. *The Selfish Gene*. Oxford University Press, Oxford.
- [21] Deacon, T.W. 1997. *The Symbolic Species: The Co-evolution of Language and the Brain*. W.W. Norton & Cy, New York.
- [22] De Boysson-Bardies B., P. Halle, L. Sagart, and C.J. Durand. 1989. A crosslinguistic investigation of vowel formants in babbling. *Child Language* 16: 1-17.
- [23] DeCasper, A.J., and W.P. Fifer. 1980. Of human bonding: newborns prefer their mother's voices. *Science* 208: 1174-1176.
- [24] DeCasper, A.J., and M.J. Spence. 1986. Prenatal maternal speech influences newborns' perception of speech sounds. *Infant Behavior and Development* 9: 133-150.
- [25] Diamond, J.M. 1997. *Guns, Germs, and Steel: The Fates of Human Societies*. Norton, New York.
- [26] Diamond, J.M. 1997. The language steamrollers. *Nature* 389: 544-546.
- [27] Eimas, P.D., J.L. Miller, and P.W. Jusczyk. 1987. On infant speech perception and the acquisition of language. Pp. 161-195, In *Categorical Perception* (S. Harnad, Ed.). Cambridge University Press, New York.
- [28] Fernald, A., and T. Simon. 1984. Expanded intonation contours in mothers' speech to newborns. *Developmental Psychology* 20: 104-113.
- [29] Feyereisen, P. and J.-D. de Lannoy. 1991. *Gestures and Speech: Psychological Investigations*. Cambridge University Press, Cambridge.
- [30] Fisher, S.E., F. Vargha-Khadem, K.E. Watkins, A.P. Monaco, and M.E. Pembrey. 1998. Localisation of a gene implicated in a severe speech and language disorder. *Nature Genetics* 18: 168-170.
- [31] Geschwind, N. 1980. Some comments on the neurology of language. In: *Biological Studies of Mental Processes*, ed., D. Caplan. MIT Press.
- [32] Gopnik, M. 1990. Feature-blind grammar and dysphasia. *Nature* 344: 715.
- [33] Gopnik, M., and M.B. Crago. 1991. Familial aggregation of a developmental language disorder. *Cognition* 39: 1-50.
- [34] Gould, J., and R.C. Lewontin. 1979. The spandrels of San Marco and the Panglossian paradigm: a critique of the adaptationist program. *Proceedings of the Royal Society of London B*205: 281-288.

- [35] Grant, P.R., and B.R. Grant. 1997. Genetics and the origin of bird species. *Proceedings of the National Academy of Sciences USA* 94: 7768-7775.
- [36] Greenfield, P.M. 1991. Language, tools and brain: the ontogeny and phylogeny of hierarchically organized sequential behaviour. *Behavioral Brain Science* 14: 531-595.
- [37] Griffin, D.R. 1992. *Animal Minds*. University Chicago Press, Chicago.
- [38] Haimoff, E.H. 1981. Video analysis of siamang (*Hylobates syndactylus*) songs. *Behaviour* 76: 128-151.
- [39] Harris, R.A. 1993. *The Linguistic Wars*. Oxford University Press, Oxford.
- [40] Herman, L.M., D.G. Richards, and J.P. Wolze. 1984. Comprehension of sentences by bottlenosed dolphins. *Cognition* 16: 129-219.
- [41] Jespersen, O. 1922. *Language, its Nature, Development and Origin*. Allen & Unwin, London.
- [42] Jonker A. 1987. The origin of the human mind. A speculation on the emergence of language and human consciousness. *Acta Biotheoretica* 36: 129-77.
- [43] Jochims, S. 1994. Establishing contact in the early stage of severe craniocerebral trauma: sound as the bridge to mute patients. *Rehabilitation (Stuttg)* 33: 8-13.
- [44] Kalmus, H., and D.B. Fry. 1980. On tune deafness (dysmelodia): frequency, development, genetics and musical background. *Annals Human Genetics* 43: 369-382.
- [45] Kluender, K.R., R.L. Diehl, and P.R. Killeen. 1987. Japanese quail can learn phonetic categories. *Science* 237: 1195-1197.
- [46] Kuhl, P., and J. Miller. 1975. Speech perception by the chinchilla: Voice-voiceless distinction in alveolar plosive consonants. *Science* 190: 69-72.
- [47] Lee, R.B. 1979. *The !Kung San: Men, Women, and Work in a Foraging Society*. Cambridge University Press, Cambridge.
- [48] Lehrdahl, F., and R. Jackendoff. 1983. *A Generative Theory of Tonal Music*. MIT Press, Cambridge, MA.
- [49] Lieberman, P. 1984. *The Biology and Evolution of Language*. Harvard University Press, Cambridge, Ma.
- [50] Lieberman, P. 1991. *Uniquely Human: the Evolution of Speech, Thought and Selfless Behaviour*. Harvard University Press, Cambridge, Ma.

- [51] Lynch, M.P., L.B. Short, and R. Chua. 1995. Contributions of experience to the development of musical processing in infancy. *Developmental Psychobiology* 28: 377-398.
- [52] Maddieson, I. 1981. UPSID UCLA phonological segment inventory database. *UCLA Working Papers in Phonetics* 53: 1-243.
- [53] Matsuzawa, T. 1991. Nesting cups and metatools in chimpanzees. *Behavioral Brain Science* 14: 570-571.
- [54] McNeill, W.H. 1995. *Keeping together in time: dance and drill in human history*. Harvard Univ. Press.
- [55] Mehler, J., P. Jusczyk, G. Lambertz, N. Halsted, J. Bertoncini, and C. Amiel-Tison. 1988. A precursor of language acquisition in young infants. *Cognition* 29: 143-178.
- [56] Miller, S.B., and J.M. Toca. 1979. Adapted melodic intonation therapy: a case study of an experimental language program for an autistic child. *Journal Clinical Psychiatry* 40: 201-203.
- [57] Mills, C. 1997. Unusual suspects. *The Sciences*, July/August: 32-36.
- [58] Milo, R.G, and D. Quiatt. 1993. Glottogenesis and anatomically modern *Homo sapiens*: The evidence for and implications of a late origin of vocal language. *Current Anthropology* 34: 569-598.
- [59] Morgan, J.L. 1996. Prosody and the roots of parsing. *Language and Cognitive Processes* 11: 69-106.
- [60] Morais, J. 1987. Phonetic awareness and reading acquisition. *Psychological Research* 49: 147-152.
- [61] Morais, J., L. Cary, J. Alegria, and P. Bertelson. 1979. Does awareness of speech as a sequence of phones arise spontaneously? *Cognition* 7: 323-331.
- [62] Nelson, D.G., K. Hirsh-Pasek, P.W. Jusczyk, and K.W. Cassidy. 1989. How the prosodic cues in motherese might assist language learning. *Journal of Child Language* 16: 55-68.
- [63] Noble, W., and I. Davidson. 1996. *Human Evolution, Language and Mind: A Psychological and Archaeological Inquiry*. Cambridge University Press, Cambridge.
- [64] Ong, W.J. 1982. *Orality and Literacy: The Technologising of the Word*. Methuen, London.
- [65] Peretz, I., M. Babai, I. Lussier, S. Hébert, and L. Gagnon. 1995. Musical excerpts: indices relating to familiarity, age of acquisition and verbal associations. *Canadian Journal Experimental Psychology* 49: 211-239.

- [66] Pinker, S., and P. Bloom. 1990. Natural language and natural selection. *Behavioral and Brain Sciences* 13: 707-784.
- [67] Pinker, S. 1994. *The Language Instinct: How the Mind Creates Language*. Morrow, New York.
- [68] Provine, R.R. 1996. Laughter. *American Scientist* 84: 38-45.
- [69] Reis, A., and A. Castro-Caldas. 1997. Illiteracy: A cause for biased cognitive development. *Journal of the International Neuropsychological Society* 3: 444-450.
- [70] Ridley, M. 1993. The units of selection, Chapter 12. Pp. 303-322, In *Evolution*. Blackwell, Boston.
- [71] Rizzolatti, G., L. Fadiga, V. Gallese, & L. Fogassi. 1996. Premotor cortex and the recognition of motor actions. *Brain Res. Cogn. Brain Res.* 3: 131-141.
- [72] Rodseth, L., R.W. Wrangham, A.M. Harrigan, and B.B. Smuts. 1991. The human community as a primate society. *Current Anthropology* 32: 221-254.
- [73] Rousseau, J.J. 1852/1966. On the Origin of Languages. In *On the Origin of Language (translated essays of Rousseau)* (J.H. Moran, and A. Gode, eds). University of Chicago Press, Chicago.
- [74] Sansavini, A., J. Bertoncini, and G. Giovanelli. 1997. Newborns discriminate the rhythm of multisyllabic stressed words. *Developmental Psychology* 33: 3-11.
- [75] Sataloff, R.T. 1992. The human voice. *Scientific American* 267: 64-71.
- [76] Savage-Rumbaugh, E.S., and R. Lewin. 1994. *Kanzi: the Ape at the Brink of the Human Mind*. John Wiley, New York.
- [77] Seckel, H.P. 1960. *Bird-headed Dwarfs*. Karger, Basel.
- [78] Seeger, A. 1994. Music and dance. Pp. 686-705, In T. Ingold, (Ed.), *Companion Encyclopedia of Anthropology*. Routledge, London.
- [79] Seki, K., and M. Sugishita. 1983. Japanese-applied melodic intonation therapy for Broca aphasia. *No To Shinkei* 35: 1031-1037.
- [80] Simmons, J.Q., and C. Baltaxe. 1975. Language patterns of adolescent autistics. *Journal Autism & Child Schizophrenia* 5: 333-351.
- [81] Skoyles, J.R. In press. Human evolution expanded brains to increase expertise capacity, not IQ. <http://www.users.globalnet.co.uk/~skoyles/brain.htm>.

- [82] Skoyles, J.R., and M. Vaneechoutte. In preparation.
- [83] Smith, J.M., and E. Szathmary. 1995. *The Major Transitions in Evolution*. Freeman, Oxford.
- [84] Steels, L. 1996. Interacting robot agents. Synthesising the origins of language and meaning using co-evolution, self-organisation and level formation. <http://www.heise.de/bin/tp-issue/tp.html?artikelnr=6001&mode=html>
- [85] Auditory radiations and cortex of the awake monkey. *Brain Research* 252: 353-365.
- [86] Streeter, L.A. 1976. Language perception of 2-month-old infants shows effects of both innate mechanisms and experience. *Nature* 259: 39-41.
- [87] Symons, D. 1979. *The Evolution of Human Sexuality*. Oxford University Press, Oxford.
- [88] Thorpe, W.H., and M.E. North. 1966. Vocal imitation in the tropical bou-bou shrike *Lanarius aethiopicus major* as a means of establishing and maintaining social bonds. *Ibis* 108: 432-435.
- [89] Tomasello, M. 1995. Language is not an instinct. *Cognitive Development* 10: 131-156.
- [90] Trehub, S.E. 1976. The discrimination of foreign speech contrasts by infants and adults. *Child Development* 47: 466-472.
- [91] Ujhelyi, M.J. 1996 Is there any intermediate stage between animal communication and language? *Theoretical Biology* 180: 71-76.
- [92] Vaneechoutte, M. 1993. The memetic basis of religion. *Nature* 365: 290. At <http://www.sepa.tudelft.nl/webstaf/hanss/nature.htm>
- [93] Vaneechoutte, M. 1997. Bird song as a possible cultural mechanism for speciation. *Journal of Memetics-Evolutionary Models of Information Transmission* 1. At [http://jom-emit.cfpm.org/1997/vol1/vaneechoutte\\_m.html](http://jom-emit.cfpm.org/1997/vol1/vaneechoutte_m.html)
- [94] von Humboldt, W. 1836/1988. *On Language: the Diversity of Human Language Structure and its Influence on the Mental Development of Mankind*. Cambridge University Press, Cambridge.
- [95] Zatorre, R.J., E. Meyer, A. Gjedde, and A.C. Evans. 1996. PET studies of phonetic processing of speech: Review, replication, and reanalysis. *Cerebral Cortex* 6: 21-30.

# THE REPLICATOR: A MISNOMER CONCEPTUAL IMPLICATIONS FOR GENETICS AND MEMETICS

By

**Mario Vaneechoutte**

Department of Clinical Chemistry, Microbiology & Immunology  
Blok A, University Hospital, B 9000 Ghent, Belgium  
E-mail: Mario.Vaneechoutte@rug.ac.be

Most evolutionary biologists and students of evolution in general have adopted the term 'replicator' for both genes and memes. There is indeed almost unanimous agreement [12]. However, current genes are not (self-) replicators and, as I will argue below, in contradiction to the current paradigm of the RNA-replicator world hypothesis, they have never been. Thus far, I came across only one source which defends a more correct interpretation of what genes are: "More pernicious, because less obvious, was the choice of the term 'replicator' for genes and other things of which copies are made. In English, the suffix -or is used to indicate that the root to which it is attached falls under the ontological category of an agent. As with the -er in 'copier', it means that which does the copying, not the sheet of paper that we copy when we feed it into a piece of office machinery. For the (passive) patient, in contradistinction of the (active) agent, we use other suffixes. In ordinary metaphysics we recognize that text gets created and modified by its author and editor, and transmitted by its copyist, perhaps using a copier. Had Dawkins [5] called his "replicators" "replicanda" his metaphysics would have been somewhat more transparent, but people might not have taken him so seriously". Quote from Ghiselin [8], in a criticism of the term replicator as it was coined by Richard Dawkins [5]. See also [9].

Indeed the suffix -or refers to an active agent, a processor. However, current DNA-genes are nothing but a material instantiation of information on how to make RNA and proteins, and perform no replication. Genes are replicated informational molecules, unable to replicate on their own, no more than enzymes can. Moreover, replication relies on the interaction of a whole collection of very different genes. Accordingly, 'memes' or bits of cultural information are processed, replicated and transmitted by human minds, photocopiers, presses, computer networks, etc. The original prebiotic genes are then better compared to plasmid genes which can be exchanged freely between different bacterial cells (processors) - OR with current memes (like printed scientific theories). They then can be regarded as elements of a language which enabled free recombination of different concepts, just like current scientific reasoning is possible because of the availability of copy true bits of information (printed matter), which can be recombined over and over again.

Partially due to the application of the term replicator for genes and memes, it has been overlooked that only the system cell as a whole enables self replication. It is important to emphasize that such a self replicating system thus far originated only once during the existence of Earth. Although it are the genes which carry heritable information, it is clear that daughter cells also need to inherit a minimal amount of processors, i.e. at least those enzymes which can transcribe and translate the DNA. So, although it are genes which are responsible for heritable information, it takes more than genes to be inherited to enable replication of genes, to make the processing system which can do so. Moreover, not only genes and some enzymatic processors,

but also membranes have to be inherited, since membranes are not synthesized de novo, but are synthesized only in close connection with the existing membrane. (This points to a very central and original role of membranes or to a 'membranes early' hypothesis of the origin of life.)

As a matter of fact, one could even consider new nucleotide strands as being a phenotypical artefact, since they are the product of enzymatic activity just as well as other cell components. Notice how this illustrates that the genotype-phenotype distinction itself has only limited usefulness, and - certainly at the cellular level - it occurs to be a somewhat artificial human-made dualistic distinction [see also 7], which in itself obstructs useful insights. One might object that for DNA-duplicates to be considered as phenotype, there is a difference with other phenotypic traits, namely that new DNA can be made only by using existing DNA as the mould. However, this observation accounts as well for membranes (see above). The only difference is the 'unlimited' [11] informational content of nucleotide molecules, but this characteristic is unlinked to replication.

I propose to adopt a terminology which is also used in information theory (and related fields like cybernetics)[see also 7] and to drop words like genotype, phenotype, replicator, interactor and vehicle, which have the disadvantage of poorly reflecting the dynamics and interactions which occur between biological processors and between biology and culture and which are in many cases strongly misguiding. It is more revealing to speak of digital information processors (e.g. polymerase enzymes) and encoded, digital information (genes). In case of self-replicators, that is cells, both processors and digital information are embedded in a housing which consists of structural molecules (proteins, peptidoglycan, phospholipid bilayers, polysaccharides ...).

From these remarks, more useful definitions then can be put forward: living cells are the only self replicators on Earth. Enzymatic polymerases, resp. human brains, copiers, presses, etc. are all processors which can function as replicators of informational molecules, resp. informational symbols. These processors are information processors. Finally, informational molecules (polynucleotides) and texts, ideas, words, behaviours are replicable elements of instantiated information. For this category, we might adopt the term 'replicanda' as it has been proposed by Ghiselin [8], although the exact translation of this Latin term is 'that what has to be replicated'.

This definition then only holds for extant chromosomal genes, since indeed the cell is compelled to copy its genes in order to produce cellular offspring. However, it should be noticed that, certainly in the case of cultural information, the successful reproduction of information processors (like humans) does not depend upon the replication of cultural information, and therefore that there is no functional need to replicate this information.

'Replicata' or 'replicates' ('that what is/can be replicated') might be a more general term for informational molecules (nucleotidic genes) or cultural bits of information (memes) which are replicated.

In summary, the use of replicator for gene or meme is twice erroneous: they are not self-replicators (remark that 'replicator' is generally used to have the implicit meaning of self-replicator) and they even are not processor-replicators. Genes and memes are replicates or bits of replicable information. For memetics this means that we have to ask the question why some information is replicated more successfully by replicators like human minds and not why they

replicate more successfully than other memes. This more scientifically correct approach is e.g. used in the paper of Heylighen [10].

## Selection, natural selection and cultural selection

These definitions of self-replicator, replicator and replicate lead us to considerations on the essence of natural selection as a special case of selection. Selection is rather a tautological principle: whenever there is variation on a theme, one of these themes will be best fitted to certain conditions, given by the environment, which also includes the other variations on the theme. Having different kinds of cars, one car will be best suited to reach high speeds on a straight road, one car will be best suited for riding a curly parcours, one car will be best suited for transporting many people, one car will be best suited for driving and parking easily in a crowded city, one car will confer an image of certainty and successfulness to its owner. In case most people live in cities and do not care too much about confirming their self image by means of the car they own, a type of small car that enables easy driving in the city will be sold most and will therefore be most successful when measured by number of copies made of it.

Specific cases of selection then can be understood by defining the theme, it's possible variations on the one hand and on the other hand the environment to which the different possible variations on the theme will have to fit.

All of these general considerations on what selection is hold as well for natural selection. However, natural selection is often erroneously considered as the only possible form of selection and it is forgotten that it is a very special case of selection, because the selection occurs among self-replicators (that is, cells), unlike cars, prebiotic metabolic cycles and ideas. Moreover, all the self-replicators stem from the first system able to duplicate. Since this duplication is not watertight copy proof, several variations on the original theme (the first cell or the theme of self-replication) are continuously produced, with as a 'welcome' side-effect that in most cases some of the variations will fit to whatever environment or whatever environmental change.

In summary, because a self-replicating system once occurred on Earth, it was possible that the original genetic code of it - which contained all the necessary information on how to make the next self-replicating system, was copied manifold. High copy fidelity, ensuring sufficient numbers of functional self-replicators, allowed also for a number of mutants to exist. Some mutants also could survive. Some mutants even could explore new environments or did better in the old environment than their parent wild types. A small percentage of erroneous descendants is an ensurance for adapting to previously not inhabited environments or for adapting to the changing environment.

These considerations stem from an information centered approach which leads to more generally applicable insights than a gene centered one.

Natural selection is about the replication of informational molecules whereby the replication of the information depends on the functionality of the organisms, i.e. the genetic information processors for which the information encodes. The fitness of the phenotype is also the fitness of

the genes, and vice versa. Cells are metabolically open, but semantically closed systems. From this definition the difference with cultural information becomes obvious.

Just like variable genetic information and variable information in general, variants of cultural information are selected as well. I referred to this as a tautology.

For example, different scientific theories are weighed by the scientific community, where after finally one or still a third, or a recombination of still another five ideas appears to solve best the problems tackled and is replicated more successfully than other theories. The difference with natural selection is that the defenders of one theory - the scientists, the information processors - do not die because their theory dies or that they do not produce more offspring because they defended a successful theory. The information encoded in a meme (usually) has no influence on the survival of the processor (a human in this case). Vice versa, the ideas of a parent of many children are not necessarily - certainly no longer at present - successfully replicated ideas. Scientists with many children are not necessarily successful in spreading their theories. Christ had no children. The evolution of cultural information does not depend on the differential survival of its processors to proceed.

By reversing the adagium that memes must be replicators in analogy to the fact that genes are replicators into the statement that genes are replicates just like memes are, we can further suggest that genes were also replicates in prebiotic life. This suggestion contradicts the current paradigm that life evolved by competition between self-replicating RNA-molecules, whereby the cell gradually evolved around the fastest self-replicating RNA-molecules by means of natural selection, i.e. as a consequence of the differential reproduction rate of different self-replicating molecules. We propose to postpone natural selection till after the origin of the first chromosome containing cell (the first self-replicator) and to consider how a much more flexible and powerful kind of memetic evolution took place, comparable to what is happening since the origin of symbolic language in the animal world.

(It should be noticed that this does not mean that we have to dismiss the several exciting findings of RNA-research [e.g. 3, 14] nor that we disagree with the 'nucleotides earlier than complex proteins' hypothesis. Only, we propose to consider these findings from a different point of view.)

Symbolisation of information thus far has occurred twice on Earth. Prebiotic life developed a manner to encode information (symbolisation) into informational molecules (nucleotide strands). In biology, informational communication first took a nonchemical way of transmission with the advent of neuronal brains in animals which could make use of nonmaterial sound wave energy and photonic energy (vision) to transmit information (the origin of memes), where after strong symbolisation of this nonmaterially transmitted information has occurred in humans by the development of spoken language. This has finally led to the re-entry of material substrates which could carry information as visual symbols, in the form of written language. I will argue that the initial role of nucleotides is comparable to that of written language. While nucleotides enabled to encode chemical interaction, symbolic language enabled to encode behavioural interaction.

The essential thing about encoded information is that it can be mixed, recombined, merged, over and over again, because of the fact that it is not a metabolic process but a code on how to

perform such a process, which has the virtue that the information is not transformed (only transcribed or translated) during interaction. It is impossible to mix processes: One can't mix the activity of enzyme x with that of enzyme y, one can't take the bacterium *E. coli* and mix it with a piece of human tissue, while keeping a functional process; one can't mix the ideas of two people by mixing their brains.

But what if these processes also are encoded somewhere? Evolution has managed many times to combine the functions expressed in different modules of enzymes with each other. This could be done because there was a genetic code for it and by recombining parts of this code, something completely new, composed of already existing components but arranged in a different manner, could emerge. Similarly we can mix the eubacterium *Escherichia coli* with human enzymes, by inserting into its genome the code for such a human enzyme (e.g. for the production of insulin for diabetes patients).

We can mix different ideas since they can exist as a memory in our mind. This process is largely enhanced by the availability of written records, since this provides a more faithful, copy true back up. One can mix as many of these ideas as one can gather (e.g. by reading). Many different lineages of information can continuously come together. Sometimes (but apparently rarely) this leads to new useful insights. Furthermore, it takes only one processor and one template to make an unlimited number of copies of encoded information, something which is not achievable when dealing with processes.

The following is just a brief exercise of how we could envisage analogies between 'memetic' evolution towards the origin of the genetically encoded cell and present day cultural evolution (which may or may not lead to a novel kind of 'memetic' self replicator, e.g. some kind of self-assembling, duplicating robot as suggested by Tipler [13]).

Not in contradiction with some current thinking [e.g. 6] we assume the existence of a premetabolic network of interacting chemical molecules, organized in smaller networks of heterocatalytic cycles, mostly embraced by membranes. Such a membrane with its internal metabolic cycling can be called a protocell (which is devoid of genetic material), which interacted intensively with other protocells and/or free-'living' cycles and available molecules. The prebiotic system as a whole can then be compared with the current ecological network, which is composed of interacting genetically encoded organisms of different complexities.

The basic proposition of this paper is that symbolisation of chemical interaction (communication) between these protocells, and/or parts of their internal networks, is suggested to have been the original function of nucleotidic information. Instead of direct metabolic interactions (comparable to direct behavioural interaction between animals), some cells also started to make use of informational molecules, which not directly took part in the interactions but which nevertheless could influence the behaviour of other protocells (or of intracellular cycling), just like symbolic sounds or written texts can influence the behaviour of other humans (or of 'intrapersonal' thinking). (Remark that information was defined by Bateson as a difference which makes a difference to 'someone' else [7].)

Written/printed information, just like nucleotide information, can then be used without being transformed (i.e. analogously processed) during this interaction: it is digital.

With respect to the evolution of information, the chemical cycling within protocells is best compared to the animal body, performing metabolically essential functions. The membrane is best compared with the neural system (incl. brains and perception organs) of animals, since it performs interactive, communicative functions. Like some of these protocells started interacting by means of symbolic molecules (informational molecules, stretches of nucleotides), some animals (humans) started interacting by means of symbolized gestures and sounds. This leads to new interactions, also with nonsymbol using protocells (accordingly with other animals... bacteria). E.g., this symbolic information may have enabled the symbol using protocells to use metabolic energy from simpler protocells, like humans started to use animal power. This is to be regarded as some kind of symbiosis, since both symbol using and nonsymbol using protocells can have metabolic profit, like the number of horses increased correspondingly with the number of humans.

(Remark that an inadvertent outcome of this kind of reasoning is that this story is about symbiosis and mutual profits, in contradiction to the 'natural selection early' approach whereby the selfish fastest replicators won the game.)

While the informational molecules may have been internalized and processed at first in order to create new ones (like humans process words in an analogous manner: transformation), the interaction between protocells and informational molecules may have lead to the production of external processors (metabolic and informational enzymes). Similarly, humans aided by words have produced different external metabolic processors (wheels, ploughs, tractors, factories) and information processors (pens, type writers, presses, copiers, and computers).

Because in this model the replication of informational molecules presupposes only the ongoing of metabolic activity in general and thus not depend on the direct 'survival' of some particular protocell - unlike in e.g. models of molecular self-replicators and chemotons [11] and unlike present day chromosomal genes, both cases where the fate of the informational molecules is tightly linked to that of the metabolic efficiency of its cell - a lot of recombinatorial freedom was available to prebiotic chemistry. Similarly, cultural information can be recombined in an at random manner, whether or not this leads to silly ideas (like this one!?), because it is not linked to the fate of the replicators (human brains, computers). In fact the interaction between protocellular membranes, nucleotide encoded information and extracellular processors may have enabled to produce large extracellular heterogeneous protochromosomes, just like the interaction between humans and printed code and external information processors like presses and computers resulted in the production of scientific theories and complex machines.

The whole community of (simple protocells and) symbol using protocells became more and more dependent on the usage of external informational molecules and external informational and metabolic processors. The same happens in our present society, which would collapse at once when language, printed information, production machines or computers were taken away. In the end, most metabolic functions and most information processing were done by digitally processing external enzymes (e.g. polymerases), so that the role of the original protocells was

merely a motivating one. Accordingly, humans are less and less needed to keep the information processing going on. We put in a search string, push some buttons and the world wide web searches the relevant information. We are more and more needed only as motivators. Current co-enzymes (motivators of enzymatic activity) accordingly can be regarded as remnants of the original premetabolic protocells.

Finally, protocells consisted of a membrane, probably flattened since at this stage it contained only few internal chemical cycles, with an externally attached DNA-protocromosome - consisting of genes encoding for very different functions - and some externally attached enzymes and RNA-molecules. At several occasions such membranes may have 'gastrulated', whereby the chromosome and the enzymes and possibly some external medium with free enzymes and other molecules were internalized (in accordance with the obcell theory [1,2]). At one such occasion the enclosed informational molecules and processors contained all the necessary information and functionality to produce a self-replicating system (remark that reverse transcriptase was not enclosed). A similar event has not yet happened in culture (self-assembling robots [13] are just one imaginable outcome).

While explaining the origin of life as a memetic process, this approach also might provide a unifying theory for the evolution of information, whereby at two events symbolisation of information occurred - once of chemical interaction, once of behavioural interaction - and whereby at one occasion this encoding lead to a chemical self-replicator (the biological cell), a memetic evolution which seems to be happening again.

This symbiotic approach may circumvent many of the current problems like the 'chicken or egg' problem of DNA/RNA/protein interaction, not (yet) solved by the RNA-replicators world hypothesis. It also more easily explains how a chromosome consisting of very different functions, and whereby none of these genes by themselves contain sufficient information to cover the complex process of self-replication, can have originated. It circumvents Eigen's paradox whereby different self replicators outcompete each other when brought together in a cell, instead of merging into a chromosome (Eigen's paradox is solved (?) only by the rather artificial stochastic corrector model [11]). The much faster and flexible memetic (cultural) evolution also might resolve the time window problem (whether real or not [see 6], which lead some eminent biologists to adopt an extraterrestrial origin of life [4]. Of course, this 'memetic origin of life' hypothesis poses its own problems.

## References

1. Blobel, G. 1980. Intracellular membrane topogenesis. Proceedings of the National Academy of Sciences, USA 77: 1496.
2. Cavalier-Smith, T. 1987. The origin of cells: a symbiosis between genes, catalysts, and membranes. Cold Spring Harbour Symp. Quant. Biol. 52: 805-824. BioMedNet Medline Search Abstract.
3. Cech, T.R. 1986. A model for the RNA-catalyzed replication of RNA. PNAS, USA 83: 4360-4363.

4. Crick, F.H.C., & L.E. Orgel. 1973. Directed panspermia. *Icarus* 19: 341-346.
5. Dawkins R., 1976, 'The Selfish Gene', Oxford Univ. Press.
6. de Duve, C. 1995. Vital dust. Life as a cosmic imperative. Basic Books. New York.
7. Emmeche, C., & J. Hoffmeyer. 1991. From language to nature. The semiotic metaphor in biology. *Semiotica* 84: 1-42. <http://www.geneticengineering.org/dna5/default.htm>
8. Ghiselin, M.T. 1994. Chapter 9, pp. 143-148. In Metaphysics and the origin of species. SUNY Press, Albany, NY.
9. Ghiselin, M.T. 1987. Replicators and replicanda. Bioeconomics and the metaphysics of selection. *J. Soc. Biol. Structures* 10: 361-369.
10. Heylighen, F. 1998. What makes a meme successful? Selection criteria for cultural evolution. First Symposium on Memetics, Namur, Belgium.
11. Smith, J.M., & E. Szathmary. 1995. The major transitions in evolution. W.H. Freeman & Co. Ltd. Oxford, UK.
12. Speel, H.-C. 1998. Memes are also interactors. First Symposium on Memetics, Namur, Belgium.
13. Tipler, F. 1995. The physics of immortality: modern cosmology, God and the resurrection of the dead. Anchor.
14. Zhang, B., and T.R. Cech. 1997. Peptide bond formation by in vitro selected ribozymes. *Nature* 390: 96-99.

# **MERGERS AND TAKEOVERS; A MEMETIC APPROACH**

**By**

**Ed Vos and Ben Kelleher**

University of Waikato, Department of Finance

Private Bag 3105, Hamilton, New Zealand.

Tel: +64 7 838 4466 #8110

Fax: +64 7 838 4145

[evos@waikato.ac.nz](mailto:evos@waikato.ac.nz)

## **Abstract**

This paper constructively diagnoses problems within the current merger & acquisition (M&A) theories and provides an alternative theory of corporate behaviour. We contend that humans are the hosts for a replicating entity known as 'memes'. Since finance based motivational studies on M&A activities have not established that this activity 'adds value' to the acquiring firm, it is our thesis that certain managers gain power through mergers and acquisitions. Thus, M&A from the point of view of the acquiring firm can be seen as driving the evolution of ideas, shaping the flow of technology, information, and tastes rather than as 'value adding'. In simple terms, managers (the meme holders) use mergers and acquisitions to enhance their power, and in gaining this power managers unconsciously provide an improved medium through which their memetic 'stories' may be replicated.

The paper first introduces popular theories surrounding M&As, the motivation for M&A is developed further in an examination of the need for managerial power, we then discuss how this power struggle relates to the 'battlefield' of corporate asset allocation allowing 'stories' to be told and replicated, and finally we ask questions and develop hypotheses that provide direction for future research.

## **1 Introduction**

This paper co-integrates theories from several academic disciplines to explain the motives behind mergers and acquisitions (M&As). Financial theory and literature noticeably dominate the view of M&As presented in this article. This paper is not meant as a complete descriptive of all individual merger and acquisition motives, rather we attempt to explain certain motive phenomena that can be observed in an aggregate state. In a theoretical sense, when considering mergers and acquisitions finance theory suggests employment of a capital budgeting model. Correspondingly, the driver of all mergers and acquisitions should be to increase shareholder value, as the general consensus among those in the field of finance, is that the principal goal of a firm should be the maximisation of stockholder's wealth (Mandelker, 1974). However, we show that practitioners may not always follow such a theory and that there are often other important factors influencing corporate restructuring that are outside the realm of merely enhancing shareholder value.

Mergers and takeovers have attracted a great deal of attention from the financial press and other media in recent times (The terms merger and acquisition "*are used interchangeably to mean any transition that forms one economic unit from two or more previous ones*" (Lubatkin & Shrieves, 1986, p. 497)). Their characteristics and resulting values pre and post formation have been studied thoroughly in the financial literature (see Jensen and Ruback, 1983, for a review).

Theoretically, a company will enter into an acquisition or merger agreement if they believe that the NPV (company A + company B) > NPV (company A) + NPV (company B), where NPV is *Net Present Value*. Or simplistically, the economic value of these firms combined is greater than the economic value of these two firms as separate entities.

Financial theory implies that acquisitions and mergers occur in the hope of positive synergistical effects, with many managers citing synergy arguments in order to justify their actions (Friedman and Gibson, 1988; Maremont and Mitchell, 1988; Porter, 1987). Reasons for these effects have been offered, such as gaining fast access to new technologies or new markets, benefiting from economies of scale in research and/or production, tapping into sources of know how located outside the boundaries of the firm and finally monopoly type advantages. Lubatkin lists M&A motivation into seven main theoretical areas:

### **Monopoly Theory: Gaining market power**

- Efficiency Theory: Operating synergies, financial synergies and management synergies.
- Valuation Theory: Bidder managers have better information about the target's financial performance than the stock market.
- Empire Building Theory: Planned and executed by managers who maximise their own utility instead of their shareholders value.
- Process Theory: Managers have only limited information and base decisions on imperfect information.
- Raider Theory: Managers creating wealth transfers from the stockholders of the companies they bid for.
- Disturbance Theory: Merger waves are caused by economic disturbances.

Lubatkin further adds that the most popular efficiency based theories are not substantiated empirically (Lubatkin, 1983), raising the question: Do mergers provide real benefits to acquiring firms?

The motivations for merger and acquisition activity seem logical and worthy of study as has been the case in the past (see Berkovitch and Narayanan, 1993, Markides & Oyon, 1998 & Trautwein, 1990). However, mergers may not lead to positive performance outcomes. The evidence to date suggests that most acquisitions cost more than they are worth and most mergers fail. Ravenscraft and Scherer (1987) and Herman and Lowenstein (1988), examine the earnings performance after takeovers and conclude that merged firms have no operating improvements. Porter (1987) found that more than half of the acquisitions by major US companies failed. Whilst Hunt states that "*most studies of acquisitions produce a success to failure rate (using accounting or finance or managerial assessments) of 50%*" (Hunt, 1990, p.70). Finally, Lorenz (1986) observed that takeovers are "*at best an each way bet*".

In Jensen and Ruback's 1983 review paper, they suggest that studies of the abnormal returns to takeover participants show that in general bidding firms seem to have no significant positive returns. Additionally, Loughran and Vijh (1997), Travlos (1987), Warnesly, Lane and Yang (1987), Franks, Harris, and Mayer (1988), Asquith, Bruner, and Mullins (1987), and Servaes (1991), find that acquirers paying with stock earn significantly negative abnormal returns. It is no surprise that this lacklustre performance in merger and acquisition returns has been summarised as unquestionably poor (Doz, 1988). In a 1970 paper that surveyed the findings of the finance research on mergers, Hogarty asked:

*"What can fifty years of research tell us about the profitability of mergers? Undoubtedly the most significant results of this research have been that no one who has undertaken a major empirical study of mergers has concluded that mergers are profitable; i.e. in the sense that they are more profitable than alternative forms of investment." (Hogarty, 1970, p. 220).*

McKinsey and Company revealed that 43% of a sample of international acquisitions failed to produce a financial return that met or exceeded the acquirer's cost of capital (Bleeke & Ernst, 1993). Non-financial studies show little improvement over Kitching's (1974) early finding that between 45% and 50% of acquisitions are considered failures or not worth repeating by the managements involved. Further support comes from Michael Porter's (1987) examination of the diversification record of large US firms over the period 1950-1986. He found that 53% of all acquisitions were subsequently divested, rising to 74% for unrelated acquisitions.

Therefore, if the discussed synergistical effects are present and financial capital budgeting techniques are utilised, why is it that bidder shareholders often never experience wealth increases? Correspondingly, in light of the historical evidence why do firms continue to undertake such seemingly inefficient capital investments?

Our aim (albeit a struggle) is to escape out of the constraining 'straight jacket' view that shareholder wealth maximisation places on the study of finance, to examine an alternative and perhaps radical foundation for the reasons behind many corporate mergers and acquisitions. It is proposed that mergers and acquisitions are both a medium through which managerial power is increased and a vehicle through which the 'corporate story' that we as consumers unconsciously absorb is skilfully replicated. The 'corporate story' represents an idea or combination of ideas that are transferred via hosts such as managers or employees within the corporate organisation. Theorists have examined this replicative dominance transfer of ideas, in which ideas that behave as competing, self-replicating entities are known as 'memes' (Dawkins, 1976).

According to Blackmore (1999) memes are basically ideas, "*they are much like genes or viruses in how they spread...we humans because of our powers of imitation have become the physical hosts needed for the ideas to get around.*" (Blackmore, 1999, p.8). Memes (or alternatively the term 'ideas' may be used for ease of thinking) are synthesised from host to host (human brains) and by way of natural selection, evolution logically fabricates memes that 'survive' efficiently and spread widely through their human hosts. As Blackmore states "*evolution may appear to proceed in the interests of the individual or for the good of the species but in fact it is all driven by the competition between genes,*" (Blackmore, 1999, p.8). This view also holds for the study of memetics (the study of memes), for memes serve no greater purpose. In a sentence they are

selfish independent ideas operating only to get themselves copied (note that selfish does not refer to any sort of active human characteristic, rather it is used only to describe the isolated functioning of memetic replication). Blackmore states that "*the selfish replicators are transmitted and copied, and they do this as often as possible, so long as they have the machinery (human brains) and the building blocks they need for that copying.... In this way evolutionary design comes about*" (Blackmore, 1999, p.13).

Although memes were not termed or theorised until recent times, the notion of replication has been examined for many years. As early as 1896 Baldwin posited the idea of a new factor in evolution 'social hereditary' referring to the imitation in society (Baldwin, 1896). Further intrigued with this notion of imitation and instruction, Baldwin later implied that natural selection was not purely a law of biology but pertained to all the sciences, including the mind (Baldwin, 1903). The obvious question thus arises, what helps a meme make further duplications of itself or conversely what inhibits it?

It is our thesis that certain managers gain power through mergers and acquisitions, thus driving the evolution of ideas, shaping the flow of technology, information and tastes. In simple terms, managers (the meme holders) use mergers and acquisitions to enhance their power and in gaining this power managers unconsciously provide an improved medium through which their memetic 'stories' may be replicated (the term story refers to the managers vision or ideas of the future (memes), and the ideas they shape through their instruction upon people, products and services). The important notion in this theory is that managers are not, except by coincidence, acting so as to maximise shareholder wealth, but are attempting to gain power through acquiring more strategic resources, thus enabling story (meme) replication. Using an analogy from science, M&As can be seen as an agar plate through which stories are cultivated, easily replicated and ultimately imitated throughout human communities.

Having introduced the notion of memes it is not the intention of this paper to examine in depth the theoretical constructs built about the study of memetics, rather it is applied to suggest that there are alternative reasons that may be offered for human action and that managers have ideas (stories) and gain power whereby they are better able to express such ideas (creating replicative opportunities), with mergers and acquisitions used as an improved environment for replication. As Powell states, "*The stereotypical competitive market is the paradigm of individually self interested, non-cooperative, unconstrained social interaction.*" (Powel, 1990, p.302). As such markets (and more specifically mergers and acquisitions) have powerful incentive effects for they are the arena with which each party (managers in particular) can fulfil its own internally defined needs and goals, in doing so providing memes with an environment that is conducive for further replication.

This discussion is developed further following a three-prong approach; first looking at the need for managerial power as motivation for M&As; secondly how this power struggle relates to the 'battlefield' of corporate asset allocation, allowing 'stories' to be told and replicated; and thirdly we pose questions and develop hypotheses, which provide direction for future research to confirm our alternative view of M&A activity relative to memetic transference.

## 2 Power

Who holds the power in this community? Is it the politicians or the highly respected citizens or alternatively is it the managers who run the corporations, which hold the most assets, have large cash flows and employ the most resources? The notion of power has alarming implications for financial theory and its overriding objective of shareholder wealth maximisation. Organisational power is the main focus in this section, as organisational power is the type of power that is derived from an individual's position in the firm and the manipulation over important organisational resources. *"These organizational resources can be tangible, such as money, work assignments or office space, or more intangible, such as information or communication access to other people."* (Riggio, 1996, p.375).

According to Mitchel, Agle and Wood, *"most common definitions of power, at least in part derive from the early Weberian idea"* (Mitchel, Agle and Wood, 1997, p. 865), that power is the probability that one actor within a social relationship would be in a position to carry out one's own will despite resistance (Weber, 1947). Pfeffer rephrases Dahl's (1957) definition of power as *"a relationship among social actors in which one social actor, A, can get another social actor, B, to do something that B would not have otherwise done."* (Pfeffer, 1981, p. 3). Both Pfeffer and Weber acknowledge that power may be awkward to define, but it is not that arduous to recognise- it seems that it is the capacity of those who posses power to bring about the effect they desire.

Power orientated behaviour refers to individual actions aimed primarily at acquiring and using power and can be catergorised into potential power and power (Kotter, 1979). Recognition of this behaviour is useful, as psychologists who have studied managers have said that power (potential or not) is relevant and important, because managers tend to have a high power motivation (Prince, 1972). Logically, the field of psychology dominates the power literature. Correspondingly, academics recognising the importance of power in the study of finance have made some attempts to integrate psychology literature with financial literature in their explanations of the financial marketplace. For example, theories such as 'Empire Building' (Lubatkin, 1983) have been assembled around the very notion of power and it's corresponding effects. This corporate control theory of M&A motivation has arisen in order to account for possible shareholder/management goal deviations. According to this theory, mergers and acquisitions are planned and executed by managers who thereby maximise their own utility instead of their shareholders. Although seemingly understudied, the common thread in the 'Empire Building' theory, specifically the maximisation of manager's goals subject to constraints put upon them by the capital markets, is most supported in the literature.

Whilst *"such alternative merger and acquisition motives have received only modest attention"* (Trautwein, 1990, p. 189), Mueller states, *"there seems little doubt that the motives behind merger activity are not simply those that lie behind the more typical investment decisions, but are overlaid to a considerable extent by management attitudes and certain rationale that does not accord with our assumed over-riding management objective of shareholder wealth maximisation."* (Mueller, 1970, p. 189). Newbold (1970) found that the reasons given by top managers for initiating merger motives did not explicitly include maximisation of shareholder wealth (Newbold, 1970), commenting and expanding on Newbolds report, Mueller (1970)

summarises by stating that, "*the reasons given by managers for merger and acquisition activity appear to be lacking in precision*" (Newbold, 1970, p. 197).

In terms of M&A activity, "*the heightened proclivity of larger firms to enter alliances has led some to conclude that the quest for market power may be an important consideration in such ties*" (Pate 1969, Berg and Friedman 1978). Hence, it is possible that many takeovers and mergers are built around the battle for internal power. They have a direct bearing on power struggles with the increased industrial might for the company translating into greater influence, prestige and pay for the executives. More importantly, M&A's are built around the battle for external power. By combining, corporations aim to enhance their market might, their innovative strength, their bargaining muscle and their ability to influence economic actors. To the extent that size is power, the merging and acquiring corporations have become more potent and the weakening of anti-monopoly intervention under open market focused politicians has opened the floodgates to a massive wave of activity in communications, financial services and drugs amongst others.

The theme still remains that managers regularly acquire and use power. They do so intentionally and wilfully, as well as instinctively and unconsciously. They exploit a variety of methods in this quest, some of which are well known, others of which are not. Indeed, so powerful have contemporary organisations and their managers become that economists such as Galbraith have overturned conventional economic theory about consumer sovereignty in brace of a revised sequence, with substantial corporations molding demand, fixing prices and planning production so as to subsume the market (Bell, 1975). In a summary of earlier studies concerned with the acquisition process, Power (1983) reported that the acquisition practice was not a comprehensively rational decision, suggesting problems such as suppressed uncertainty, lack of planning, political influences, varying process participants and no agreed upon acquisition criteria. Song (1988) gathered evidence that supports the assertion that senior executives' background plays a role.

*"A successful manager, by the nature of his assignment as a manager, has access to considerable power over human beings in an economic sense, in an educational sense and also in a propaganda sense"* (Mueller, 1979, p. 162). Thus, controlling managers usually succeed in securing compliance with their decisions simply because they possess the organisational power, which is obviously enhanced through further M&A activity. In interviews carried out by Kotter, he states that "*although no executive would admit it, at least publicly, they all spent considerable time in activities that were at least partially aimed at acquiring or attaining power for themselves*" (Kotter, 1979, p. 4) further adding that "*Managers were willing to admit privately that power dynamics were an important part of their work and that seemingly political behaviors are usually considered a way of life at the top levels of the organization.*" (Kotter, 1979, p. 385).

This of course does not mean that power is the only goal of managers, or that power is a fixed pie that businesses and individuals fight to divide, or that human interactions are reduced to a power nexus. Rather, it is merely to show that the efficiency based theories do not adequately explain corporate M&A behaviour, suggesting that it is power not reasoning that often drives the market and those in it. Extending this reasoning toward the influence of memetics we suggest that as managerial power increases, ideas that are transferred via such managers are ascribed

greater credibility. Therefore, where managerial power plays a substantial role in a merger or acquisition ideas are transferred through the manager into the new combined entity, such that their story is better able to be told. The interaction of ideas and managerial power may be illustrated if a correlation can be established between 'managers who have power' and the 'replication and sustainability of ideas or stories' for which they are agents (carriers). This issue shall be visited later in a discussion on how thought contagion is influenced by power.

### 3 Story

Who decides whose story is told? For instance, Blackmore asks, "*why do we have fax machines? Why Coca-Cola cans and wheelie bins? Why windows 98 and felt-tip pens. Because we want them is not a satisfactory answer because we need them is clearly untrue*" (Blackmore, 1999). In support of Gelb (1997), Williams (2000) reasons that the application of memetics and business offer potential explanations as to "*why some theories take off and come to dominate in our own culture for periods of time, perhaps even when there is a lack of empirical evidence to support them*" (Williams, 2000, pg. 273).

A further study, which brings together business and memetics, provides interesting insight into the area of market efficiency. Frank (1999) applies memetics to financial markets by asking the question "*Do markets evolve toward efficiency?*" Based on memetic theory Frank reasons, "*If there are many possible competing financial memes and only a small percentage of those memes are economically sound, chances are that the most psychologically appealing memes are not the most economically sound...*" (Frank, 1999, pg. 9). It would therefore make sense to rule out "economically sound" as a criterion for ideas or stories that are likely to be successful. Following from the previous section, it is our proposition that the stories seen in society are the ones told by people with the most power (replicative power- derived through organisational power).

Logically, memes that induce thought contagion behaviour in their carriers that tends to reduce rival memes will be fitter than those that do not provoke such behaviour (able to survive longer and replicate further), since they will have more resources for themselves. The distinct outcome being that a group of hosts with various memes will tend towards homogeneity resulting from the imposition of the stronger meme and elimination of all non-conforming memes (Boyd & Richerson, 1985). This notion has interesting connotations for the influential role of money in persuading individuals and groups to conform, as powerful companies (those with strong market positions) have the ability to coerce and initiate self-fulfilling prophecies. That is, if they say that something will happen then it often will, simply because of the power they possess. It is also directly applicable in corporate control situations where management teams compete for increased market power in order to dictate the flow of information and resources.

It has also been suggested that memes group together to aid survival. Such groupings of memes are referred to as 'meme-plexes' (Blackmore, 1998). This grouping of memes makes sense in terms of natural selection; if a story meme could group with a defence meme, it could decrease the chance of outside memetic attacks (i.e. the fight for limited space in the mind), thus enhancing its replicative ability. Perhaps this is why we see managers of possible target firms set in place popular anti-takeover strategies to help shield their story memes. 'Poison Pills', 'Lock-Up' strategies 'Pac Man' defences, 'Shark Repellent' defences, 'Sale of the Crown Jewels',

'Scorched Earth' Strategies and 'Supermajority' amendments are only a few of the prevalent anti-takeover defence mechanisms that have arisen in recent times. These anti-takeover defence strategies occur frequently in the business world, even though it has been proven extensively that on average target companies are paid a substantial premium and shareholders of these companies experience positive wealth gains when they are taken over, (Jensen and Ruback (1983), Loughran and Vijh (1997), Travlos (1987), Warnesly, Lane and Yang (1987)).

Power is a cultivator through which thought contagion breeds, evidenced by various ideas that spread through a society by those hold power, reverberating the notion of memetic transfer.

Remember that "*memes are ideas that propagate themselves around the world by jumping from brain to brain, memes are stored in human brains and passed on by imitation, individuals learn from society by imitation instruction. Ultimately, human life is permeated through and through with memes and their consequences*" (Blackmore, 1999, p. 6).

Historically religion has been used as an example of thought contagion (memetic transference), with large powerful groups creating different beliefs in god. In the past, religion that promoted large families were successful because they created more people to adopt the faith from their parents, this example draws resemblance for the memetic advantages of M&A activity, with religion and ceremonies very much likened to business ideas and practices, both of which are spread by one person copying another, creating large fellowship and belief.

As discussed, the way that we see and understand the world is guided by thought contagion (or the replication of memes). Certain characteristics enable thoughts or ideas to enhance their replicative ability. As stated by Dawkins (1976), the criterion for a successful replicator (successful in the sense that they create multiple duplicates) consists of three elements: Fidelity, Fecundity and Longevity. When examining mergers and acquisitions on this three-factor (fidelity fecundity and longevity) success scale, the theme remains that greater market strength places companies in a position where they are further able to supply the market with their ideas.

Logically, mergers and acquisitions allow this to happen as they have the enhanced power and distribution channels to make a vast number of exact copies (fidelity & fecundity), with the pooled research, design and marketing capabilities to defend against competitors and ensure continued expression (longevity).

The overall fitness of a meme can be determined by identifying whether it will maintain within an individual's memory and spread to other individuals or be eliminated. M&As illustrate the effect of this as they impose changes on business environments, overpowering the acquired companies and enhancing their story. De Jong (1999) illustrates how institutions act as filters for new concepts, which filters reflect dominant behavioural and conceptual practices. Concepts that survive organisational restructuring (M&As) may become institutionalised, forming part of the story being told.

When acquisitions and mergers occur the acquiring company is also able to increase its organisational networks, which aids replicative ability. As noted "*an organisation's power is determined less by its internal resources than by the set of resources it can mobilise through its*

*contacts. The more such contact the firm has, the better it is 'plugged in' to the key task and influence process of the industry" (Galaskiewicz, 1979, p. 478).*

Resulting from this, once those in power decide the direction of their environment, in general it is very tough for people to stop the flow and assume the 'old way' of life, (or if you like the replaced meme way of life). It is important to note that the competitive game here is not played at the level of the firm, but rather at the level of the idea itself. Victory in this 'game' means that the idea becomes adaptive (adaptive in the sense that the idea or meme creates a contagion type effect with a higher probability of 'survival' and 'reproduction').

In terms of managers replicating their meme driven stories (individual fitness) perhaps we should invoke the analogy of two people sitting on respective ends of the rope in a tug-o-war competition, each heaving like mad to serve their own selfish replication, with the stronger competitor eventually able to drag the weaker competitor: power creating compliance! Correspondingly, M&As are an excellent launch pad for memetic transference, for the power they build helps the propagation of ideas and the continual replicative muscling of such ideas.

Clarifying the relationship of thought contagion and power, M&A's are transactions in which organisational power is transferred, where the acquiring firm holds the power and in turn memetic transference sees that acquirer's story is replicated. In biological terms Mokyr states that, "*the environment into which seeds are sown is of course the main determinant of whether they will sprout*" (Mokyr, 1990, p.299). Whether or not comparisons between evolution and cultural or management progress are valuable has been the focus of a substantial and spirited debate in the literature, as Mokyr states "*the parallels are inevitably incomplete and do not provide the researcher with sharp analytical conclusions.*" (Mokyr, 1990, p. 275). Additionally, it must be noted that at present we cannot predict exactly how stories will evolve over time, we can only stand back and watch with a greater understanding as they unfold. We can however find deep, underlying reasons governing the unpredictable flow of such activity. Being at least in part, that those in power have the higher ability to dictate the flow of information and resources, ultimately shaping the products, services, systems and ideas that the economy maintains.

## 4 Future Direction

Our major assertion is that memetic transference is enhanced through M&As, where M&As are driven by managers need for power. Our stance however, has not come about through a clearly defined road map of empirical results. Rather, we have presented an alternative view of corporate restructuring, suggesting that M&As help to shape the evolution of ideas within society. In this final section we pose some key questions and present four hypotheses that provide direction for further research of this subject. Our intent is to offer a new thread of motivation toward understanding the complexities of behavioural finance.

Fundamental questions that one may naturally be compelled to ask are: Is it in fact managerial need for power that drives M&A or are memes themselves the drivers of such corporate activity? Contrastingly, do memes motivate managers to acquire power so that they can be replicated? Are memes indeed better replicated through M&As? If M&A's are driven by stories rather than the need for power, does the story being told have any relation to subsequent performance of the

newly combined corporate entity? Does industry classification offer any explanatory value in the replication and sustainability of stories that are told through M&A? Do M&As and their resultant stories refute our theories of memetic transfer by creating an entirely new story that cannot be explained by current logic?

Establishing quantifiable variables to test when addressing issues in behavioural finance is integral to producing acceptable conclusions. We propose the following testable hypotheses and offer reasoning as to how each one supports the framework of M&As and memetics.

*Target firms replicate the ideas or stories of the acquiring firm.*

In support of our discussion on power we reason that the dominance of the acquiring firm in a takeover transaction promotes replication of its stories and ideas. A finding in support of this assertion would suggest that M&As are an improved medium for memetic transference, where replication of the dominant entity's stories and ideas are enhanced.

*M&A's with less debt in their capital structure provide a superior medium for memetic transference.*

Given that there are agency concerns surrounding M&A, it has been posited that higher levels of debt augment management discipline, by keeping managers honest and focused on fulfilling their company debt obligations (Jensen, 1976, Jensen & Ruback , 1986). As debt requires servicing, it acts to constrain managers such that less debt creates greater management sovereignty, ultimately enabling managers to merge or acquire other firms and by doing so they improve the replicative chances of the memes that they hold.

*M&A's with low managerial equity provide a superior medium for memetic transference.*

Empirical evidence also supports the notion that managerial ownership acts to align managers' interests with the interests of shareholders (Shinn, 1999). Equity ownership serves to constrain managers against wealth destroying activity, which may take the form of an unsuccessful M&A. However, low equity ownership firms are more likely to experience higher agency costs such as managers seeking to gain power through M&As therefore providing memes with greater replicative opportunities.

*Where executive compensation increases in relation to share price performance, stories told by such managers' display enhanced replicative ability.*

Managers of firms that experience positive price performance are commonly rewarded with greater compensation in the form of bonuses, increased salaries, and more frequently stock options (Ofek and Yermack, 2000). Such executive compensation brings with it the credibility for the manager and his or her ideas. We reason that the credibility ascribed to the manager serves to enhance the propagation of stories and ideas for which the manager is an agent (carrier).

## 5 Conclusion

One of the most often asked questions in the M&A literature is why. Why do companies merge with and acquire other companies? From a strictly financial perspective, managers who make these investment decisions are meant to be undertaking wealth maximizing activity. However, the finance literature is replete with evidence that this activity only increases the wealth of the target shareholders. So what is in it for the acquiring firm? Why do the managers of these firms persist in this activity when it is not profitable?

Several theories have developed to explain this ongoing activity. Theories relating to efficiency and wealth enhancement have been largely discounted. Power based theories, on the other hand, have gained considerable support: Monopoly, empire building, and management entrenchment. But again, we must ask: Why?

Why is the power story so important? Why is it that managers are able to 'finesse' their shareholder wealth maximization responsibilities and instead seek power? We suggest that insight can be gained by an understanding of memetic transference. Memes and meme-plexes, which are successful, are those who replicate. M&A increases management power so that memes can spread more widely. Access to more resources provides memes with more chance to selfishly attempt to replicate themselves. When observed from this perspective, M&A activity appears to be an arena where power is not the end goal, but rather the means to the end. Memetic transference is the end goal.

Empirical testing of memetic transference is never easy. But the testing of several resulting hypothesis about M&A activity offers an opportunity to begin this process.

## References

- Asquith, P, R.F. Bruner, & D W Mullins, 1987. 'Merger Returns and the Form of Financing', working paper, Harvard University and University of Virginia quoted in K.J Martins' 1996 paper: 'The Method of Payment in Corporate Acquisitions, Investment Opportunities and Management Ownership.' *Journal of Finance* **51**, 1227-1246.
- Baldwin, J. M., 1896. 'A New Factor in Evolution'. *American Naturalist*, **30**, 4441-51, 536-53.
- Baldwin, J. M., 1903. '*Darwin and the Humanities*'. Review Publishing, Baltimore, MD.
- Bell, D., 1975. '*Power Influence and Authority*'. Oxford University Press, New York.
- Berg, S., & Friedman, B., 1978. 'Joint Ventures in American Industry: An Overview.' *Mergers and Acquisitions*, **13**, 28-41.
- Berkovitch, E. & Narayanan, M.P., 1993. 'Motives for Takeovers: An Empirical Investigation', *Journal of Financial and Quantitative Analysis* **28**(3), 347-362
- Blackmore, S., 1999. '*The Meme Machine*'. Oxford University Press, USA.

- Boyd R. & Richerson P.J., 1985. 'Culture and the Evolutionary Process'. Chicago University Press, Chicago.
- Bleeke, J. & Ernst, D., (eds), 1993. 'Collaborating to Compete: Using Strategic Alliances and Acquisitions in the Global Market Place'. New York: John Wiley.
- Dahl, R. A., 1957. 'The concept of power.' *Behavioural Science*, **2**, 201 – 218.
- Dawkins, R., 1976. 'The Selfish Gene.' Oxford University Press. New York.
- Doz, Y., 1988. 'Technology Partnerships Between Smaller and Larger Firms: Some Critical Issues.' *International Studies of Management and Organisation*, **17**(4), 31-57.
- Frank, J., 1999. 'Applying Memetics To Financial Markets: Do Markets Evolve Towards Efficiency', *Journal of Memetics - Evolutionary Models of -Information* **3**(2): 87-99.
- Franks, J.R., Harris, R.S., & Mayer, C., (1988). 'Means of payment in takeovers: Results for the United Kingdom and the United States', quoted in K.J Martins 1996 paper: 'The method of payment in corporate acquisitions, investment opportunities and management ownership'. *Journal of Finance* **51**, 1227-1246.
- Friedman, A., & Gibson, R., 'Phillip Morris Co. is bidding \$90 a share for Kraft Inc. in \$11 billion tender offer'. Wall Street Journal, 18 October 1988, p. A3.
- Galaskiewicz, J., 1979. 'Exchange Networks and Community Relations.' Sage, Beverly Hills, CA.
- Gelb, B., 1997. 'Creating 'memes' whilst creating advertising', *Journal of Advertising Research*, **37**(6), 57-9.
- Herman, E & Lowenstein, L., 1988. 'The Efficiency Effects of Hostile Takeovers' in: John Coffee Jr. Lowenstein and Ackerman, eds 'Knights, Raiders and targets: The Impact of The Hostile Takeover'. Oxford University Press, New York.
- Hogarty, T., 1970. 'Conglomerate Mergers and Acquisitions: Opinion Analyses.' *St. John's Law Review*, **44**, Special edition, 317-327.
- Hunt, J., 1990. 'Changing Patterns of Acquisition Behaviour in Takeovers and the Consequences for the Acquisition Process.' *Strategic Management Journal*, Vol. **11**(1), 69-77.
- Jensen, M. C., 1976. 'Theory of the Firm: Managerial Behaviour, Agency Costs and Capital Structure'. *Journal of Financial Economics*, **3**, 305-360.
- Jensen, M. C., & Ruback, 1986, 'Market for Corporate Control'. *AEA Papers and Proceedings*, Vol, **76**. No.2.

Jensen, M.C and R.S Ruback, 1983. 'The Market for Corporate Control', *Journal of Financial Economics*, **11**, 5-50.

Jong, W.M. de., 1999. 'Survival of the institutionally fittest concepts', *Journal of Memetics - Evolutionary Models of Information* **3**(1). <http://>

Kitching, J. 1964. 'Winning and Losing with European Acquisitions'. *Harvard Business Review*, **45**, 85-101.

Kotter, J, P., 1979. 'Power in Management'. Amacom Publishers, USA.

Lorenz, C., 1986. 'Takeovers- at best an each way bet', *Financial Times*, 6<sup>th</sup> January, p. 12.

Loughran, T. & Vijh, A.M, 1997. 'Do Long Term Shareholders Benefit from Corporate Acquisitions?' *Journal of Finance*, **52**, 1765-1790.

Lubatkin, M. 1983. 'Mergers and the Performance of the Acquiring Firm.' *Academy of Management Journal*, **8**(2), 218-225.

Lubatkin, M. & Shrieves, M., 1986. 'Towards Reconciliation of Market Performance Measures to Strategic Management Research'. *Academy of Management Review*, **11**(3), 497.

Mandelker, G. 1974. 'Risk & Return: The Case of Merging Firms'. *Journal of Financial Economics*, **1**, 303-335.

Maremont, M. & Mitchell, R., 1988. 'Pillsbury could be a Grand Coup for Grand Met', *Business Week*, 17 October, p.30.

Markides, C. & Oyon, D., 1998. 'International Acquisitions- Do they Create Value for Shareholders?' *European Management Journal*, **16**(2), 125-135.

Mitchel, R.K., Agle, B.R., & Wood, D.J., 1997. 'Toward a Theory of Stakeholder Identification and Salience: Defining the Principle of Who and What Really Counts.' *Academy of management Review*, **22**(4), 887-910.

Mokyr, J., 1990. 'The Lever of Riches'. Oxford University Press, New York.

Mueller, R. K., 1970. 'Risk, Survival and Power'. Amacom Publishers, USA.

Newbold, G. D., 1970. 'Management and Merger Activity', Guthsted, Liverpool.

Ofek, E., Yermack, D., 2000. 'Taking stock: Equity-based compensation and the evolution of managerial ownership'. *The Journal of Finance*, **55**(3), 1367-1384.

Pate, J.L. 1969. 'Joint venture activity: 1960-1968'. *Federal Reserve Bank of Cleveland Economic Review*, January, 16-23.

- Pfeffer, J. 1981. 'Power in Organisations'. Pitman Marshfield, MA.
- Porter, M. E., 1987. 'From Competitive Advantage to Corporate Strategy.' *Harvard Business Review*, **65**(3), 285.
- Powell, W.W., 1990. 'Neither Market Nor Hierarchy: Network Forms of Organisation.' *Research in Organisational Behaviour*, **12**, 295-336.
- Power, D. J., 1983. 'Acquisition Decision Making', *Mergers and Acquisitions*, **18**(2), 63-65.
- Prince, G. M., 1972. 'Creative meetings through power sharing'. *Harvard Business Review*, **50**(4), 47 – 52.
- Ravenscraft, D. & Scherer, F., 1987. 'Mergers, Sell-offs and Economic Efficiency'. The Brookings institution, Washington DC
- Riggio, R., 1996. 'Introduction to Industrial/Organizational Psychology', 2<sup>nd</sup> Ed., Harper College Publishers.
- Servaes, H., 1991. 'Tobin's q and the Gains fromTakeovers'. *Journal of Finance*, **46**, 409-419.
- Shinn, E. W., 1999. 'Returns to acquiring firms: The role of managerial ownership, managerial wealth, and outside owners'. *Journal of Economics and Finance*, **23**(1), 78-89.
- Song, J. H., 1988. 'Diversification Strategies and the Experience of Top Executives of Large Firms'. *Strategic Management Journal*, **3**(4), 377 – 380.
- Trautwein, F., 1990. 'Merger Motives and Merger Prescription'. *Strategic Management Journal*, **11**, 283-295.
- Travlos, N.G., 1987. 'Corporate Takeover Bids, Methods of Payment and Bidding firms Stock Returns.' *Journal of Finance*, **42**, 943-963.
- Warnesly, J.W. Lane, W.R. & Yang, H.C., 1987, 'Gains to Bidder Firms in Cash and Securities Transactions,' *The Financial Review*, **22**, 403-414.
- Weber, M., 1947. 'The Theory of Social and Economic Organisation.' Free Press, New York.
- Williams, R., 2000. 'The business of memes: memetic possibilities for marketing and management', *Management Decision*, **38**(4), 272-279.

**A REVIEW OF: JASON POTTS (2000) THE NEW EVOLUTIONARY  
MICROECONOMICS: COMPLEXITY, COMPETENCE, AND ADAPTIVE  
BEHAVIOR**

**Cheltenham: Edward Elgar, 2000 (ISBN 1 84064 543 1)**

By  
**Jack J. Vromen**

Department of Philosophy  
Erasmus University, Rotterdam

Jason Potts's *New Evolutionary Microeconomics* is at once an ambitious and imaginative piece of work. It aims at providing a micro-economic foundation for all schools and traditions within heterodox economics. Anyone familiar with the 'family quarrels' within heterodox economics will immediately grasp that this is no small feat. To achieve his aim, Potts brings together various bits and pieces of notions, ideas, theories and models, dispersed over many different disciplines. Instead of letting his readers get lost in the bewildering complexities and perplexities of the various strands of literature that he draws on, Potts gently takes his readers by the hand and navigates them through the labyrinth. Potts is able to do this, because he found a key to connect all the strands. This key he calls the geometry of economic space. Instead of the orthodox Walrasian notion of an integral space (related the mathematical notion of a field), in which all elements are connected to one another, Potts proposes a non-integral space as a unifying notion for heterodox economics. Adopting a non-integral space makes it possible to account both for particular interaction and for time irreversibility. Replacing an integral space by a non-integral space implies that the focus shifts from elements to the connections between elements. Accordingly, what primarily evolves in Potts's evolutionary microeconomics are connections between elements rather than the elements themselves.

This is an analysis at an extremely high level of abstraction, and Potts is aware of it. But already here, at this high level of abstraction, doubts begin to creep in. Orthodox economists often accused heterodox economists of not getting beyond putting forward the vague and unproductive notion that 'everything depends on everything'. And, indeed, many heterodox economists have strong leanings toward a holistic, non-reductionist type of approach. But if we are to believe Potts, then the accusation is more appropriately levelled against mainstream economists themselves. For the notion of integral space that Potts assigns to mainstream economics seems to imply the belief that every element is connected with all other elements in the field.

At a lower level of abstraction, I also found Potts's depiction of *hetero economics*, his own preferred notion of the micro-economic agent, questionable. Connections here are not defined interpersonally, but intrapersonally, as the cognitive technology an individual agent has at his disposal for combining his resources. Potts here explicitly invokes the image of Robinson Crusoe, an image that heterodox economists often have objected to. Only later on other individuals that the individual agent may interact with appear on the scene. Potts simply assumes that the logic by which agents decide whom to interact with is the same by which agents interact with the environment. This seems to run counter to how many heterodox economists view the relation between individual agents and social structure. These economists stress that the social environment in which agents are embedded affects the ways in which agents think, argue, behave

and interact with one another. This is taken to imply that the social environment needs to be taken into account right from the start. Potts pays no attention either to cultural evolution, to cultural transmission of memes (or similar units of cultural transmission) and to how this affects human behaviour. Once again this calls into doubt whether Potts really comes up with a microeconomics that is acceptable to heterodox economists of all stripes.

This does not diminish my admiration for the accessible way in which Potts introduces quite complex theories and notions, however, and especially for the comprehensive and creative way in which he links these theories and notions with one another. Graph theory, complexity theory, chaos theory, self-organisation theory and the notions of genetic algorithms and dynamic efficiency all get a succinct, but clear coverage. Without eschewing some elementary technical matters altogether, Potts succeeds remarkably well in showing in an informal way how these theories and notions can complement one another. Potts makes clear why he believes that the future lies in multi-agent computer simulations. All this is done in vivid prose. Potts's book demonstrates that complex and highly abstract notions and issues can be discussed in an engaging and entertaining way.

The only hesitation I have here is that sometimes it seems that Potts reads too much in mathematical formalisms. Potts sometimes argues as if each mathematical formalism has unique ontological implications. I think this is a bit naïve. Mathematical formalisms can be given many different interpretations. And sometimes the one mathematical formalism can be translated into another. For example, contrary to Potts's general point that the ontological implications of the mathematical notion of a field are altogether different from those of graph theory, it seems that a field can be described in terms of graph theory (as Potts himself acknowledges; see note 4, p. 81). Minor reservations as these notwithstanding, I warmly recommend Potts's book to anyone who is in for an intellectually adventurous attempt to synthesise various pioneering non-mainstream strands of theorising and modelling.

# **THE ECOLOGY OF INTENTIONS: HOW TO MAKE MEMES AND INFLUENCE PEOPLE: CULTUROLOGY**

**By**  
**Adam Westoby (1944-1994)**

**Foreword**  
**By**  
**Daniel C. Dennett, Director, Center for Cognitive Studies**

Whether the meme meme deserves to flourish depends in part on how well it can account for itself. Adam Westoby's "The Ecology of Intentions" makes a fine double contribution to this reflexive task, both as analysis and example. If you read it, you will see for yourself the variety of original and incisive ideas about memes that prompted me to become its willing vector. And if you read it, you will yourself provide another data point measuring the power of memes--whatever their intrinsic virtue--to spread by harnessing human minds to the task of their further replication. The draft before you is not just unfinished and unpublished; it is full of blemishes, gaps, half-baked ideas that distract us from the best ideas in it. Life is short, so why will you read it? Perhaps to see how Westoby's memes got as far as this (the Zahavi Principle at work in the memosphere).

These memes exited their nest through an extraordinary bottleneck: the relentless activity of a single unparalyzed finger on a word processor keyboard. That does not make them good ("for us"), but it does make them powerful; the task of getting them out dominated Adam Westoby's last years. Shortly after his death, his brother, Mark Westoby (some of whose own work in biology is briefly discussed by me in Darwin's Dangerous Idea, pp. 234-5), sent me the manuscript and the accompanying "brief background" account, with a request that I advise him about how to get his brother's work to the right minds. Some months later I found the time to look at this uninvited but curiously inviting package, and began to put in motion the process in which you are now participating. I considered undertaking the editorial task required to fit Westoby's manuscript for regular publication, but rejected it; the task was huge, and I had my own pile of projects to tend to. In the end, after sharing the manuscript with a few of my favorite advisors and informants on such topics--Richard Brodie, the late Ben Cullen, Aaron Lynch, Alex Rosenberg and Don Ross--and considering their advice, I offered to make the rough draft available as a Working Paper of the Center, which is the form in which you now see it, and in which it may be cited. Mark Westoby has provided me with further items about his brother and his work, which appear as appendices here. I never met Adam Westoby, nor have I met Mark yet, but some of their memes are now among mine, and I am pleased to have them on board.

The Center for Cognitive Studies has become something of an informal, self-selected depot for current work on memes, and I am sorry to say that I have been simply unable to filter, evaluate and transmit the material that has been sent to me so far. Piles of manuscripts lie as yet unread on my shelves--other authors should be warned of this--and Westoby's vehicle is the first that I now send on its way with this equivocal--but effective?-- endorsement. I wish I had had the opportunity to discuss it with Adam Westoby, since I would have tried to adjust his vision of several of the philosophical topics he boldly tackles, and also would have asked him to expand

on the passages I found most suggestive and illuminating. But that is just a selfish reaction; every reader can execute one editorial function or another on this rich set of materials.

We at the Center would appreciate being informed of any discussions and commentaries of these ideas.

### **Brief background to the manuscript "Ecology of Intentions"**

Mark Westoby

March 1995

The manuscript is of about 50,000 words and was written by my brother Adam Westoby. The current version dates to July 1994. Following that time, Adam's remission from myeloma ended and he died in November 1994. The July 1994 version represents a revision following comments he received on an earlier version, dated December 1993.

The manuscript takes the "meme" metaphor and applies it to a range of topics in cultural history, education and political economy. The metaphor is also developed beyond its usual form, notably by grafting "intentionality" on to it.

A brief outline of Adam's personal history may help to explain the mixture of ideas and knowledge he brought to the manuscript. In 1964 he went up to Balliol College Oxford, initially to read physics but switching within a month or two to Politics, Philosophy and Economics. He graduated with BA Hons I in 1967, and during 1967-8 did a one-year BPhil with a thesis on Haeckel's relation to Marx. By the mid-60's he was active in Trotskyist politics. Much of his time went on activism (newspaper sales, union organization etc) and on Marxist theory for some years. During this period the question of the nature of the State in the Soviet Union (also Cuba, China and the soviet satellites) was a key issue among Trotskyists. By the mid-70's he was no longer active in Trotskyist organizations (there had been a series of splits). His interest in the structure and political nature of these states continued, however, especially in the three major books of 1981, 1985 and 1989 (see short publications list below).

From 1970 he was appointed to a Lectureship in the Dept of Education of the Open University, a position he held for the remainder of his life. He developed readers and course materials in the sociology and economics of education.

In 1973 he was diagnosed with syringomyelia, a progressive spinal cord disability. He underwent a number of operations to slow its progress, but by the 80's he was spending much of his time in a wheelchair. From the late 70's on he lived with Sabitha Hasan, a lawyer (now judge) who had been disabled as a child by polio. They had two sons, born 1984 and 1987. During the 80's he was involved also with his father's progressive disability and eventual death (1988) from motor neuron disease. His experience of his young children, combined with his sharp-edged relationship with his own father, further shaped his thinking about the transmission of ideas. In 1989 he was diagnosed with multiple myeloma, contracted pneumonia while under chemotherapy, and was not expected to survive. He emerged alive, though much weakened, and in clinical depression during 1990-91. By this time he could move only in a powered wheelchair,

could sit up only for a few hours at a time, and could use a word processor only with one finger of his right hand.

During 1989-90 the whole face of Soviet-bloc politics had changed, and it had become impossible for him to recover his former expertise in this area. At different times from 1989 onwards I had conversations with him about memes (my own background is in evolutionary biology), extending first over many weeks in hospital, then days at a time at his home, interspersed with phone and email conversations from Australia. The manuscript "Ecology of Intentions" takes the metaphor of memes and applies it to a range of themes from Adam's continuing interests, including some large ones (the subjectivity of meaning; money and the social surplus).

## **THE ECOLOGY OF INTENTIONS: HOW TO MAKE MEMES AND INFLUENCE PEOPLE: CULTUROLOGY**

**Draft: please comment**

By  
**Adam Westoby**  
75 Chambers Lane  
London NW10 2RN  
tel 081-459 4668  
fax 081-451 1641

### **Note to readers of this draft**

Thank you for commenting. This note is to help you do so more effectively.

The date of this draft is printed above. It supersedes previous drafts.

Errors. Please point these out. Do not fear that your explanations will be too elementary.

Main text and footnotes. I hope to make the main text as readable as possible, consigning details, problems etc to the notes. Help in this regard will be especially welcome.

Footnotes. These are a miscellany of:

- References to sources, sometimes ill-remembered
- Notes and reminders to myself, and
- Short or unsatisfactory versions of what might go in the text.

Drafting/editing comments are placed in square brackets [thus]. They are mainly intended to be intelligible to me.

References/bibliography. This is very much a working version, both overgrown and incomplete.

Conversely, the fact that I have annotated an item does not imply that I have read all of it. Some items got there via book reviews etc.

## In brief

This draft explores some ideas that see culture as organic.

Viewed as living things, cultures as a whole, and their parts, depend on at least two types of adaptation:

They must be adapted for inhabiting human psyches; and

They must survive the mortality of individual humans.

Culture must be adapted for direct or indirect transmission from living humans to other living humans; much of it is learned behaviour. It must be adapted, crudely speaking, for both parasitism and reproduction. Cultural objects must dwell within individual psyches, and they must pass between them.

Much of what follows is a sort of "natural history" of cultural objects, paying particular attention to these two features and the interplay between them.

I also have a practical - or, if you prefer, ethical - aim: to help us discuss what forms of culture we humans can most happily share our world with.

## Phases and metamorphoses

Because cultural objects need both to get inside individual humans, and to pass between them, they exist (like carbon-based forms of life) as sequences of differentiated phases. This has some general consequences which it may be useful to sketch at the outset:

"Mental" phases, as experienced by human subjects, alternate with phases which have evolved to transmit well to, from and between subjects. We humans experience the contrast between mental and external phases as a gulf between subjective and objective. My discussion concentrates on evolution of the transmissible, objective phases of cultural objects.

Cultural forms arise - like embryos and organisms - by evolutionary accretion. Historically later forms are modifications of earlier structures, and actual structures embody the consequences of earlier development. The cultural forms experienced by present day humans incorporate long evolutionary histories.

This, in conjunction with the relatively fixed scale of individual humans (for example, our limited brain size and life span) imposes a cumulative process of abbreviation on cultural transmission. For example, modern school curricula must recapitulate in a few months developments that first took shape over thousands of years of cultural history.

This places a premium on forms of transmission which yield advantages in compression of this sort - such as, for example, extended symboling, reusable artefacts, or money.

Corresponding to "parasitism" and "reproduction" in cultural objects are two enigma-areas of the human sciences:

- the problem of meaning, or that of the subjective versus the public; and
- the problem of social reproduction.
- These problems are linked. Indeed I might recast them as follows:
- the problem of meaning may be approached as that of the separation out of persons, and the development of memory; and
- the problem of social reproduction is that of the cultural development of neophyte persons, or that of education.

## Memes

I return to these problems in more detail later. However one approaches them, they are problems of complexity, of wholes being more than the sum of their parts. If I mention them now, it is only to introduce an "atomic" device which I think can help us make them a little more manageable: memes.

This derives not from physics but from biology. To begin with evolutionary biology analysed variation and natural selection among organisms. More recently its focus has shifted towards genes - the units of biological inheritance. Biologists remain divided on how best to think about relations between genes and organisms, and about how far life is naturally organised as organisms.

When questions analogous to these are transposed to human cultures they give rise to severe difficulties. I try to make progress by sidestepping them, at least initially. To do so I borrow a term invented by the biologist Richard Dawkins: memes, the units of cultural inheritance and selection.

What sorts of things are memes? Dawkins defined memes by analogy with genes. Just as, in biological reproduction, genes replicate themselves and "jump" from parent to offspring, surviving beyond the limited lifespan of individual organisms, so we may think of memes as a new form of life, leaping from mind to mind in human cultural transmission. Memes are "patterns of information that can thrive only in brains or the artificially manufactured products of brains - books, computers and so on". Memes replicate themselves (Dawkins holds) like genes: "by imitation ... in the broadest sense". They share certain fundamental characteristics with genes: "longevity, fecundity and copying-fidelity" (1976, 208). However, memes also vary, and the more successful variations proliferate. {Footnote 1}

Memes range from relatively slight - a few whistled notes of a popular tune - to great interconnected assemblages of memes. Such collaborating assemblages of memes are analogous, Dawkins argues, to the "survival machines" (that is, organisms) in which biological genes congregate to propagate themselves from generation to generation. We could, for example,

regard "an organized church, with its architecture, rituals, laws, music, art, and written tradition, as a co-adapted stable set of mutually-assisting memes" (Dawkins 1977, 212). I shall come back to the question how memes are incorporated in larger cultural objects. For the moment I ignore the distinction between simpler and more complex memes and treat them all interchangeably. The letter "i" is a meme; so is a paperback, a supermarket orange, a character in a film, a funeral, a screwdriver, a military command, a firework display, a quadratic equation, or a political party. Writers such as Daniel Dennett have pursued the idea that memes proliferate just because they are good at proliferating, not necessarily because human beings want them to. At first, Dennett concedes, we resist the idea of our brains "as a sort of dung heap in which the larvae of other people's ideas renew themselves, before sending out copies of themselves in an informational Diaspora". But this view captures an important truth. And it raises the question "Who's in charge, according to this vision, we or our memes?" As Dennett points out, there is no simple answer, and in his discussion the question leads (among other places) back to the enigma of the human self. (Dennett 1991, 202-3)

An advantage of the meme meme (as Dawkins rightly calls it) is that it gives us a relatively commonsense way of looking at problems of subject and object, personal and public, without collapsing one into the other. Thinking of memes as life forms which parasitise or domesticate humans helps avoid one-directional teleology. Memes make use of humans in senses just as real as those in which humans make use of artefacts and ideas. And it helps us to avoid humanist prejudices: by keeping memes conceptually separate from people, it allows us to examine some of their intrinsic characteristics, which cannot be reduced to those of human individuals.

My core argument is that the principles of cultural selection, as encapsulated in the notion of "memes" and their evolutionary interplay, may have much to offer the human sciences - a little as the principles of natural selection have increased the understanding of biologists. I discuss various ways in which memes affect people, in which people affect memes, and - particularly important - in which memes interact with each other in the larger ecology of memes.

Memes serve as my working answer to the question "What is a cultural object?" I know that there are powerful objections to thinking about cultures as collections of objects, and I do not mean to dismiss these. (Later on (Chapters) I look at some of the problems thrown up by a memetic view of human - and other - cultures.) I simply want to suggest:

- That such a view can illuminate some aspects of culture; and
- That some of its insights may help us to develop cultural forms with which we can more pleasantly share the world.

I take it as widely agreed that human culture is not yet so "user-friendly" as to be open to no improvement. I want to suggest ways in which we could helpfully think of this - as the problem of breeding better-domesticated memes.

By taking Dawkins' meme as a starting point I don't mean to dismiss its various near-synonyms and connected ideas. I discuss some of them ("culturegens", "brain bugs", "viruses of the mind" and the epidemiology of beliefs, etc) in Chapter X. One of the things I should like to do is

"evolve" (that is, vary and select from) the meme meme, to help it function and proliferate a little better. {Footnote 2}

## Varieties of meme adaptation

"Tell me, how did you love my picture?" - Sam Goldwyn

What types of meme might we expect to be successful?

Let us look at some common meme adaptations, using the convenient shorthand of biological functionalism. What sorts of things might memes do if they "wanted" (and for our purposes they all do) to increase their success among the population of memes? More specifically, what sorts of things might memes do:

- In order to make better use of human beings, but also
- In order to succeed better in their interactions with other memes?

Memes' environments, after all, consist largely of other memes.

One - but only one - possible way is for a meme to adapt so as to be more palatable or more desired by more human beings more of the time. A very large number of human artefacts are of this sort.

Consider the orange laid out in neat peeled segments next to my keyboard. It is an outcome of combined genetic and memetic selection, and allows us to compare them.

What is chiefly interesting to me about my orange is its convenient packaging of pulp and juice, high in sugars and distinctively flavoured. I enjoy eating it; it refreshes my concentration; and it forms a healthy alternative to biscuits.

The orange however, began its career with quite different functions. Fruiting developed (starting about million years ago) as a secondary adaptation of the sexual reproduction of flowering plants. Natural selection acted in favour of variants which surrounded their seeds with tissues appetizing to animals. Two factors, in particular, acted through the foraging behaviour of animals to drive the evolutionary development of fruity organs around seeds:

- Better seed dispersion allowed offspring plants to populate the surrounding terrain more widely
- And proliferate better (for example by competing less intensely for light) improving seeds' chances of rooting by depositing them on the ground well mixed with animal excrement

Natural selection, however, explains only certain basics of my orange. It has also been greatly changed, in the last few moments of its evolutionary history, by cultural selection. The package of genes has become a meme.

In the few centuries that orange trees have been cultivated, humans have applied increasingly knowledgeable selection, driving evolution at an accelerated rate in pursuit of enhanced desirability to humans. The characteristics that have been modified or emphasised by human selection include:

- Colour: attractiveness to animals has been enhanced to attract consumers
- Sweetness and taste: quite different selective pressures have been applied to my eating orange than, for example, to oranges for marmalade or Cointreau
- Seedlessness: the original *raison d'être* of the orange's evolution has become an inconvenience. The very few seeds that may make it into my intestinal tract will pass on into an urban sewage system in which they have no hope of germinating.
- Uniformity: the varied shapes and sizes of fruit on wild orange trees have yielded to the astonishing similarity of Grade I eating oranges. This allows them to be shipped on carton with standard indentations, to be sold for a standard price, and to reduce the cost of distribution. The uniformity in the supermarket tray arises from cultural selection both between orange tree genes, and, at grading, between harvested fruit.

And, before the orange slides along the check-out counter to become mine, further cultural selection occurs, in which orange tree genes from round the world - Spain, Florida, Morocco and South Africa - compete to reproduce. They do so, however, not through excrement-smeared seeds, but through quite different memes - such as coins and balance sheets.

## Derivative memes

Markets - with their continual revolutionising of both products and tastes - provide one modern, general forum of meme selection. In markets, money forms the institutional environment in which commodity-memes complement and compete with each other (through market research, product development etc, with their feedbacks on production and consumption). Comparisons between biology and commerce are now commonplace and terminology naturally exchanges: market research creates economic "niches"; superseded commodities become "extinct", etc.

Palatability or desirability to humans is not necessarily direct. In markets memes can be desired because they give access to other memes - for example, because they may be exchanged for them. If a type of meme emerges which can easily be exchanged for many other types, it can become desirable independently of any use of its own, emancipating its desirability, so to say, from particular desires. Money is such a meme, and the recomposition of desire as avarice, and the sense of the magical quality of money which humans now acquire in childhood, is one of the great accomplishments of memetic evolution. Intention becomes separated from all specific desires and all particular desirers; replicating its value, it passes, endlessly, from hand to hand. The emergence of money memes has had profound effects on the overall ecology of memes, forcing other memes to evolve effective symbioses with money, or face extinction.

(Some parenthetical questions: why do some variations of money succeed over others? Why does bad money drive out good? Essentially because simulating or diluting the currency are variations which shift the balance of probabilities as between the two things that holders of money may do with it: hold on to it, or spend it. The paradox is real partly because it is

apparent. {Footnote 3} One of the reasons why you can't take money with you when you die (if you try, it ceases to be money) may be that this destroys its exchangeability - put more generally, its intentionality. I return to the question of intentionality later. {Footnote 4}

The method of reproducible experiment in science can be viewed as another pervasive meme, in some ways like money. Very occasionally it gets subjected to analogous tests: Is it falsifiable? Or forgeable? Can it be replicated? It circulates tirelessly among expanding numbers of human individuals and institutions, insinuating itself into mental life wherever it can. Where it does it transforms our methods of thinking, and with it the practical world we create. {Footnote 5}

Why, then, does bad science not drive out good? Essentially because - like the spider's web, or the virus' sneeze {Footnote 6} - it has (to a degree) successfully imposed its own organisation of development, its own replicative history, on part of its environment, as "norms". Science's price for economic and intellectual freedom is - among other things - the capture or exclusion of other memes. Similar relations hold for other normalising memes, and their obverses: deviance, disability etc.)

In invoking money, or science, we also implied small galaxies of other memes (property/theft, ownership, contract, price, etc, in the first case; and reason/mysticism, evidence, law, order, etc, in the second). This is quite natural, as memes, like genes, go around in gangs. {Footnote 7} Indirect palatability can take other forms. Human beings also compete, and a meme's unpalatability to one person can be the basis of its desirability to another - rather as an insect repellent or an antibiotic is a meme designed to be objectionable or lethal to other species. An eviction order, like a bullet, is a type of meme which proliferates by making itself desirable to (some, but not all) human beings.

Certain memes, however, seem to be irreducibly unpalatable. Perhaps, as sociobiology occasionally seems to suggest, it is negative prescriptions - "Thou shalt not commit adultery" - which are most like this. This sort of meme has various options open to it to increase its circulation. It can link itself as tightly as possible ("Till death do us part") to other memes that *are* highly palatable - parental enjoyment, domestic bliss, home cooking, and so on. Marriage, incorporating and preserving its prohibitions and irritations, is such a meme-amalgam - analogous to the fruit that disperses many types of seed.

For some unpalatable memes, though, the pill simply won't carry enough sugar. At this point a meme's adaptation may need to work more interior changes. It may, for example, try to persuade carriers by distinctively human mechanisms: not that it would, on balance, be *pleasant* to ingest it, but that they *ought* to do so. Law, morality, political correctness and good taste are meme-ingredients of this sort. Various subsidiary means have developed to encourage these. Some appeal to human beings' wish simultaneously to conform to and maintain status within the group: socialisation, and its modern philosophical expression, universalisability. Others reinforce themselves with a moral carrot - eternal salvation - or with a stick conscience, hellfire, etc.

Yet linking a meme to its "should" need not be a grim business. Fashion, as much as morality, is fuelled by others and their expectations. Many delights of flirtation - and of social reproduction in general - are "frequency dependent" - accentuated by being generally relished: *Cosi fan tutti*.

{Footnote 8} And many memes proliferate by causing their human vehicles to internalise purposes beyond themselves and their kin: in a word, to embrace meaning in life. {Footnote 9} Considering meme adaptations will also lead me to ask - though not to resolve - some more opaque questions. Let me mention a few of them here:

Memes have numerous adaptations to their symbioses with humans. How have humans adapted to their memetic environments?

Why, even in solitude, do we feel shame, or joy, or laughter? What drives the evolution of our species' taste for meaning in life?

Why are we caught in the momenta of moods which seem to arise outside and despite ourselves?

Is the relative speed of meme evolution changing? In the past generations upon generations of our ancestors changed so gradually that it was imperceptible. Yet today teenagers inhabit a world unrecognizable to their parents. What has accelerated?

As this Cook's tour of culture suggests, Dawkins' meme promises (or threatens) to apply to a rich variety of cultural phenomena. It may indeed do this, but only by behaving as DNA-based life has done in colonising the biosphere - by evolving very diverse forms suitable for particular niches and problems.

## Facets of memes: an overview

Like biologists, our task is thus to make some sort of preliminary classification of memes' features - only now slightly more detailed than the distinction between parasitism and reproduction with which I began. I have found it useful to distinguish 12 principal facets, to each of which the general principles of natural selection apply. What follows is, of course, intended as a shopping list, not as dogma.

Infectiousness. Unlike genes, memes can pass not only from parent to offspring, but between unrelated humans. They are like "infections" or "viruses" which proliferate by jumping from one human mind to the next. Good examples of this sort are fashions in dress, crazes that sweep school playgrounds or jokes that make the rounds. The crucial adaptations are transmissibility and memorability. {Footnote 10}

Teaming. For this characteristic it is no longer adequate to treat memes as leaping or infecting from individual humans to other individual human "atoms". The "meccano" aspect of memes links individual humans together in shared institutions and purposes. A football team is a good example of this aspect. This sort of meme links a number of individual humans together for common purposes. This happens through bonding mechanisms - a little as atoms link together to form molecules, and as certain types of molecules link to form macro-molecules capable of self replication.

Bonds. But there is a crucial difference. Human "atoms" link to each other via intentions and expectations. In football, each player wears a distinctive shirt to make clear which team he

belongs to. The shirt is a visible shorthand for his intentions - which goal he is trying to kick the ball into. In football teams and other memes of this sort individual humans act in concert by treating each other types of as tools or dolls - but as walking, talking, hoping, planning, feeling dolls.

Feelings (team spirit, in this case) are what allow human dolls to link to other dolls, and together manipulate the world. We do not need uniforms to recognise each others' feelings and act through them, but we do need some links. Uniforms depend on feelings at least as much as feelings depend on uniforms.

Immortality. This feature also subordinates individual humans to memes, but across time rather than space. Memes can endure across many human generations. A church provides a good example. So does a language or a state. So - to take more modest instances - does a song, or an accent. Rather as a candle flame, or an organism, can remain the same while most or all the atoms which make it up are changed, such memes endure historically through the replacement of the human atoms which compose them. A football team largely lacks this characteristic, since it is assembled for a period and for purposes shorter than the typical human life-span. But a football club has it - changing its players from season to season and its fans from generation to generation. For many people their football club, indeed, binds them more tightly than their church. I term this the immortality or eternal flame aspect of memes.

Memes within memes. Memes are constituted not just of "raw" humans, but also of other memes. Memes ingest, parasitise, inhabit, invade, include and conjugate with other memes. If we examine either the external or the internal environment - the anatomy - of a meme, much of what we see consists of other memes. (Though the distinction between inside and outside is an even more difficult one for memes than it is for organisms. {Footnote 11})

Values. Memes link humans by means of values as well as feelings. The members of the football team (and the club) are integrated by the common value they place on kicking the ball between the other side's goal posts - and, through and beyond that, on getting the particular vehicle for doing this nearer to the head of the league table. Values may also include commandments, rules, ideals, orderings and faded dreams. Change key values and you may destroy the meme. We can imagine a Martian studying our football match. S/he/it might easily draw the conclusion that if twenty-two grown persons wanted to propel a leather sphere as often as possible through a wooden hole, they should do so in concert. That would, however, be failing to see the memes for the people. Like memes, values derive from other values.

Meme vehicles. Successful memes persist across human generations by assembling into complexes which are capable of self-reproduction: societies and cultures. Societies are themselves memes, which incorporate other memes and memetically "inscribed" human beings and many other types of artifact. To endure, societies must reproduce values and feelings, as well as human beings. The meme vehicle, taken as a whole, must reproduce itself, including its atoms and bonds.

Selective extinction. Many of humankind's earlier cultures are extinct. This is not only because they originated a long time ago, but also because meme selection and evolution confer relative

advantage on (some) of the more elaborate variants. Similarly, the primitive nucleic acids through which life on earth first developed have long been displaced by more complex descendants - cells, organisms and species. The evolution of culture resembles that of life in that it confers advantages on (some) latecomers. Culture, like life, expresses combined and uneven development. {Footnote 12}

Scratchpad effect. A related point in studying memes, therefore, is that their evolution involves erasures and extinctions, as well as additions (new speciations). The cultural scratchpad incorporates traces of many vanished memes. As in biology, reconstructing history, much of which is invisible to our present eyes, is an essential part of understanding. History - change over time - greatly complicates classification, but it is essential to explanation.

Genotype/phenotype. Memes lack any single physical replication mechanism such as carbon-based life makes use of in the DNA helix. Consequently memes lack organic life's more clear-cut distinction between genotype and phenotype.

The "culturological principle". Any explanation or definition is itself a meme, with its own history and evolution. Earlier, less adapted, forms of biological explanation, for example, try to abstract from the problems of the new arising in time - but have, consequently, more limited success. Modern evolutionary biology arises from centuries of memetic evolution, from creation myths, through efforts of classification, to the recognition that no life form is final. The same principle applies to all memes, including all explanatory ones. I call this memetic principle the "culturological principle" (in allusion to the "cosmological principle" by which astrophysicists remind themselves never to assume that the Earth's position in the universe is a privileged one).

Humanism. Our view of memes is coloured by certain characteristic illusions or necessary appearances, analogous to creationism or vitalism in our thinking about organic life. One is the conviction that memes get their purposes from entities other than themselves, even from us. Another (in a sense the obverse) is that memes arise solely from other memes (or that values cannot derive from facts, or culture cannot derive from nature). It is possible - indeed common - for a human individual to entertain both these memes at once.

These features of memes are neither exhaustive nor exclusive. They do not, even together, define memes. And we cannot draw wholly watertight distinctions between them. For example the borderline between feelings and values is extremely fuzzy. I have nonetheless found it helpful, in trying to think about memes, to separate out these 13 points, and I refer back to them in the following chapters. This is why I have given them shorthand names.

## Culture and psychology

"Culture is the passion for sweetness and light, and (what is more) the passion for making them prevail" - Matthew Arnold, Literature and Dogma

What sorts of problems might the meme meme help with? One, I think, is the apparent autonomy of culture.

The transmission of culture means, among other things, its passage from one human mind to another. A cultural trait is one that passes reasonably successfully, in recognizable form, from mind to mind, and is to that extent independent of particular minds and psychologies. A standpoint prominent in the human sciences at least since Marx and Durkheim is that of the autonomy - relative, or less so - of culture. It has been greatly encouraged by mass manufacture and mass education.

On this view the psyches of new born human infants are highly malleable - blank slates, or "general purpose" software - ready for the learning which their cultures of upbringing subsequently imprint upon them. A concomitant principle of this independence of cultural "software" from neuro-psychological "wetware" is an emphatic anti-psychologism, such as that which Durkheim expressed:

"The determining cause of a social fact should be sought among the social facts preceding it and not among the states of individual consciousness."

However, over the last few years ideas from human biology, in the form of "evolutionary psychology", have laid siege to views of human psyches as blank sheets of paper on which cultures autonomously evolve their messages. "Evolutionary psychologists" take aim at social science's assertion, or assumption, of cultural autonomy. They argue it is wrong to suppose that social and cultural regularities are all *sui generis*, not dependent on human beings' evolved biology. The life sciences are now starting to explain many important components of human psychology as adaptations of the hunter-gatherer life lived by human beings until the relatively recent evolutionary past (up to about 10-12,000 years ago, when agriculture began to take root in the Near East). Human beings' asymmetries of sexual preferences; our patterns of male jealousy and of female adultery; the propensity of young men to form aggressive coalitions for war; children's play fighting; attachment (and grief) between careers and children; our sex differences of spatial perception; our colour categories; our body language; our tastes for salt, sugar, fatty foods and open landscapes - all these (argue "evolutionary psychologists") may in principle be understood as universally human, "species typical" adaptations of hunter-gatherer life in a largely untouched nature. {Footnote 13}

The programme of "evolutionary psychology" raises problems of its own. {Footnote 14} Can we separate out genetically transmitted adaptations from culturally transmitted ones? Can we combine genetic and cultural transmission within a single ("co-evolutionary") framework? I discuss such problems below in the context of the sociobiologists' bête noir, the "Standard Social Science Model".

## Cultural diversity and biodiversity

A related problem is that of cultural diversity. Nowadays our world is very different from the natural surroundings of our hunter-gatherer ancestors. Almost everything that urban humans see has been made or modified by other humans. Consider tangible artefacts alone. Many of these are mass-produced objects, but their variety of types rivals that generated by organic evolution. The contemporary individual, it is estimated, may encounter (and need to discriminate) 20-30,000 different specialised objects of common use during her or his lifetime (Petroski, 1993,

23). {Footnote 15} If our active vocabulary has greatly contracted, compared to our hunter-gatherer forbearers, with respect to living species (see, e.g., Levi-Strauss,), it has much expanded with respect to made objects. {Footnote 16} Indeed Basalla estimated that the number of distinguishable technical inventions now patented exceeds by a factor of two or three the number of different flora and fauna species identified by biologists.

How does technodiversity come to rival biodiversity? Knowledge of the basic machinery regulating evolution among genes has transformed biology. We now think in terms of the chemical mechanisms of genetic replication, of the discrete nature of inheritance, of the sources of variability and selection, of specialised reproductive organs and the isolation of germ lines, of speciation, sexual recombination, and kin selection. Such concepts are fundamental to our understanding of the diversity of life-forms based on DNA-RNA replication. And they form the groundwork which evolutionary psychologists bring to bear on humans' universal, "species-typical", psychological architecture.

The history of biological classification suggests some of the difficulties. Early classifications tended to distinguish species pragmatically - as, for example, in creation myths, or in "herbals" which grouped plants according to their medicinal uses. {Footnote 17} Only gradually did biological thought come to distinguish species in the modern manner (Aristotle, for example, did not always view them as reproductively separated) or to see their distinctive organs in terms of their functions for the organisms themselves. {Footnote 18} From there it was a further step to the Darwin-Wallace view of the evolution of species, and of the biological present as a "still" from a much longer historical film. Darwin developed his theory with almost nothing of our present systematic knowledge of inheritance {Footnote 19} or of the molecular patterns which transmit genes. And there remain arguments among biologists on the relative uses of classifications by adaptations, by evolutionary descent or by reproductive separation. {Footnote 20}

By placing memes at centre stage our understanding of culture can gain additional help from biology. Genes are difficult and arbitrary of definition, but that does not make them less real or less fundamental to our understanding of biological evolution. I suggest that we allow that memes are similarly difficult to define, but nonetheless treat them as real. Studying the natural history of memes and their selection may even help nature/nurture debates to yield more fruit.

Several approaches have already been developed, though they have attacked the matter from different angles and used a variety of terminologies. Population biologists have refocused their tools on learned or cultural traits ("memes" or "culturegens") and have developed formal models of the incidence and spread of learned behaviours in communicating populations of organisms. The transmission of learned behaviour has been studied in a variety of animal as well as human populations (Bonner 1980, Delius 1991). Models have been developed of cultural and genetic "co evolution" or "dual inheritance". {Footnote 21}

There is a conceptual difference between a form of learned behaviour as a cultural trait, replicated by transmission from human individual to human individual, and the part which the behaviour plays in the replicative cycles of a cultural complex. Singing *Silent Night* is behaviour learned by multitudes of small children; the song's collective rendition is a commonly-

encountered element in Christian congregations' rituals of winter renewal. The direct replication of memes from person to person is part - but only part - of the processes through which institutions persist across generations.

[WHAT is the meme, virus in this example?]

[Historical studies of technological and cultural change have often treated it as broadly analogous to histories of biological organisms. {Footnote 22} And crucial qualitative distinctions have been drawn between different *types* of meme, and the relationships between their evolutionary characteristics, and those of the humans who carry them. {Footnote 23} {Footnote 24}]

What sorts of phenomena might an evolutionary ecology of memes help us to understand?  
Is the social superorganic: a meme's-eye view?

In my opinion the most interesting is the central conundrum in response to which ideas of the autonomy of culture have spread in the social sciences: the compulsiveness of the cultural and social. Cultural beliefs and practices have a variable and often arbitrary character. They combine this with their compelling, obligatory, "instinctive" action through and on individual human beings. "Our" institutions precede us and prevent us choosing. Students of society thus often posit the social as something independent of and - at least logically - prior to the individual, and reject ideas of already-formed human animals. Indeed many social scientists err in the opposite direction, conceptualizing the cultural and cultures as things developing independently of individuals, and reifying culture alone as the propellant of individual action. This is a view which, in its pure form, "evolutionary psychologists" regard as fantasy. {Footnote 25}

Suppose, however, we set aside for a moment our questions about how the social and cultural drive individual humans in the ways that they do. Let us also set aside our even larger questions about what memes essentially are, or if they really exist at all. Suppose, instead, we ask questions of the following types:

- What sorts of adaptations might memes incorporate in the course of evolution in order better to equip themselves to proliferate among individual humans?; and
- Why may this or that characteristic of a meme have been selected as an adaptation to its (then) environment?

These are two distinct questions, since much of the environment of many memes consists of other memes.

Rather than puzzling over the problem of how memes drive individual people, let us step back a little and concentrate on the distinct question: what drives meme evolution - the selection of traits to become more frequent *among memes*?

In borrowing such concepts as *adaptation*, together with connected ideas, such as *exaptation* from evolutionary biology I do not mean that they are easy of definition there. (An exaptation is, roughly speaking, an adaptation which, over subsequent evolutionary time, has assumed a

different function from that which initially gave rise to selection pressure in its favour. An example is the ear-bones of mammals, which began as the gill-structures of our remote aquatic ancestors. Exaptations are rather important among memes.) The idea of an adaptation is, however, so central to understanding why organisms (and memes) are as they are, that it is unavoidable. {Footnote 26}

## Memes' and genes' environments

What is the environment of a particular meme? Or, to borrow a further idea from evolutionary biology, what formed its "environment of evolutionary adaptedness" - that is, the (historical) environment in which the meme's (or organism's) characteristic adaptations arose and became incorporated. The environment of evolutionary adaptedness may be very dissimilar from the present environment. A central contention of "evolutionary psychology", for example, is that the environment of evolutionary adaptedness which has shaped the "species-typical" psychology of today's humans is not today's environment, but the environment in which our hunter-gatherer ancestors lived before the agricultural revolution. {Footnote 27}

To anticipate by a little the answer to the corresponding question about memes' environments of evolutionary adaptedness: for most of the memes we are familiar with not only do (or did) their EEAs consist largely of other memes, but those other memes were themselves already adapted through memetic evolution in highly complex ways.

One - but only one - of the ways in which a meme's adaptations can increase its meme's success is by helping its incidence in a human population. The compulsive - though arbitrary - power of culture has much to do with this: a meme tends to succeed among its coexisting and competing memes if, for whatever reason, it can predominate in individual humans and among human populations. In this indirect sense memes are democrats; in order to get selected they must "try" to find ways of getting and staying in human beings. In similar fashion, genes "try" to get incorporated in species (or "gene pools").

## Selection, invention and intention

There is an important distinction between the environments of evolutionary adaptedness of biological organisms, and those of memes. Biological adaptations arise, are selected for, and become embodied in the organism because they are successful for the organism in its current environment (though they may then endure long after the environment of adaptedness changes - and this includes the other adaptations gathered in the same organism). Natural selection is "blind"; it "tries out" solutions, selecting or discarding them, just on those problems which it can feel under its hands at the moment. The functional design of organisms is not the result of any supraorganic intention. Insofar as we may speak of organisms "trying" out variations, this is a convenient, but metaphorical, shorthand.

With memes, however, it is otherwise. They do frequently adapt to the future as well as to the present and past. Successful memes are (almost by definition) those that adapt most successfully to their actual future. Memes are intricately bound up with intentions, in a way that genes are not. My introduction of intention is a small but significant mutation of Dawkins' meme-concept.

Dawkins conceives of memes as consisting essentially of similar copies in living human brains, and as replicating through imitation. I broaden the matter slightly: one - but only one - of the forms intention can assume is imitation.

An example or two may help illustrate the difference. The meme threatening hell fire, as Dawkins points out

"...is highly effective. It might almost have been planned deliberately by a Machiavellian priesthood trained in deep psychological indoctrination techniques. However I doubt if the priests were that clever. Much more probably, unconscious memes have ensured their own survival by virtue of those same qualities of pseudo-ruthlessness which genes display" (Dawkins 1977, 212).

There are two distinct questions about memes and consciousness here. The first is: Are (human) consciousness and intentions integral to memes' propagation? And the answer, for many memes, is Yes. The second question is: are memes conscious? Dawkins implies that the answer to this question is No. But for some types of meme, at least, the answer may be a complicated sort of Yes.

Take on of the more complex memes in which the hellfire and other memes have been incorporated - the Protestant reformation. When Luther composed his theses attacking the Pope's marketing of indulgences to finance St Peter's, and nailed them on the door of the church at Wittenberg, he was intentionally propagating one complex of memes (armed with people - people who were, in turn, to be armed with Luther's new-forged memes) to compete with another: setting, so to say, one family (or coalition) of memes to catch another. The Roman church had encouraged the commercialization of indulgences as a solution to certain future problems posed by prior memes: Pope Leo X's extravagance and his need for funds to complete St Peter's; and, on the part of the purchasers of indulgences, the dread of hellfire and the intention (or at least hope) of avoiding it in exchange for cash. Fusing the attractions of salvation and economy, Luther's recombined complex of doctrinal propositions spread through Germany "in a fortnight". {Footnote 28}

Notice that the memes involved varied along a sort of spectrum from rather sophisticated ones, that had evolved among and in relation to other memes which themselves already had an elaborate evolutionary history - like Luther's denial that the Papacy could delegate the remission of sins to professional pardoners - to relatively "primitive" ones - like the fear of pain or fire. We may think of parallels, perhaps, with the increasingly complex chemical history of biological evolution, or of analogies with higher organisms, many of which have evolved wholly within elaborate cycles of adaptation and dependency on each other - for nutrients, for shelter and support, for reproduction, and so on.

## The evolution of intentionality

Intentionality is involved in generating scientific as well as religious memes. Here is how another biologist, Wolpert (1992, xii), sees scientific innovation, contemplating both his own experience and the overall evolution of science:

"Since science is unique, it is to be expected that scientific creativity has its own special characteristics quite different from those of the arts...Scientific genius is often characterized by a "psychic courage" which requires scientists to include in their ideas assumptions for which they have very little evidence."

Perhaps significantly, Wolpert sees natural science as arising out of the lineage of Western Christianity, with its penchant for logic among a paucity or confusion of evidence {Footnote 29}.

He sums up the problem thus:

"The puzzle lies in how scientists decide which experimental data or which theoretical construct they are willing to give up when these are in conflict" (Wolpert 1993, 116)

The decision what to cling to and what to abandon is not peculiar to science. However, it arises for memes in a way that it does not for genes.

Memes raise a fascinating but difficult question, with roots in the histories of biological reactivity and animal behaviour: that of the emergence of intentionality, self-consciousness, and rational thinking as a function of organic life. {Footnote 30} This is a question I skirt round.

More generally, I steer as clear as possible of general questions about ultimate or essential sources of change in memes. At this stage of our knowledge a piecemeal "natural history" approach, describing memes and some of their adaptations and relationships, seems to me the more useful.

But turning away from theorizing about origins and toward investigating the historical details of particular adaptations, does not mean escaping from the problems of intention. Human history involves the continual inception of new intentions in at least two senses:

- Old humans age and die and are replaced by younger humans who absorb similar patterns of intention; broad, habitual, patterns of intention are (approximately) replicated from generation to generation;
- Among these habitual patterns new types of intention arise, and some of these proliferate powerfully.

Any overview of memes and the patterns of intention linked with them require us to consider types of intention that only arise in fairly recent, historical times. However there may be a compensatory advantage to this: we have fuller and more direct evidence on intentions for the more recent past. Provided we avoid getting hung up on questions of essential or ultimate origins, and concentrate on particular adaptations, we may take advantage of the fact that evidence on the historical past is very rich compared with that on Pleistocene or earlier times. Since then there have arisen many psychological processes and practical techniques that were not around earlier.

Can thinking about memes help us understand them? Can memes and their evolution be studied as phenomena distinct from DNA-based evolution - a little as our understanding of organic life is distinguishable from (though consistent with) physics and chemistry?

We need to take account of intentions. But we must avoid getting bogged down in questions about what intention (or its associated notions: meaning, consciousness, free will, etc) essentially is (or, for that matter, biologically was). Instead I take a fairly matter-of-fact view of intentions and meanings, and ask more prosaic and immediate questions about how they are transmitted, preserved, disguised, enforced and so on. One point I underline is that intentions need not be either individual, mental or subjective. They take effect through minds, but to pass from one mind to another they frequently take extra-somatic forms: for example, in an artefact made by one subject for use or consumption by others. The physical phase of such a meme embodies intention, not only when it explicitly expresses intention (as in a No Entry sign, or an "If ... then ... do ..." instruction in a computer programme), but when its impersonal purpose is implicit in its nature (as in a lollipop, or an airliner)

One advantage of intention is that it enriches the evidential record, at least compared with that left by genes and organisms, whose remains survive mainly in fossils. The primary reason for this is that many memes are made to last - by a combination of invention and selection. This applies to intangible as well as tangible memes; the song survives the singer, though the temple may crumble. Organisms, on the other hand, are designed to reproduce, senesce, and die. Very few individuals provide lasting remains; only their genes are immortal, recombined in their descendants. Whatever the difficulties of interpreting it, the trail of historical evidence is of great importance. As far as carbon-based life forms are concerned, an estimated 99% of the species that have lived to date are now extinct (though, granted, classifying memes among "species" is even more difficult). {Footnote 31}

## Limits on change

That history is such an important source of evidence does not contradict the fact that it acts on the present only indirectly, through the structures it has already produced. {Footnote 32} Evolution "tinkers"; it must start from the functional equipment already in existence. We may sometimes feel, as the traveler who got lost in Yorkshire was told: "If tha' really wants to get to Bradford, tha'd best not start from here". But the fact is that all evolution, biological and cultural, starts from where it's presently at, and all adaptation is modification of, and to that extent incorporates, previously existing forms.

In culture, as in biology, ontogeny remains rooted in phylogeny - though not in the same way. For languages based on the Latin alphabet, for example, the QWERTY keyboard, though suboptimal in several respects, has become ontogenetically incorporated in typewriter and microcomputer design. Efforts to step round or short-circuit its irrationality run into memetic obstacles embodied in humans' enculturation and training. {Footnote 33} For memes as for genes, co-evolution - interdependent natural selection of multiple entities - leads seamlessly into irrationality, extravagance and invention. {Footnote 34}

In biological structures variations are more likely to be viable - that is, not lethal to the organism or its offspring - the later they come in the organism's development. {Footnote 35} What the genetic code of an organism specifies is not a depiction of the organism at any future stage of its development, but a developmental programme for the organism's life cycle (including the all-important and oft-repeated "loop" specifying reproduction of the organism's genetic code). The chances of a random change being made early in the program and leaving subsequent steps viable or functional are less than for changes later in the developmental sequence. If we think of one of those many storied buildings one can construct with dominoes on a hard table, it is clear that a very slight disturbance of a domino nearer to the base is more likely to bring the whole thing tumbling down than a similar disturbance to one of the upper dominoes. {Footnote 36}

With memes, however, intentional redesign substitutes, in part, for chance variation. The possibility thus arises (or, to be exact, becomes vastly more probable) of substantial alterations to the base of the structure: substituting dominoes near the bottom, replacing some of them with completely different types of building blocks, or even removing superfluous ones altogether. Consider the following account of meme adaptation. It is Lorenzo da Ponte, explaining how he turned Beaumarchais' *The Marriage of Figaro* into the libretto for Mozart's opera:

"The duration prescribed as being usual for dramatic performances, a certain number of characters generally introduced into the same, and some other prudent considerations and exigencies imposed by morality, place and spectators, were the reasons why I did not make a translation of this excellent comedy, but rather an adaptation or, let us say, an extract.

"To this end I was obliged to reduce the sixteen characters of which it consists to eleven, two of which may be performed by a single person, and to omit, apart from an entire act, many a very charming scene and a number of good jests and sallies with which it is strewn, in place of which I had to substitute canzonettas, arias, choruses and other forms, and words susceptible to music, things which can be supplied only by verse, but never by prose." {Footnote 37}

Human intentions allow meme adaptation to "cut through" or "streamline" existing structures, finding alternative routes to existing or imagined outcomes, in a way that biological redesign cannot. {Footnote 38} {Footnote 39} They also encourage *exaptations* - the turning of existing adaptations to new uses. Indeed, a great deal of cultural evolution consists of inventing new ends which existing means can be converted to propel. {Footnote 40}

## **Invention and necessity**

When they embody intentions, memes in some degree detach them from individual human agents. In one famous passage of his *Decline and Fall*, Edward Gibbon asks about the adaptations to human intentions that made Christianity so successful:

"Our curiosity is naturally prompted to inquire by what means the Christian faith obtained so remarkable a victory over the established religions of the earth. To this inquiry an obvious but satisfactory answer may be returned: that it was owing to the convincing evidence of the doctrine itself, and to the ruling providence of its great Author. But as truth and reason seldom find so favourable a reception in the world, and as the wisdom of Providence frequently condescends to use the passions of the human heart, and the general circumstances of mankind, as instruments to

execute its purpose, we may still be permitted, though with becoming submission, to ask, not indeed what were the first, but what were the secondary causes of the rapid growth of the Christian church?" (Gibbon, XV)

The "secondary causes" of any successful ideology includes myriad of individual human intentions, coherently oriented. Intentions may be embodied in very specific and (especially where expressed in symbolic systems) apparently minor ways.

A further theological meme - that of the Trinity - furnishes an example of this - and of intention's streamlining effects.

Most of today's Christian memes of God descend from the doctrine of the Trinity adopted by the Council of Nicea (325), which proliferated as orthodoxy in the mystery at the heart of the Nicene Creed. How was the tripartite God to be defined in a way which did justice to its family origins, its administrative functions, and the divinity of Christ? The theologians settled upon an ambiguous but precise formula: a single God existing in three persons and of one substance, the Son being perpetually generated by the Father but simultaneously co-eternal with Him.

But they did so only after years of disputation, in which arguments and counter-arguments sprouted vigorously, producing doctrinal tangles which included most of the imaginable ways in which the components of the Trinity could be put together.

Thus one of the key decisions of Constantine's Council at Nicea was procedural rather than substantive: to reduce the thickets of previous controversy, and consider just two alternative Greek terms, differing by a single letter, for describing the identity or resemblance between Father and Son: either St Athanasius' assertive *homoousion* ("of one substance") or the more cautious *homoiousion* ("of like substance") favoured by the supporters of Arius of Alexandria.

The single *iota* separating them marked a distinction (as Gibbon faintly exaggerates) "invisible to the nicest theological eye". More visibly, it was expressed in the formidable array of anathemas, against Arianism and other heresies, behind which the Nicene theologians sought to protect their preferred formula. The concentration of intention - in this case organizational and political, as well as theological, shaped formulaic choice and heightened doctrinal creativity. The decision that the doctrinal edifice as a whole should balance on the inclusion or deletion of a single letter was recognition of the many (and often poorly-understood) alternative combinations of formulae that had been advanced and found wanting, many of them before the disputants were born. Barkow (1989, p250) writes of meme variation:

"Some categories or subcategories of cultural information do seem to lend themselves to a "meme" or "culturegen" approach; for example, whether I brush my teeth with a rotary motion or up-and-down and side-to-side. Others, such as my perception of the nature of the Holy Trinity, do not seem to lend themselves to particulate approaches."

This needs qualifying. What may be susceptible to "particulate approaches" are particular variations. These can appear quite local within the overall meme-complex they inhabit, but have very general, and partly intended, consequences for the larger meme.

The point is that apparently trivial symbolic changes can contain large intentions with profound repercussions. When the iota was reintroduced "The whole world groaned and marveled to find itself Arian". The Arians succeeded partly by increasing the infectiousness of their meme, setting their theological ideas to music as popular songs. In the 320's common folk from Alexandria to Constantinople could be heard singing ditties to the uniqueness of God the Father. {Footnote 41})]

God memes have one distinctive and interesting feature. While they evolve, the object to which they refer, being eternal and unchanging, does not. This furnishes us with a sort of rudimentary experimental control. When God-memes undergo rapid change, we can reasonably attribute this to the memes' circumstances. In the context of our Arian example, the candidates needed to gain the assent of the senior ecclesiastics promulgating doctrine - a requirement that reduced the variants to two, rather simply related, contenders. {Footnote 42}

Selection among biological structures is absolutely "blind", genetic variation arising wholly by chance, while memetic adaptation is only relatively so. The fact that cultural adaptation has intended as well as unintended consequences focuses variation, injecting an element of conscious design and accelerating the incorporation of successful features and the attenuation of unsuccessful ones. The distinction between genetic and cultural evolution is only relative. They combine in domestication, the foundation of much human culture, and, more recently, in the possibilities opened up by direct examination and manipulation of our and other species' genetic material.

The idea that acquired characteristics cannot be inherited is rather basic to biology - now almost one of its defining characteristics. But what phenomena the principle applies to, and in what ways, remains an empirical question. A large part of the study of culture involves investigating how acquired characters are transmitted - and how transmissibility gets acquired or invented.

It is certainly possible, as Delius (1991) argues, that the apparent ability of memes to transmit changes acquired within individual brains (and its corollary ideas, of intention, creativity etc) is an illusion, arising from the very limited extent to which we understand meme selection within the individual's memory and its underlying neural structures. Thus what we call "intention" is in fact a concealed process of random variation and selection within a long-lived host of memes with a much shorter lifespan, analogous (say) to intra-host evolution of micro-organisms' antibiotic resistance. Edelman's (1992) "theory of neuronal group selection" - which draws on our understanding of how immune systems learn - has similarities with this view. {Footnote 43}

I remain agnostic, concentrating on processes of meme transmission between humans, and continuing with the commonsense, or "folk", view of meme change within individuals as (at least partly) "intentional".

## **Group selection?**

For gene-centered biology the vehicles for genes, and therefore the "candidates" for selection among them, are individual organisms, not species. Traits such as genetically transmitted altruistic behaviour (alarm calls at predators) or altruistic organs (mammary glands) or even

whole organisms (sterile worker bees) - which benefit other individual members of the same species can therefore evolve only where the individuals benefitted share genes with the individuals expressing the trait: Hamilton's (? 1964) kin selection, or inclusive fitness. {Footnote 44} This modern, gene-centered view has largely replaced earlier habits of thought which supposed that traits could be selected "for the good of the species" {Footnote 45}. More generally, the appropriate units of selection remain difficult and debated matters in evolutionary biology.

But human beings are sometimes nice to some of those who are not their kin. And some are nasty to those who are close kin. What is more, some people would like to be nicer than they generally are. And prisons, psychiatric hospitals and committees contain yet others who would be even nastier than they are, if they were allowed.

This is where an advantage of treating memes as real comes in. Memes can evolve (and many of them have evolved) powerful and sophisticated means for coercing, motivating, persuading and repelling human beings - fruit, saliva and teeth, so to speak. True, we are still a long way from understanding the "biochemistry" through which these adaptations work. {Footnote 46} But there is less doubt that they exist. Organs - such as patriotism or prisons - that would require group selection to evolve among individual humans may, nonetheless, evolve from variations among memes.

One common characteristic through which natural selection between memes takes place (though it is certainly not the only one) is differences between memes in the numbers and qualities of human beings they succeed in "recruiting" or "ingesting" - a very rough analogy, perhaps, to differences between individual animals of the same species in foraging or predation efficiency. Patriotism (regarded as a virtue, even by the other side) emerges from a long process of competition among loyalties on a variety of territorial, linguistic, etc bases, where the size of groups competing for individuals' loyalties ranges from the family to humanity as a whole. Prisons remind us that recruiting or ingesting support is often indirect. Prison restrains and punishes those it incarcerates, but it *succeeds* by being a morally appealing device for the politically powerful. Its nutrition comes both from those incarcerated and, more importantly, from the values and intentions of the population at large.

These examples contain a more general point: the compulsions of our social world may become more intelligible once we get into the habit of seeing them as efficient meme adaptations, rather than as final causes of individuals' behaviour. {Footnote 47} Naturally, there are many things, such as love, or war, which we can get to grips with only by combining accounts of memetic and genetic adaptation. {Footnote 48} Many memes develop a sort of "territoriality" - personal, not spatial - whereby they include or repel humans according (chiefly) to their memetic characteristics. (The quotation marks are essential, as notions of species, individual organisms, and hence territoriality, do not transfer easily from genes to memes.)

Applied biology - antibiotic drugs, for example - raises the question of who is being domesticated by whom: humans by bacteria, or bacteria by humans? Similar questions arise with humans and memes. Human populations, and their culturally and morally differentiated groups, act as both prey and predators of memes. From our point of view some moral memes - for

example - are frighteningly supra-organic, unpredictable and often rapacious. But from their point of view, memes are just getting on with the everyday business of life. And some memes - some formal organizations, for example - engage in elaborate intercourse with each other. Like you and me at breakfast time, such memes are not concerned with the feelings - still less the intentions - of their sausages, or the atoms (or humans) of which they themselves are composed. They are too busy telling each other their plans for the day, or last night's dreams (and, perhaps, persuading themselves that their dream-work was perfectly proper). We can get inside their skins only by regarding them as being as real as organisms, or persons.

## Meta-memes

The assemblages in which memes unite for survival and propagation do not necessarily have physical or even biological foundations. Memes can hang around together for more general reasons, and in more abstract ways. In many cases this is because they have gotten used to each other in the past - as in the specialisms of academic knowledge or the division of labour more generally. But new and unexpected affiliations can also form. Like upwardly mobile football hooligans, memes seem to dissolve and reassemble their ranks for each season (indeed for each match, even each penalty). They form and reform as gaudily clad but often hostile crowds: orthodox versus heretic, U versus non-U, entailed versus excluded, true versus false, and so on and so forth.

Naturally, the patterns of meme-crowds are also memes, and get transmitted around as such. Myths, magic, religion, mathematics etc are (so to speak) "carnivore" memes whose life patterns reflect the general feeding, nesting or swarming habits of other forms. These more specialised meme machines - which, of course, may compete vigorously with each other - are reinforced by numerous forms of invention or recruitment.

Many memes go in for differentiated infection or domestication of human beings, forming a sort of "priesthood" of specially conditioned human types, and evolving markedly different morphologies within their life cycle for more effective spread - that is, for becoming endemic or epidemic in the human population. This point sounds abstract, but is actually quite familiar to the humans who specialise in the initial spread of the more abstract, "spore" memes: ideologists and intellectuals. As Plekhanov expressed it (in one of this century's more successful memes):

"Propaganda is providing many ideas to a few people. Agitation is moving many people by means of a few ideas. And agitation depends on propaganda." {Footnote 49}  
Similar infective and innoculative technologies are fundamental to education.

Memes may succeed by organizing their human disseminators in new and highly differentiated ways, rather as modern military forces employ most of their personnel on administrative, logistical and technical tasks, and only a minority on combat duties (or, perhaps, as parasites may inhabit different hosts at different phases of their life cycle). The military simile is apt, since one important function of such meme meta-machines is to contain and reduce the opposing meme meta-machines of rival meme-complexes. Rivalry is defined by reference to yet other - logical, ideological, and theological - meme complexes.

Here, as in many other meme contexts, truth or falsity are not necessarily the important things, and sometimes they are completely irrelevant. Adaptations are selected according to the maxim:

"My meme, right or wrong!" National anthems and advertising, the defense of orthodoxy against heresy, dictatorship and the doctrine of cabinet responsibility, a national curriculum and a guillotine, hypothesis and experiment, even the conventions of peer review - are various forms that memes' equipment against their competitors have evolved. I do not mean to suggest that truth is unimportant, merely that it is just one category of meme (or meta-meme) among many, surviving as best it can in a testing environment.

## **Other adaptations**

"What are you famous for?" "For nothing. I am just famous." - Iris Murdoch, Flight.

Memes do a wide range of things in their efforts to propagate. They rhyme, scan and get themselves set to music; some make themselves sweet, or sour, or both; some exaggerate their obnoxiousness or durability ("The word of God endureth for ever"); some become succinct or memorable, or simply turn up the volume; some parade gorgeous tails - or tales; some sprout wings and take off; some lumber after their mates, or pursue their prey, on thousands of footnotes.

One of the interesting and relatively recent things memes have done is try to emulate the more exact replicative techniques (and the built-in obsolescence) of biological life. This was first and most effectively done with tangible mass production - pins, cars, wedding rings, and so on - but similarly potent technologies are fast developing for moral and aesthetic memes. For both there is a trade-off between durability and cheapness - partly analogous to that between longevity and fecundity in organisms. {Footnote 50}

Many of these cheaper and more plentiful memes require the human beings who use them to be reshaped a little. Schools ensure that hearts and minds develop in accordance with proper templates. {Footnote 51} The development of standardized, quality-controlled production lines seems to require - for memes as much as for genes - the protection of the germ-line from the clamor and accidents which afflict the soma: insulating the design shop and the specialists who work there from the hubbub of the proletariat on the factory floor.

## **Memes and genes**

"Population genetics .... may be defined as that branch of epidemiology that deals with infectious elements transmitted exclusively from parent to offspring." - GC Williams, Natural Selection, p13.]

As meme selection proceeds, the main mechanism emphasised by human sociobiology and "evolutionary psychology" - the adaptive pressures of hunter gatherer life on genetic characteristics of the human psyche - never disappears. (It could do so only if memes switched species.) But as a mechanism of adaptation it acts so slowly and "blindly", it is so relatively inefficient, that its effects are increasingly smothered by more recent, meme-derived and meme-

dependent, mechanisms. Indeed, so relatively rapid is cultural transmission that even memes which deliberately and drastically lower individuals' inclusive fitness (contraceptive technologies, for example) can, when favourably enmeshed among other memes, saturate a human population in little more than a generation.

Ball (1984, 147) suggests a fourfold typology of memes, classifying them according to their adaptiveness for (a) themselves and (b) their carrier-organisms:

A symbiotic meme promotes behaviour that is adaptive for itself and adaptive also for its organism... ("Sex is fun; pass it on.")

A difficult meme promotes behaviour that is maladaptive for itself but adaptive for its organism... ("I know how to do it but can't explain it.")

A parasitic meme promotes behaviour that is adaptive for itself but maladaptive for its organism... ("Praise the Lord; send money; tell your friends.")

A bad meme promotes behaviour that is maladaptive both for itself and for its organism... ("Get away from it all; be a hermit.")

Meme selection tends to favour symbiotic/parasitic varieties, and Delius (1991) further disaggregates this category into its subdivisions in the biological study of symbiosis:

- Mutualism, from which both species gain fitness
- Commensalism, where the guest gains without cost to the host, and
- Parasitism, where the guest gains at the expense of the host.

In such terms a successfully disseminated contraceptive technology - or the thirst for upward social mobility which encourages it - is a "parasitic" meme: one which promotes behaviour that is adaptive for itself but maladaptive for its organism. Memes may vigorously recruit to the ranks of "their" organisms, even when the effect on the organism is to reduce its life-expectancy greatly (as in patriotic or kamikaze-style heroism).

Moreover, as the density and diversity of memes increases, forms can emerge which, though psychologically very powerful, have only slight direct connections with the psychological drives shaped before the Neolithic revolution. Similarly, while it is true that new species evolve to occupy vacant niches, it is equally true that most niches are opened by other species, which define them: the woodpecker presupposes the tree. (Develop on organism and environment.)

The appetite for post-compulsory education, for example, is only remotely connected to the mechanisms emphasised by "evolutionary psychology". It increases generation on generation, despite the fact that educating individual humans significantly reduces their inclusive fitness.

The effect outweighs the increase in parental investment per child. And (as high rates of dropout suggest) the level of human organization (gene, individual, nuclear family, extended family, peer group) of the units of selection is a complex problem.

Uneven, often countervailing, "co-evolution" of memes and people complicate attempts to classify the relations between memes and people. Dennett categorizes memes three ways: good, pernicious, and mixed. His categories are related in a general sense to Ball's, but more pragmatic:

"...many - most, if we're lucky - of the memes that replicate themselves do so not just with our blessing, but because of our esteem for them. I think there can be little controversy that some memes are, all things considered, good from our perspective: ... such general memes as music, writing, education.... The Marriage of Figaro, Moby Dick, and returnable bottles...Other memes are more controversial; we can see why they spread, and why, all things considered, we should tolerate them, in spite of the problems they cause for us: shopping malls, fast food, and advertising on television. Still others are unquestionably pernicious, but extremely hard to eradicate: anti-Semitism, hijacking airliners, computer viruses, spray-paint graffiti." (Dennett 203)

Dennett's "our" and "we" contain a problem. I mentioned it earlier when I referred to "humanist" illusions in studying memes. Why are "pernicious" memes so hard to eradicate? (Those that are readily eradicated cease to be memes at all.) Even "pernicious" memes are good from some peoples' perspective; that is the major reason why they are so hard to eradicate. One person's meme is another person's poison, but the reverse also applies. Graffiti and racism would not need control if there did not exist reservoirs of support for them in the human population. Many memes divide people, rather as diet divides species, or reproduction divides sexes. {Footnote 52}

A "pernicious" meme (in Dennett's terms) is one that changes categories (in Ball's terms) - as you shift from organism to organism or from perspective to perspective. Many memes are of this type, or contain such components. Ball's typology classifies behaviour socially according to its effect on memes (how far does it promote the meme-type's replication and proliferation in the society of memes), but only individually for people (how does it affect the individual's functioning). This may be why his example of a symbiotic meme ("Sex is fun; pass it on") seems slightly unconvincing. Societies persist and expand even with powerful memes prohibiting and restricting sexual activity. Indeed human societies, as institutions for raising single, socially-competent offspring through protracted juvenile dependency and heavy parental investment, depend on such memes.

Putting Dennett's and Ball's typologies of memes side by side underlines the fact that a framework for understanding organism-meme co-evolution must take account of the replicative machineries of each. Dennett's categories do not really distinguish mechanisms of meme replication or variation in memes' replicative effectiveness; Ball's do, but only to the extent of distinguishing memes that transmit well (by whatever means) from those that do it poorly, and transecting that distinction with a comparably coarse distinction of adaptiveness for the individual human carrier.

It seems to me unlikely that a satisfactory framework can be developed by multiplying such distinctions: the number of categories would explode far faster than their discriminative potential. I return to this problem below, when I discuss the "standard social science model" versus "evolutionary psychologists" collections of specific adaptations. {Footnote 53}I shall also

return to the question, implicit in Ball's and Delius' "parasitological" classifications of memes: what sort of memes, and in what sorts of ways, can be considered analogous to organisms or species?

## Meme replication and imitation: humans and other species

What can we say about how memes reproduce and proliferate themselves? The differences with biological organisms are at least as important as the similarities. Memes, like genes, do replicate (as Dawkins has it) by "imitation...in the broadest sense". But we must be cautious against interpreting "imitation" *over* literally. {Footnote 54} Although imitation may be the principal means of cultural transmission in other species than humans, in humans it is only a particular case. (I leave aside the interesting problems posed by human training of other species. {Footnote 55})

When scientists studying primate behaviour watched a troop of macaques trying to extract food grains that had been mixed in sand, it was a middle-aged female who first puzzled out how to separate out the grains from the sand (she did it by putting the mixture in water, so that the sand sank, while the grains floated). In this species the innovation spread through the troop by imitation, becoming generalised in a few generations. {Footnote 56} We might describe it as becoming "fixed" as an adaptation of the monkeys' culture - if it were not for the difficulties of applying notions of adaptation and speciation to culture.

Other cultural behaviours among animals pass between individuals by more restricted imitation. Local "dialects" of bird song, for example, are transmitted from mature adults (with by-then fixed song patterns) to fledglings during the latter's "sensitive" periods (Catchpole 1986, Lynch and Baker 1993).

Among human beings equipment for meme replication and transmission has become specialised and highly elaborated. Even where cultural transmission is more-or-less replicative, it may still be rather indirect, the "teacher" not actually doing what the learner is learning to do - and sometimes the latter not doing it either. Moreover replication of behaviour is only one type of cultural transmission. When we ask empirical questions about some of the arrangements memes have evolved to propagate successfully, we should bear in mind that meme propagation does not take place only by imitation - or, like genes, by the use of physical templates and moulds (though it is sometimes does - as in printing or factory production more generally). {Footnote 57}

Moreover, even insofar as replication is involved, there are other fundamental differences. Memes adapt not only particular variation by particular variation, but also by rewriting of the codes in which they are inscribed and transmitted. This can be quite literally the case. The evolution of language, writing, or the alphabet, has revolutionized the overall ecology of memes almost as fundamentally as the cell has revolutionized that of life. In both cases a "pause" for investment in structure pays off as faster evolution subsequently. But one difference is that memes have already spent a significant part of their histories "tinkering" with the codes they themselves are inscribed in, and, what is even more important, devising new materials to write on and with; this is a matter that genes have only recently started to address.

## **Variation and differentiation among memes**

What are the sources of variation in memes? We already know, in general terms, that the sources of memetic variation are wider than those of DNA-based life, since they include both random replication errors ("mutations") and deliberate or partly deliberate variations of design. Does the biological analogy, at this point, career off into unmanageable dissimilarity?

Not, I think, if we are careful. Let us prune the problem down with a narrower analogical question from biology: what sorts of arrangements have genes (and memes) evolved in order better to control their own variation and to increase the evolutionary advantages they can derive from it within their particular environment? For genes, remember, the particular environment is very largely made up of "adjacent" genes. What sorts of adaptations have gene-complexes and their survival machines evolved in order the better to regulate their own evolution?

Biologists regard two general forms of differentiation as particularly important among genes: speciation, and sex. Organic life separates its forms into isolated gene pools. And, within many of those pools, it generates opposite processes which continually stir up and remix the gene pool from generation to generation. One of the most important of these is sexual recombination.

Speciation, often in combination with sex, bifurcates the "bush" of evolution. There are numerous problems in the relations between (biological) speciation, and sex, and evolution. But our main question is: can we point to analogous mechanisms for cultural - memetic - evolution? I believe we can.

One of these we can describe, as a first approximation, as combined specialisation and concentration of memes. It based on the tendency - indeed the necessity - of memes to "cluster" according to similarities and complementarities of function. This tendency is so general that it is easy to overlook it. Selection among and in favour of more robust meme complexes - and therefore of larger and more complex ones - tends to eliminate loosely-bound memes, just as natural selection acts against the re-emergence of very short DNA sequences. {Footnote 58} {Footnote 59}

## **The human structure of culture**

Some types of clustering are rather closely grounded in cultural evolution's biological basis - that is, in the lifestyle of a socialising savanna primate with an erect gait, reversible thumbs, moderate sexual dimorphism, single offspring, a prolonged fetal/juvenile phase, a large cranium, and a typical life cycle of a few decades. Just as genes build up their "survival machines" - organisms - over vast periods of time as slow, imprecated accumulations of adaptations to long but finite DNA sequences, so culture develops through the more - and less - gradual succession of meme adaptations. This sort of "clustering" - perhaps "chaining" would be a better term - seeks, often literally as well as metaphorically, to exploit the finite capacity of individual human beings by combining them in organised ways. The sexual and occupational divisions of labour are central examples. Each new modification starts from the shapes set by previous cultural evolution. Neither "clustering" nor "chaining" are ideal terms; perhaps "cohesion" or "integration" would be better.

You can get a cultural quart into the biologically given pint pot - but not immediately, and only through a variety of indirect means. You can stretch the pot just a little (though the anatomical compromise between brain size and human birth canal has already set a tight limit on this). You can redesign the pot a little - which may also involve reorganizing the pottery. More important you can increase your overall meme capacity by overflow and interface devices: ducts, siphons, membranes, tubes, pumps, valves and so on connecting the pots one to another and allowing - or driving - their contents to move to and fro between them.

Even more important you can work on the contents of the pots, compressing, concentrating, summarising, generalising, enriching, so to speak, the cultural soup, so that any given configuration of pots has a greater nutritional capacity. And, perhaps most important of all, as the cultural broth cooks and condenses, parts of it will take solid shape, "caramelising", so speak - and (to conclude the domestic metaphor) these crystalising solids offer - among other things - possibilities for making new types of pot and - no less significant - new types of connections between pots. Of course, many of these constructions are relatively unsuccessful and are sooner or later redissolved into the memetic soup. Others, however, endure or get replicated. And as they do, they vary.

Human beings are certainly not the only organisms to link intentions or goal-seeking behaviour over time or across individuals. {Footnote 60} But they are distinctive in the elaborate chainings of intentions that they sustain and wield. In us memetic evolution has chanced across a species that can bond together vast linkages of memes. They and we form "strategic" structures that, pushed at one node, produce action, both intended and unintended, at enormous spatial, temporal and personal distances. When the ideogram-reading Korean factory worker stenciled on the packaging of my personal computer the roman lettering that guided it to my desk, s/he also allowed me to quote to you Aristotle's definition of man as a political animal, and caused this paragraph to carry its meaning from me to you.

Most of these great "chainings" of intentionality are invisible or incomprehensible to the individual humans who link them at particular loci, but through socialisation they come to participate in them with impressive facility and confidence. One of things this has encouraged (or selected for) is our ability to recognise and pursue - individually and in concert - large and remote purposes. Hence our species' itch for meaning in life.

## Literacy

One of the most familiar forms of clustering among memes is expressed as cultural differentiation and occupational specialisation between people. Consider one important type of this, which also fixes and transmits intentions in tangible form: writing and associated skills. Like all culture, writing and reading have evolved through increasingly complex differentiation. Originally they were arduous, highly specialist skills, never mastered by more than a tiny, relatively privileged, section of the male population, who learned their scribal skills in a form of lifelong apprenticeship. {Footnote 61} Today fluent literacy is general, and has integrated itself into numberless other technologies.

Between then and now there has occurred a cumulative series of memetic innovations, of which one of the most important is the invention and imposition of the phonetic alphabet. This reduced writing's total of "atomic" symbols to a couple of dozen, and now allows it to be mastered by most six year old humans. There are many others, not least the development of the key means through which I (though distant from you, and perhaps dead) am enabled to speak to you: silent reading, and modern western punctuation. {Footnote 62} To return to our metaphor of pots and cooking, the history of writing is one of experimentation and successful innovation in making connections and interflows between pots and between connectors.

The Greco-Roman alphabet, for example, has proved a remarkable success through simplifying the meta-connection between spoken and written language. In contrast, Chinese ideograms lack the direct linkage between script and pronunciation. Spoken Chinese evolved large differences between geographically separated areas of China, with the result that officials who corresponded without difficulty could, when they met, sometimes not understand each other through speech {Footnote 63}. In recent times there has been persistent pressure on Chinese to be more conformable with the Latin alphabet. {Footnote 64}

Writing is an example of equipment that meme-complexes have evolved to reduce and correct copying errors in replication. Writing reinforces memes by extending their persistence through externalised physical texts, which fulfil some of the functions of transmission between people in non-literate cultures. It slows (or stabilises) "drift" in transmission; subsequent inventions, such as printing, further strengthen error correction. Printing does this more efficiently, essentially by increasing the resources devoted to error-checking. Printing with moveable alphabetic type, which reduced copying errors in scholarly publishing by an order of magnitude, type made Europe's scientific revolution possible. {Footnote 65}

The history of writing, and meme evolution in cultures embodying it, suggests something else. What once filled a few pots right to the brim (literacy) has now, through distillation and some other chemistry, become an unremarked minor ingredient in the contents of virtually all pots. But - since culture, too, abhors a vacuum - the spaces thus released (and already suffused, so to speak, with literacy) are soon filled with new - and newly differentiated - contents. These processes of concentration, accompanying the growth of the total of human knowledge, are the obverse of the continual recomposition of specialisms.

## Academic knowledge

The modern academic division of knowledge offers an example. The boundaries between different disciplines and sub-disciplines are rigid in the short term, but more mobile in the long term. But focus, for a moment, not on the detailed placing of divisions, but on their general scale. Very approximately we can discern limits to the scale at which specialist academic disciplines can "deposit" themselves in and around individual humans. One minimum limit is set by the human life span, the period of maturation and working life, and the existing forms of specialist culture embodied in education - the duration and scope of the undergraduate curriculum, graduate training, etc. The "Not my period..." by which the twentieth century medievalist affirms and distinguishes his credentials is the lineal - we might almost say embryological - descendent of the elongated apprenticeship of his cuneiform scribal ancestor.

Humans' physiological scale sets a variety of such limits. The span of consecutive adult attention sets the optimum duration for a dramatic performance, around one-and-a-half hours - which proves also to be about the same time typically allowed for classes in higher education. In each case the amount, and complexity, of the matter for exegesis must be constrained in order to be effective. If the producer wishes to push the duration, s/he is well advised to punctuate the proceedings with intervals, or, in extreme cases - Wagnerian opera, for example - with dinner breaks.

When a segment of culture - a population of memes, like science, or academic learning - comes to be finely divided among humans, both the memes and the humans must be organised. The memes are grouped in part to realise currently available economies of human effort involved in learning (and therefore teaching) them. Academic specialisation and the internal organisation of curricula express adaptation of this sort. Their groupings bear no very direct relation to what memes are "about"; the arithmetic developed on fingers later gets applied to football coupons and physics. But one thing the groupings *are* compelled to stay in close contact with is their own past histories. At this organizational level academic specialisms are largely held together by their shared historical descent in joint "survival machines" - subjects, textbooks, institutions, etc. But there is a positive side, in that crowding promotes synthesis: the importance of a scientific work (as the mathematician Hilbert remarked) can be measured by the number of previous publications it makes it superfluous to read.

Other limits of scale are set by the natural history of institutional life, together with some elements of game theory. For a subdiscipline to nucleate out requires, on a typical faculty board, a minimum coalition size of about three or four, or perhaps a little more. Two people is rather rare; perhaps it is too few for comfort, or for political triangulation. Barkow (1989, pp 12-4) emphasises ways in which interdisciplinary frictions and methodological spaces between disciplines may mutually sustain each other round historically arbitrary "cores". The mechanisms, he holds, are rooted in academics' appetite for reputation (see also Whittle 1985). Macro factors also play a part: the viability of journals, the scale of experimental work and the numbers and funds required for it, and so forth.

## **Cellular knowledge**

Academic understanding and science, like much of the rest of human knowledge, develop "cellularly". Of course there are extra-cellular materials and tissues as well as intra-cellular ones, but understanding what goes on around and between cells depends heavily on our understanding of individual cells and their internal mechanisms. Human biology and organizational history set elastic but finite limits both to the possible functions and sizes of cells, and to the next modifications of existing organs. Bodies of theoretical knowledge become institutions, not only in a bricks-and-mortar sense, but in the self-regulating sense of consisting in a settled division of skills, training and effort among (replaceable) human beings.

Tensions persist, especially at the boundaries of different disciplines and - equally importantly - at the potential boundaries between disciplines which are presently far apart. But everyone now uses, every day, myriad artefacts, conventions and theories which they are quite incompetent to understand or maintain, and which to most users are "black boxes". Does this mean that common

sense is being overtaken and absorbed by specialised knowledge? Well: yes and no. Specialised "research" organs continually ingest new material. The products of some of this, digested, then pass along general purpose channels to power and build the organism. The cathode ray tube, or slogan, or drug, or electronic charge card, or tax loophole - meme - proposed today by a few specialists gets embodied as tomorrow's daily, commonsensical implement.

## Science

"Let her and Falsehood grapple; who ever knew Truth put to the worst, in a free and open encounter" - Milton, Areopagitica

Milton's conditions for truth are poor ones for science. The memes of theoretical natural science, as Wolpert (1992) points out, are highly "unnatural" memes, remote from "common sense". Like cattle or sheep, they have been bred for generations into the forms preferred by their domesticators (of whom some of the most important are other memes). Testability, generality, uniform vocabulary, unambiguous meaning, internal consistency, and so on - even taken singly such traits are rare memes, and to assemble them all requires long intentional selection. The domesticated memes of theoretical natural science, having embodied such significant adaptations to artificial circumstances, could no longer survive reintroduction to the wild. They can live and breed only with the aid of rather complex arrangements to sustain them. The cultivation of theoretical science (like keeping sheep) has come to rely on auxiliary breeds, such as scientists - rather like sheepdogs, who keep the flock together and bark at intruders. By comparison, much social science consists of more "common sense" memes, less "deformed" by domestic breeding. They more resemble semi-domesticated breeds which forage freely on the mountain slopes in summertime, but are herded in for the winter. {Footnote 66}

One of the most interesting of domesticated organic species is humankind: interesting because humans are self-domesticators. We serve as both sheep, and sheepdogs, and instructors of the sheepdogs. To instruct each other and ourselves we deploy memes. We are now far too far domesticated, too dependent on memetic cultivation, to survive in the wild. {Footnote 67}

### **Do memes have "survival machines"?**

When I referred to individual people as "vehicles", each transporting a collection of memes, I was using a rather approximate analogy. For gene-centered biology the fundamental unit on which (biological) natural selection operates is the gene. Organisms - the expressions of developmental programmes inscribed in genomes - are differentially selected, but what is important is that some genes get selected in preference to their rivals. Organisms are "survival machines" built by combinations of genes (genomes) as their means for survival, replication and proliferation.

In multi-celled organisms, the genome builds "survival machines" at two interdependent levels. At an aggregate level the genes specify a developmental program for an organism. But, more immediately, they specify a developmental program for each cell, with the site and context of each cell interacting with the genes to determine the cell's destiny as a component of the whole organism. So genes are carried and expressed in "survival machines" at two levels of

organization; individual cells develop so as to work effectively within a whole phenotype; the gene's chances of being transmitted reside both in the attributes of the whole organism and in the attributes of individual cells.

How far could we use an analogous model for memes? For example, could we think of an assemblage of memes being brought together in the socialisation and education of an individual human, then of the interplay of suitably prepared individuals in memetic "organisms" - as conventions, rituals, institutions? Culture, on this view, would be reproduced through processes of replication on (at least) two levels: individuals, and institutions.

The parallel has its points, but also its limitations. Certain memes (a common language, for example) are involved in the formation of all or virtually all individual humans within a given (national or linguistic) culture. Like the common genome in cells' nuclei, it is repeated almost identically in all the "cells" of the "organism". But the particular part that a cell is called upon to play within the life history of "its" organism is affected not only by the developmental program written in the genome, but on the time and the anatomical and physiological context under which the developmental algorithm of the individual cell starts running and gets committed (i.e. where and when it is formed). The genome of every cell plays the same basic melody, but on an instrument, at a pitch and volume, etc which depends not only on what is in the score, but where within the orchestra the instrument develops. Like multicelled organisms, the operation of an organisation as a whole depends on a complex differentiation among parts which are formed on a modular basis. {Footnote 68}

Perhaps, though, the orchestra metaphor has allowed us to slide too easily from genes to memes? One key issue is how far we can identify things which, replicated across individual humans, help us account for the characteristic, enduring qualities of institutions (or organisations, or cultures) across generations.

Darwinian evolution is not confined to organic life-forms, nor is it dependent on particular mechanisms of either replication or reproduction. Organisms, like cells, descend from similar parents, often by the more roundabout processes of sexual recombination. Persons are culturally assembled from a range of other people and memes, some closer, some remote. For example, children acquire much of their first language in imitative, repetitive play within the nuclear family: the "mother tongue". However, they acquire literacy later, in forms which are critically dependent on earlier language development, within formal organisations specialised for that purpose (schools), overwhelmingly composed of other humans who are not close genetic relatives. Literacy develops in the individual human partly by processes of reciprocal imitation with others (peer groups), and partly through functionally specialised humans (teachers).

## Views about meme vehicles

In fact, writers about memes have rather various views of what relationships hold between memes and entities that serve as their "survival machines" or "vehicles". Dawkins, in his original formulation, starts from the view that memes replicate by imitation. He then speaks at various points of a variety of different things which can transmit memes: individual humans (or their brains), ideas, songs (as in Auld Lang Syne), physical artefacts (such as hymn books), symbolic

systems (such as languages), formal organizations (such as churches), and social institutions (such as priestly celibacy).

Dennett argues similarly:

"Genes are invisible; they are carried by gene vehicles (organisms) in which they tend to produce characteristic effects ("phenotypic" effects by which their fates are, in the long run, determined. Memes are also invisible, and are carried by meme vehicles - pictures, books, sayings (in particular languages, oral or written, on paper or magnetically encoded, etc). Tools and buildings and other inventions are also meme vehicles. A wagon with spoked wheels carries not only grain or freight from place to place; it carries the brilliant idea of a wagon with spoked wheels from mind to mind. A meme's existence depends on a physical embodiment in some medium; if all such physical embodiments are destroyed, that meme is extinguished... "Meme vehicles inhabit our world alongside all the fauna and flora, large and small. By and large they are "visible" only to the human species, however. Consider the environment of the average New York City pigeon, whose eyes and ears are assaulted every day by approximately as many words, pictures and other signs and symbols as assault each human New Yorker. These physical meme vehicles may impinge importantly on the pigeon's welfare, but not in virtue of the memes they carry - it is nothing to the pigeon that it is under a page of the National Enquirer, not the New York Times, that it finds a crumb." (Dennett, CE, 203-4)

Let us look at some of these alternative meme vehicles in a little more detail. To what extent are they actually alternatives - rather than being, say, different phases of a developmental sequence? By looking at this we may be able bring the role of "imitation" in meme replication into clearer focus, and unpack some of the ways in which memes "leap" from one subject's mind to another's.

## Physical artefacts

"No part of the walls is left undecorated. From everywhere the praise of the lord is drummed into you." - Nikolaus Pevsner, London. "Who *is* the Potter, pray, and who the Pot?" - Rubuyat of Omar Khayyam

Even in some of the "simplest" forms of artefact-making there can occur great stability of the artefact's form over many (human) generations. The stone tools of pre-agricultural humans, for example, evolved fast in terms of biological time, but very slowly compared with today's rates of technological change. Their characteristic forms were highly stable in the shorter term. They developed over millennia through sequences of styles in the end-product - for example, flint arrowheads. The overall evolutionary development through recognizable styles often allows archeologists - like paleologists - to date and source particular examples rather exactly. The evidence as a whole points to increasingly competent and exact processes of imitation.

Most (physical) artefacts arise from processes of replication and imitation. But the two are not the same. What, precisely, got imitated in the manufacture of early stone tools? It is a reasonable surmise that what was being imitated, from human to human, elder to neophyte, was not (or at least, not only) some finished example of the artefact itself (the flaked and shaped arrowhead) but rather the whole developmental sequence that led to it: from the searching out and selecting

of suitable pebbles (all of which would be quite different from the finished product), via rough shaping, finishing, selection and perhaps grading, to final use. With our modern, elaborated division of labour between people the developmental processes of an artefact - considered as one sequence, from location of raw material to finished artefact-in-use - has been broken into numerous lesser, recombining, parts and much of it has passed beyond any one individual's view. If we want to reconstruct it we must do it in our mind, with the help of some very recent artefacts (including the technical and social sciences).

What gets imitated, therefore, is not immediately the finished, developed artefact, but its process of manufacture or development. And one of the intermediate products that must be reproduced is the artefact maker. {Footnote 69} Imitating another's artefact is, first of all, imitating the Other, and that imitation is itself a developmental process. The education of an individual in a skill - the skill (or more exactly, the ordered sequence of sub-skills) for producing an artefact-meme, in this case a Neolithic hunter's particular style of arrowhead - requires that the particular meme's specialised reproductive equipment - the apprenticeship of the arrow maker - itself be reproduced. Neolithic sites can often be recognized by their accumulations of unfinished tools and related debris, by-products of the joint production of tools and toolmakers. Education, including the imitation of other already-educated humans, is how a technology, a language or a culture endures across generations. As Dennett puts it:

"The haven all memes depend on reaching is the human mind, but a human mind is itself an artifact created when memes restructure a human brain in order to make it a better habitat for memes. The avenues for entry and exit are modified to suit local conditions, and strengthened by various artificial devices that enhance fidelity and prolixity of replication: native Chinese minds differ dramatically from native French minds, and literate minds differ from illiterate minds." (Dennett, 207)

Much change in artefacts, as in other memes, is marginal. It arises by design but not by inspiration. The usual propellant is dissatisfaction with existing forms and the consequent, iterative search for an improved form. Henry Petroski, writing on the history of design, argues that, contrary to design theory's conventional wisdom that "Form follows function" (design theorists being, so to say, the ideologists of successful designers), it is more accurate to say that "Form follows failure", and to see the inventor primarily as a critic, not as an artist (Petroski 1993; cf Oscar Wilde, *The Critic as Artist*) {Footnote 70}

## **Life cycles and art cycles**

"God is really only another artist. He invented the giraffe, the elephant and the cat. He has no real style. He just goes on trying other things." - Pablo Picasso (*Gilot and Like, Life with Picasso*)

Even in creative modern art, much modification of memes remains marginal. Here is how David Hockney improved his memes to help them gain better access to human minds, "the haven all memes depend on reaching":

"I realized that was the way most people were going to see this work, in this form of reproduction, because the number of people who could see the real piece, which is ten feet across, would be quite small. If you sell five thousand posters, not only are five thousand people seeing it, but maybe ten other people see each poster, so it means a lot of people. Therefore, I thought if that is the form it is going to be seen in mostly, let's make sure we do it as well as we can technically. That's why my posters began to be rather good. We took a little more trouble just making sure the work was photographed very well and the printing was of good quality. I was surprised that so few artists seemed to take an interest in that. I think now more and more artists are becoming aware that reproductions of their pictures are the way most people see them. After all, artists should have known that, because that's the way they themselves saw a great number of pictures." {Footnote 71}

Hockney goes on to describe how memes not only enter minds, but re-engrave them - how repotting remakes the repotter:

"For me, reproductions of my work were something that became more and more important, beginning with my first big book, Hockney by Hockney. I realized that not only had I got my work at my fingertips, I could flick through it, flick through my life as it were; I realized too that others could flick through my work, and I think this is also a part of art." (Hockney, 1993, 115)

Only in the fairly recent past (about the last two centuries) has human society become information-rich through memes capable of mass manufacture or transmission. Where, as recently as the last century, individual humans competed for news or entertainment, now the boot is on the other foot, and an "info glut" of media competes furiously to be noticed by as many individuals as possible. {Footnote 72}

Post modernity is society in which the key scarce resource is public attention (Baumann). But competition for this resource takes place among memes whose basic structures evolved in information-poor societies. A fifteen second television advertisement may assume fluent literacy in its audience, flashing its message in glimpsed, oblique word-fragments. But the alphabet it uses evolved nearly 3000 years ago in the eastern Mediterranean in peasant societies where a large majority of people could neither read nor write, and a large proportion had never *seen* a written artefact. The conservative element is not the physical artefact, whose design and production methods have been many times revolutionized, but the symbolic conventions with which human generations engrave their successors. {Footnote 73}

Even apparently "simple" artefacts, such as our Neolithic arrowhead, raise the questions: What gets replicated? What reproduces? What, if anything, corresponds to the biological organism in population genetics? One of the difficulties in reconstructing the development programs of memes is that they involve the confluence of other and successive tributary inputs. "Tributary" itself oversimplifies: many intermediate products serve as input to several outputs. And many output-memes can be produced by a variety of combinations.

Over time combinations vary and alternative combinations get selected - evolve - so as to "enhance fidelity and prolixity of replication". Selection takes place partly through and partly

outside human intentions. An important part of technical evolution is change in the artefacts used to produce other artefacts: finger-molding of clay, the potter's wheel, the glass factory.

Innovation in the end product is often an almost incidental result of selection among variants at earlier stages of a developmental sequence. Additions to the repertoire of intermediate artefact-memes - the wheel, iron smelting, paper, the steam engine - bring change virtually throughout the system of production.

## How do meme-vehicles reproduce?

Where, then, in all this, should we identify memes and their vehicles? Is a completed artefact a "stably co-adapted complex of memes", or is it just one phase of the former's life-cycle, its "vehicle" or "survival machine"? If the latter, where are the memes - invisible, according to Dennett - that persist by replication through the life-cycle of the organism? How are memetically-structured (trained) human brains integrated within artefact-memes' development sequences? In particular, what gets replicated when the imagined idea of the pot gets shaped on the potter's wheel, or (to invert the problem of signifier and signified) when repeated practice at the elbow of a skilled potter educates a new potter?

Two differences between genes and memes (or, at least, the artefact forms or phases of memes) seem to me particularly important. In the first place memes in general have no physical "stencil", analogous to nucleic acid sequences, preserved through transcription and replication right across the life cycle; the mechanisms of replication are quite different. In fact DNA does not replicate itself directly: its double strands first separate into two single strands, each of which then bonds to free bases to re-form the double structure of DNA. There are more general reasons why nothing can replicate itself directly, producing itself, like the Christian god, from out of itself; the same thing cannot act simultaneously as maker and made. Replication (and its conscious form, imitation) must occur through at least a double transcription. {Footnote 74}

Secondly, artefact-memes can tolerate far more radical changes in the developmental sequences from which they issue than can genes and their survival machines. Major technical changes and substitutions can be introduced, a long way ahead of the finished product, which are not lethal but, on the contrary, increase both the efficiency of replication and the quality of the product: paper in place of parchment; animal muscle, steam, or electricity instead of human effort.

These two differences lead us into further problems. If biological replication takes place at two general "levels", the cell and the organism, how many levels can we point to for memes? What is a meme's equivalent of the organism's life-cycle, in which development takes place through a structured sequence of forms, including fertile phases but ending in the death and disintegration of the organism? Indeed, can we draw any generally useful distinction between genes and their vehicles or survival machines? Let us look at some examples of (complex) memes and the problem of analysing them into their conceptual components.

Memes must replicate. For them to do so there must be adapted equipment or arrangements in place. But must memes be self-replicators? Physical artefacts, on the whole, are not - if we leave aside particular and arguable examples, like machine tools, or software for writing software. Arrowheads do not reproduce arrowheads, and the inputs to a baked bean cannery don't include

tins of baked beans (they do include uncooked beans, though slightly fewer emerge tinned at the other end). The meme-artefact has - so to speak - externalised its reproductive arrangements (the physical inputs and the meme-structured bean and can makers and their motives), and these are renewed through separate cycles of replication: the production of beans and tomatoes, steel, paper labels etc, and the cycles of socialization, education, exchange and employment that deliver the right mix of skilled and motivated workers to operate enterprises (for making beans, for making bean and can making equipment, and for making bean makers). Artefacts are produced by means of artefacts, commodities by means of commodities, people by means of people, and each of these have histories of memetic evolution. People and their psyches (and beans and their life cycles) also have histories of biological evolution, fixed as adaptations in almost all individuals.

Perhaps, though, it is more helpful to think of artefact-memes not so much as having externalised reproductive apparatus, but of the physical output as being part - just one phase - of a larger and longer self-replicative cycle. Biologists are used to thinking of organisms as self-reproducing machines, and of their design features as selected to be functional from this point of view. Could we not, by extending meme-processes, identify replication cycles which correspond more closely to those involved in biological reproduction?

If it makes good sense to biologists to think of a bird as an egg's way (or its genes' way) way of making another egg, or of flowers and their pollinating insects as parts of common reproductive cycles, does it not make equal sense to think of the species of bird as a particular type of nest's - or song's - way of replicating into another nest or song? And, by extension, can we not think of the bean cannery and its auxiliary agriculture as (another, intersecting) replicative route involving beans? {Footnote 75} Plants have "learned" over evolutionary time to use the distinctive behavioural patterns - the proto-motives - of insects, birds and mammals as they feed, nest and excrete to pollinate their flowers and disperse their seeds; similarly artefacts have adapted to use humans' complex patters of behaviour and motivation as reproductive equipment.

In fact, with higher forms of moral meme, the idea is already familiar to us. We have no especial difficulty in thinking of generations upon generations of human beings in the service of art, science, the mission of the church, or the juggernaut of progress. Evolutionary biologists do not think - except as a deliberate approximation - of species' evolving in a static niche, but of the co-evolution of many species (or, if you prefer, of arrays of many niches, and of the species which define and compose them {Footnote 76}).

As far as physical artefact-memes are concerned, we cannot point to common physical structures at the base of their forms of replication. Moreover their forms of reproductive separation and isolation are quite different from DNA's - another aspect, from one point of view, of the externalization of their replicating equipment. Replicating cans of baked beans need never involve starting with other cans of baked beans - though it well may, if the canners retain their appetite for baked beans. More generally the replication of artefacts, commodities and people takes place by means of artefacts, commodities and people - but different ones.

Cells, organisms and species can each be regarded as different routes for complex molecules' self-replication, and self-replicators' self-replication, and so on. DNA's replicative cycles vary,

and selection takes among the variants, speciation being an advanced result of the process. The effects of biological evolution include many striking examples of functional integration and trade off between competing variants and meta-variants of nucleic acid sequences: exquisite balances among genes between the needs of survival and replication; intimate and exact co-evolution of functions among several species; ecologies as dynamic equilibria of (specialised and reproductively segregated) populations of molecules, cells and species.

Species, however, like individual organisms (or birds' songs, or candles' flames) are not static objects, but temporary processes. There is always the possibility they may go extinct or subdivide. For this reason the biologist GC Williams emphasises the clade as the general unit of selection. A clade corresponds to a set of entities of common ancestry (monophyletic). It may be a strain, a species, or a larger category. Clades "compete" at many levels, but

"The ultimate prize for which all clades are in competition is representation in the biota"  
(Williams 1992, p25) {Footnote 77}

Recall Dennett's similar observation about memes:

"The haven all memes depend on reaching is the human mind..." (Dennett, 207)

But is Dennett's formulation a necessary truth about memes - that they can only infect human minds? Dennett adds that a human mind is itself merely a particular type of meme-inscribed artefact. Culture consists not only of minds but of many other artefacts. Cultural and technical evolution has seen memes develop increasingly elaborate, varied and circuitous routes outside minds, many of them inscribed in codes that evolved first in human minds. Could memes emancipate themselves from human beings altogether? They show tendencies to do in various directions, from the permanent transfer of memes from human culture to that of other animals [tits and bottles, bower birds], to incipient pattern recognition and intelligence among computers.

The meme-family of conspecific aggression among social mammals (and its inhibitory limitation) long predates hominids. But with the aid of humans it has "domesticated" many other vehicles: human artefacts from stones to nuclear weapons; other species (especially horses); more general memes of cohesion among humans (from parental investment to proletarian internationalism); and numerous more-or-less complex formal organizations, from the warrior band and the air force, to space technology and the Red Cross. Warfare, having engraved itself among our memetically transmitted preferences, seem to be proliferating by emancipating itself to an increasing degree.

Perhaps Dennett's insistence on the human mind as memes' irreducible "haven" is a residue of the "mentalism" or "humanism" that has - understandably - proliferated in memes about memetic evolution? Conclusively to establish the need for Dennett's "haven" we would, in a sense, need to prove a negative: that no meme could remain aloft indefinitely without perching for rest in a human mind.

There is nothing inherently impossible in the idea that the lineages of patterning that form life could be coded in other replicative structures than nucleic acids. Cairns-Smith has developed the

speculation that the early history of life was formed on clay crystals, only subsequently being transferred to organic molecular chains.

[Perchless swallow. Human history of memes just an early segment of their learning curve?]

## Money

"Money is like a sixth sense without which you cannot make a complete use of the other five" - W Somerset Maugham, Of Human Bondage ["Life is too short to do anything for oneself that one can pay others to do" - W Somerset Maugham, The Summing Up "As I take my shoes from the shoemaker, and my coat from the tailor, so I take my religion from the priest" - Oliver Goldsmith, quoted in Boswell, Life of Johnson (DQ)]

Organic life may almost be defined as patterns which have taken nucleic acid chains as their means of replication. By analogy, memes can provisionally be thought of as replicators which have taken vast chainings of human intentions as their means of replication - presupposing life just as life presupposes replicating molecules. However, while an individual intention necessarily has a psychological ingredient, a chaining of many intentions is not necessarily intentional - though it may be. This - the anonymous necessities of the social - is a problem I return to. First, however, I glance at some ways in which individual intentions can and do unknowingly bind together - money: law; crowds; and formal organisations.

If artefact-memes do not start from a common evolutionary substance - as life has, in large-self replicating molecular chains, varying and being differentially selected as they replicate - perhaps they may nonetheless have evolved, or be able to evolve, an analogous common substance? Something of this sort, I think, already exists, in the form of money. The differences - the external character of the common substance, and the reversed evolutionary ordering - correspond to the facts that while large carbon molecules and their macrostructures replicate themselves internally, from themselves, memes are replicated externally and by other memes (and by people). One - relatively late - result in their evolution has been the evolution, through exaptation, of a meme specialised in exploiting a particular niche: barter, leading into more generalised exchange.

One way of viewing life is as a vast set of interlocking mutually dependent routes for proliferating variant sequences of DNA. Similarly, a monetary economy consists of a vast set of intermeshed alternative pathways replicating and expanding money. In the first case the common basic replicator is a starting point, in the second it is a result, but the effect is somewhat the same. Being external and a relatively late product of evolution, the common replicator can also evolve to fulfil external functions - such as equilibration across the overall mass of value expansion. The adjustments this brings into play also tend to maximise the rate of expansion of the whole mass of money. {Footnote 78} Indeed macro-economics is, in the view of one influential school, simply artefactory production with diminishing marginal returns - only writ large.

Money is also central to one of the ways artefact-memes expand the mass of value - through accelerating, creating, surveying and filling new openings in the overall replicative structure. Biological life, originally equipped only with direct, exact replication plus occasional (random)

mutation, has had to evolve special, subsequent arrangements: speciation, to allow its ocean of genes to be divided into separate, specialising "gene pools", and, within many of those pools, sexual recombination to stir the genes more vigorously. Being superimposed on underlying, more fundamental, properties of the genotype which were selected for distinct reasons at earlier evolutionary phases, biological speciation and sex produce relatively awkward and slow-changing innovation mechanisms - compared, say, with some cultures, such as market economies. Viewed as a replicative mechanism money, partly through coming later in evolutionary development (insofar as we can make such comparisons across "zones"), also proves better at innovation. {Footnote 79}

When evolution works on intentions it tends to do so "backwards", selecting and thus modifying the intentions and expectations which issue from psychological processes via selection of means serving intentions further "down the line". An important class of social arrangements consists of "economic" meme-complexes which have the effect of shaping the intentions of some persons (slaves, workers, officials) in such ways that they meet the needs or desires - respond to the intentions - of others (e.g. owners, consumers, citizens etc) who are so remote they have never even heard of those to whom they respond. Individual humans are socialised and placed in their social classes partly by their minds being "engraved" (Dennett, check) by socialising memes which cause them to respond stereotypically to day-to-day memes.

Meeting the desires of some through the intentions of unrelated others is an important separation. It contrasts with the predominance of kin investment in many other species. {Footnote 80} Its effect is to attenuate the balance between effort and satisfaction that regulates the life of individuals among most higher animals: the fact, for example, that many animals cease foraging or predation once satiated. De-coupling the location - the human site - of the desire from the effort through which it is realized is what permits a social surplus to develop, and with it - together with other conditions, in particular the transmission of learned behaviour - luxury and invention.

But memes adapt intentions not only by expanding across more people the efforts they bring into play, but also by shaping expectations. "Ascetic" memes {Footnote 81}, or the restriction of expectations and the schooling in "deferred gratification" that goes on in the socialisation of every child, evolve as well as luxurious ones - and these stimulate their own forms of invention. These domestications of intention are also the work of memes - and memes which also often require of their human carriers/vehicles considerable effort in order to limit the consumer's calls on the efforts of others. The relation between asceticism - or luxury - for one person and the burden on others is often complex. As Gandhi was reminded by a follower, it cost a great deal of money to keep him in the poverty he had grown accustomed to. {Footnote 82} Conversely, even monarchists get drawn into debating whether the Royal Family provides value for money.

Asceticism of a kind can also be had by expanding the number of memes' consumers proportionately more than the total cost. The twentieth century electronic media do this spectacularly well, concentrating vast social resources on individual templates, which are then replicated at small or negligible marginal cost. Even the most elaborately designed templates, however, are often elaborations on ancient psychic equipment. The appetite for gossip about kin and familiars, which evolved as a key but high-cost - one-to-one - aid to inclusive fitness in our

ancestors' pre-literate (and pre-pecuniary) hunter-gatherer bands (as Barkow, AM, argues), did much to propel the telephonic revolution, and has been incorporated (with crucial modifications) into many of the most successful genres of television - not only chat shows, soaps and sitcoms but also, for example, global news.

Money is a crucial adaptation of such memes, not only in the general sense that the technologies have evolved within moneyed industrial societies, but in the nitty-gritty competition between programs. TV advertising sells on the basis of charges [CPM?]: numbers of viewers weighted by their disposable income.

Global culture is a function not only of money, but of "species typical" adaptations: the fact that humankind's genetic unity entails (as "evolutionary psychology" accurately points out (CT)) compelling unities of anatomical and psychic design. Nowadays "Nikes are worn by cannibals." [Check ref Twitchell]

[Section on trashing of taste, mass media and unit cost of hominid gossip (Barkow). Cost per thousand dollars (weighted sum): money as regulator of the most intangibles]

## Work

One obverse of money's meshing together unlike, by connecting their like linkages, is an omnium gatherum notion of "work" as something detachable from the experience of the human animal who performs it. In reality the division of labour is a separation of lives:

"There is no greater modern illusion, even fraud, than the use of the single term work to cover what for some, as noted, is dreary, painful or socially demeaning and what for others is enjoyable, socially reputable and economically rewarding. Those who spend pleasant, well-compensated days say with emphasis that they have been "hard at work," thereby suppressing the notion that they are a favoured class. They are of course allowed to say that they enjoy their work, but it is presumed that such enjoyment is shared by any good worker. In a brief moment of truth, we speak, when sentencing criminals, of years at "hard labour". Otherwise we place a common gloss over what is agreeable and what, to a greater or lesser extent, is endured or suffered." (Galbraith, 1992, p 33)

Socialism shares the difference and consequent fiction. Haraszti (1977) came to think of workers and their managers in the "workers' state" as distinct species, partly co-inhabiting common space {Footnote 83}. From time to time social science fiction has projected a future in which the reproductive isolation of classes becomes (or is made) absolute and speciation occurs. {Footnote 84} But the differentiations could also, and perhaps better, be expressed in terms of differences among the species' of memes by which humans are domesticated. {Footnote 85}

The related point that the social organism becomes complex at the expense of individualism has often been made. Here is Mikhailovsky's essay on progress in [187?], attacking the paeans to industrialization of Herbert Spencer and the social Darwinists:

"If society makes a transition from homogeneity to heterogeneity, then the process of integration in the citizens which corresponds to this transition must proceed from heterogeneity to homogeneity. In short, individual progress and social evolution (on the model of organic evolution) are mutually exclusive, just as the evolution of organs and the evolution of the whole organism are mutually exclusive....In an organism it is the whole that experiences pain and pleasure, not the parts; in society it is the parts that experience pain and pleasure, not the whole" (quoted in Edie, Scanlon and Zeldin, II, 180-1)

## Law

Law and property, as systematic means for modifying and redirecting the efforts of others, precede and underpin the emergence of money. Legal norms endure by replication. The origins of "justice" in the sense of fairness and impartiality lie less in moral considerations than in the evolution of legal and quasi-legal memes for the more efficient administration and operation of law. The operation of law goes far wider than the tiny minority of situations in which it is explicitly invoked or enforced; it regulates and circumscribes much of everyday life because legal memes evolve in such a way that they easily enter into the habits and motivation of most of the population, even where they do not crystallize as explicit awareness. The law acts as both efficient and final cause of law-governed behaviour, not so much because subjects select among their intentions to eliminate unlawful impulses, but because their intentions are formed already "engraved" by the extant legal memes. An important element in the design of legal memes is the jurisprudential principle that "A law is best the less it must be enforced", and the fact that temporary or local lawlessness (the breakdown of law and order) is identified with situations in which coercion comes into play as a means of regulating behaviour.

Modern legal systems have evolved to the point that they are administered by highly trained, socially and occupationally segregated, legal professions and judiciaries, themselves complexly internally differentiated. In reality ignorance of the detail of the law is the almost universal condition, an ocean of vagueness, dotted with human islands of professional competence. These islands perform the social function of "radiating" awareness of the legal structure through the population as a whole. For the structure of legal memes to act effectively and continue it must generate - or more exactly, perhaps, emanate - popular quasi-legal memes which can readily enter into the habits of the population as a whole. The legal fictions that knowledge of the law is general or that ignorance of the law is no excuse - is nowadays essential to the integrity of legal systems. It is reflected in other general, but more practical, maxims: that minor or unintentional breaches of the law are not subject to sanction, that legal reform and administration should aim at commonsense consistency etc. And it elicits specialised adaptations - compliance officers, legal departments etc - in complex legal personalities. Just as language acts not through particular points of application or by the *seriatim* utterance of words, but through providing a structure within which meanings are transmitted, law operates not through particular episodes of enforcement, but through the general dispositions within which law-regarding habits and intentions crystallize {Footnote 86}.

But awareness of what? The simple image of law as objective, or at least as texts on tablets, dissolves on closer examination. Law which is stated in broad, popular but definite terms on one skin of the onion is, seen in more detail, an effect of memetic consensus among specialists. Often

enough particular cases discover fissures in the consensus at one level, which must then be resolved by reconstruction into more detailed consensus, new memes embrained, so to say, in the legal profession or judiciary or particular segments of it. The common law notion that judges, aided by lawyers, discover a pre-existing law is recognized by those engaged in it as the necessary appearance it is: an appearance transmitted and made necessary by the requirement that laws survive by propagating themselves in a form securing broad compliance - a compliance which is itself a condition of law being nourished by the economic system.

In its developed form law, like money, acts as a permanently adapting and structuring framework of intentions shaping other intentions, acting through a continual flux of memes. Money and law each embed themselves in social relations as an apparent structure of transmitted values and rules. In each case the individual feels her or himself to be, and as a consequence really is, directed by institutions experienced as "things" (but with intentions). More closely examined, though, these things are relations, transmissions, interchanges among people - relations which persist and replicate themselves, both severally and as the complex wholes they build up into.

Their persistence and replication are strongly affected by natural selection, and consequently tend to evolve in their survival and propagation functions. They become, in the perceptions of individual humans - which is one of the most important things with cultural transmission - "reified", appearing as encrusted and coercive to the individual, human relationships transformed into things. And the thing for one individual is formed of the human relations of others.

{Footnote 87}

Law and legal memes are good examples of how natural individual selection among memes can produce effects - behaviour coordinating genetically unrelated individuals - which would require group selection to arise biologically. The legal meme with the greater capacity to replicate itself - principally by eliciting compliance - across the human group and over time is the one that tends to get selected.

## Crowds

In moving from money to law we shifted from a type of complex meme with a more-or-less tangible expression to a type expressed in less tangible social institutions - a change, so to speak, in the staining and magnification employed for our sociological microscope. Let us continue to vary the optics. Consider the performances of individuals in two moving but intersecting crowds of rush-hour commuters as they cross the concourse of a rail terminal. Each individual is oriented towards their next turning point - the exit, the kiosk, or whatever - and walks towards it at a speed whose elastic upper and lower limits are set by average capacities.

Yet collisions are far fewer than for two streams of particles on similar routes at similar velocities without forces of mutual repulsion acting at a distance. Between the individuals composing the crowds there take place continual, rapid, unreflecting - in a sense microscopic - identification, classification and tracking of others, and an estimation and reassessment of their intentions, integrated with the subjects own continual adjustment and readjustment of her own pace, posture, gaze and intention. Each individual can be considered as radiating a continuous stream of memes, and as continually selecting and processing memes from others. This generates

interactions (of which the simplest is repulsion) between individuals and their intentions which are negligible at separations of more than a few meters, but are a markedly inverse function of distance at the lower ranges, so that near contact evokes a sort of reflex of recoil. Crowd behaviour as shaped by this meme-flux contains symbolic ingredients that differentiate it from herd behaviour. The moving, calculating commuter is continually recognising others according to symbolic and normative as well as physical-biological criteria. The symbolic criteria which come into play are those of the culture - of which the crowd forms a momentary agglomeration and concentration.

## Norms

Dress and other accoutrements are among the most important of these. A currently-recognized uniform is more-or-less essential to interrupt or deflect the commuter stream. Even in Rome an individual dressed as a centurion is presumed to be on his way to a fancy dress party. One good measure of women's' inequality with men is the effort a female uniform wearer must project to successfully steer the crowd. On the transit systems of San Francisco or Paris the white stick carried by the person with impaired sight serves (like a green light for her, and a red for others) as an instant symbolic signifier to leave additional margins of space and time, and causes others to adjust very rapidly the image they hold of the stick-wielder's visual field, state of knowledge, and intentions. The effect is both familiar and symbolic or cultural - the white stick is an effective modifier of crowd behaviour in New York, but is unintelligible among crowds where a different symbolism exists, or where people with seriously impaired sight seldom go, or go unescorted, in public, as in much of the Third World.

The effect is also familiar in the sense of being the almost effortless consequence of recognising a symbol, rather than the outcome of any chain of deduction from particular circumstances of the situation. In deductive logic thick glasses should evoke the same sort of response as the white stick, only perhaps less marked. In reality the meme flux has evolved rapid-symbolling routines through which wearers of glasses alone are assigned to the category normal, and those with white sticks are disabled. Disability is expressed symbolically in the white stick, or the exaggerated, "warding away" gestures used to usher a wheelchair user through the crowd.

The evidence of impairment or the appliance to compensate for it has become a symbol - a radiator of stereotypical memes to others. One advantage of the stereotyping is economy of effort - rapid recognition and avoidance. Receivers of such stereotypical memes are rewarded by a direct gain in symbolic status. The white stick affirms the sighted avoider as normal and at the same time gives him an opportunity - more, or less, ceremonious or fleeting, seized to an extent that time and personality determine - for a demonstration of status and chivalry in face of one of guaranteed lower status. {Footnote 88}

Earlier symbols of this sort had such functions even more explicitly. The tolling of the lepers bell, the beggars' stumps at the cathedral door, served to highlight one's own wholeness and salvation. Alms cemented the reassurance; the monetization of charity, as excoriated by Luther, grew from such everyday exchanges of status. Without the symbol the norm would scarcely exist. In Hitler Germany's the Star of David, or the pink triangle of homosexuality, were memes of warning, but also of reassurance. {Footnote 89} As with the segregation of women in the

mosque or synagogue, differences or supposed differences of biology are crystallized into memetic reaffirmations of humanity. But where the "minority" is a smaller proportion the psychological effect, to be general, must be correspondingly more intense. {Footnote 90}

In the anonymous, factory-dressed commuter crowd guarantees of relative status are all-too-scarce. The hunter-gatherer's anxious monitoring of self-vis-a-vis others for relativities in dominance flicks over a torrent of unintelligible, momentary others, negotiating a route through society like a smoke-signaler in the fog. Much of the crowd's suppressed tension arises from the fact that each of millions of individual subjects is simultaneously thirsting for sociability and dominance.

General clues - of age and personality, education and affluence, gender and motivation - exist, with their signals of inquiry and empathy, assertion and deference, and enter into the close-range interactions in mass society's crowds, generating the close-range "strong" forces. Up close the forces of repulsion among citizen-strangers are strong. One of the things the human mind is domesticated to do is negotiate the individual's route through a bombardment of normative claims, combining minute deftnesses of interaction with fine judgments of clearances and tolerances. In the crowd others are all potential animals or machines, animated but unfamiliar, in a potential human-jam, and uncertainty makes it necessary for psychology - and thence action - to expand the clearances. {Footnote 91}

The human crowd differs from common types of animal herd in other domesticated species, such as sheep, where physical contact and ricochets do not raise stress levels so much. {Footnote 92} The strength of close-up forces once human body-spaces overlap {Footnote 93} can be judged by the violence of the measures sometimes necessary to overcome them and pack the commuting atoms more densely. The Tokyo subway has special employees whose job it is to pack more people into the cars, squeezing bodies up against each other - and raising mute meme-exchange and repulsion to extraordinary intensities. In medieval towns, grown up round foot transit, those of high social status travelled with a retinue - "retainers" as much as body-guards - whose function was to buffer the protected individual against the continuous jostling of ordinary members of the crowds. The forces brought into play when crowds are obstructed soon overcome lightweight and hence symbolic inanimate barriers; the football crowd or demonstration must be held back with special retainers, whose powers of repulsion are amplified by special memes - the uniforms, drill and training of bodies of police.

## **Structuring the crowd: uniforms**

The adoption of a uniform is a good example of memes achieving the group selection effect that genes cannot. Uniform memes emerge, like multi-celled organisms, only when simultaneously replicated on a fair scale - where more casual signals of status and function prove insufficient to maintain a coherent, lawful meme-flux ordering a crowd. {Footnote 94} Consider the differences between (a) two counter moving crowds of commuters flowing face to face into each other, and (b) two hostile battle formations of infantry engaging in hand-to-hand combat - the differences not only of overt behaviour but of moment-to-moment psychological adjustment. The multiple, more accidental attire which serves to orient the individual in the first case is insufficient in the

second. A basic memetic adaptation of military action until quite recently was the bold stereotyping of the other as enemy or friend, object of lethal assault or potentially fatal altruism. A bayonet is a memetic weapon. When the forces of uniforms and patriotism grow weak it no longer functions to order, becomes just "a weapon with a worker at each end".

Animals other than humans are innocent of war (as distinct from, say, group predation or mobbing), not just because they lack the technological memes but because they have insufficient differentiation and recognition of individuals, incapable of sustaining command - one centre, coordinating others against yet others. Command is also stereotyped for easy recognition and compliance: the officer's insignia; the regimental colours. In war officers' bright insignia designed for recognition by comrades of other ranks may - like the peacock's tail - render the wearer especially vulnerable (because he is recognisable by the enemy). In the trench warfare and massed assaults of the First World War officers' uniforms made them particular targets of small arms fire, and over time battle dress differentiation was reduced to the minimum which would enable recognition by one's own side at close range.

Crowds gather at different places for different purposes, practical and ritual. The same individuals who will patiently tolerate a mother weighed down with shopping and toddlers, or an old person fumbling with packets and change in a supermarket queue, would elbow the same individual aside if she attempted the quite different tempi of the city rush-hour. The promiscuous body contact of the Calcutta market place or the pilgrim horde at Lourdes would be downright offensive on the pavement of the Via Condotti or the Rue St Honore. Individual social clues are also important: indicators of higher socio-economic status - the Harrods carrier bag rather than one from Tesco - may purchase space through amplified avoidance. {Footnote 95}

## Formal organisations

The anonymous, unstructured crowd of commuters' functions and flows by the continuous person-to-person transmission of memes. It is also a (less immediate) result of replication among many other types of meme: those of factory or bureaucracy; of suburbanization; of electrified means of communication and transport; of large-scale commercial and state organisation; of revolutions in information technology. Many of them involve the routine recognition, classification and use of other humans as predictable stereotypes - in the extreme, quasi- or virtual machines. Human minds (as Dennett has it) are engraved by memes to process other memes. Ease of recognition and use can give enormous replicative advantages to the meme complexes within which they function: the semi-skilled worker as a substitutable component of the assembly line; the post to be filled by the appropriate qualification; the lubricative effect of civic consciousness and civil rights in the crowd of citizen-voters-commuters; the orderly flow of the mass into office or Auschwitz.

There are, of course, also efficiency-penalties to substitution which override individuation, especially where group tasks are complex, have moving goals, or make stressful demands on cooperative skills. The hoplite phalanx was a major innovation in infantry tactics, not only because it fused bodies, shields, armour and weapons into an orderly, relatively safe array, but because it did this on such a scale that most of the young men of the phalanx, drawn from the same village or district, knew each other as individuals and comrades. Attic citizenship,

comradeship and eroticism in the phalanx aimed, by different means, at what the mass military schooling of Spartan education also sought; the integration of individual motives and impulses into machine-like mass action. {Footnote 96}

Formal organisations often find - by evolutionary trial and error - that they cannot dispense with the need for intimacy. At the tactical level modern military organization is a complex compromise between the dictates of equipment, unity of command, and training, battle hardening - and loyalty on the approximate scale of the hunter-gatherer group. The bureaucratic battle group cannot fully emancipate itself from its ancestry in the phalanx or posse. Troops may risk all for comrades where they would not put their lives on the line for commander, nation or fellow-citizens in the abstract. {Footnote 97} Similar considerations of scale and familiarity apply to the designed or traditional work crew, the surgical team, the school class, or the invisible college of academic specialisation. {Footnote 98} The family frequently collapses as the site of child care and socialisation, but when it does the state, saddled with infant citizens, can often do no better than reconstruct it.

Modern complex organisations are frequently built round explicit statements of intentions, mapped into organizational structures. Intentional change can then occur (in part) by linked modifications of goals and structures. Or it may occur - as our *homoousion* example (above, p ) suggested - in more local, inexplicit ways. In all cases where intentions are transmitted between people unintended consequences of social action arise - it is just a matter of their nature and extent.

## The case for memes

What do our examples suggest? Let me briefly make the case for the "meme's-eye view". Psychological and cultural phenomena form - that is what monism means - part of a seamless universe. As such, they can be understood from various points of view. What is most effective depends on what specific phenomena we want to understand and, equally important, on the means of understanding available to us from meme-evolution (history) up to date. To take examples from our present historical location: we can view the human world from the standpoint of physics and chemistry (for example, in dentistry or pharmacology); we can see it as one specialised type of molecular-patterning replication (as in biology); we can think of culture as the interactions of individual human organisms (as in anthropology or sociology). Or, I'm suggesting, we can think in terms of the natural history and evolution of memes, chiefly of meme-forms specialising in the direct and indirect "domestication" of human beings.

These standpoints do not exclude each other. They require occupational differentiation among humans, and they are thus, to a degree, separated from each other, but the barriers are neither waterproof nor fixed. On the contrary, to regard the universe as seamless means that among the tests applied in meme-selection is the requirement (itself evolved as a complex meme) for logical consistency between the memes employed by different standpoints.

Looking at the world in meme terms is not new. History does it insofar as it seeks to be explanatory, or populates its narrative with ages, classes, styles etc. But its treatment has often been casual, anecdotal - or, worse, narrowly human-focused and pragmatic, failing to see the

memes for the people. The second part of my case for a (more systematic) meme's eye view is the claim that the study of memes can be effectively developed if it borrows models and insights from the existing natural sciences, particularly from evolutionary biology.

When we try to understand why the commuter crowd makes space round the man with the white stick, why the audience rises for the national anthem but weeps for the aria, why wife and husband, corporation and customer, TV and viewer, trick and treat each other as they do, why the bull and the Bolshevik see red so differently, or how rival personal computers compete, thinking in terms of the natural history and evolution of memes can be useful. This is not to deny that humans (or their sticks) consist of agglomerations of atoms and tissues, that they grow under the control of nucleic acids, or that speech and music accord with the physics of vibrating bodies, but just to suggest that a meme's eye view may have a distinctive purchase on these things.

## Meme memes

This brings me to the third component of my case for memes, which is that of relative advantage. When we argue about ideas and theories one of the things we are doing is selective breeding of memes. And in defence of Dawkins' meme memes it can be said, at the very least, that their immediate competitors - the social and psychological sciences - have not recently shown anything like the pace of successful adaptation of natural science memes. There is, so to speak, a niche waiting to be inhabited. Insofar as explanatory power is an important criterion of selection, the environment for meme-theory looks a favourable and under populated one.

One objection to placing too many hopes in meme-theory may be the underdevelopment of psychology. We are still far short of understanding the macro- and micro-mechanics by which memes "parasitise" humans. But we can draw some comfort from the history of biology. Darwin, Wallace and their early followers were able to sketch an essentially accurate theory of natural selection even though they did not understand the physical machinery of heredity and variation, were unaware of Mendel's genetics, and were inclined to the false view that some naturally acquired characteristics could be inherited. They were able to develop their theory by drawing on what was available - evidence from fossils and geology, work on biological classification and the study of species' distributions, embryology, and the experience of commercial breeders and growers. Today, the study of human culture and psychology is not really short on empirical evidence; the problem is to locate it within unifying principles. Unfortunately meme-evolution has equipped us with powerful inhibitions against doing this in its differentiation of academic and scientific specialisation.

## One universe

Our universe can be thought of as selected selective processes favouring patterns that endure. {Footnote 99} In terrestrial conditions "elementary" wave-particles combine in various configurations - the more stable of which form nuclei. Among nuclei, some few dozens are stable enough that (after much meme-evolution) we now identify them as isotopes of distinct elements. The longer-lived among them - together with their longer-lived fission-products - form the atomic "stuff" of our world. Their electron clouds form - according to their total charge - more or less stable configurations. The more stable are inert elements; the less stable are prone to

bind to others as compound molecules. At low momenta many molecules enter into crystal lattices; several possess more than one stable lattice pattern, and which crystal structure spreads may depend (among other things) on which is there to be replicated in the first place {Footnote 100}.

[Can Dawkins' (in *The Extended Phenotype*) be heard as saying that matter is simply the occurrence of successful forms? See as quoted in Smith, *Genes, Sex and Evolution*, p 106 ff] Notice that natural selection does not require replication - at least, not in any strict or narrow sense. Water's action on granite separates the quartz grains from feldspar and mica, preserves and transports them, and organises them, finely graded, on beaches (Van Valen 1989).

Carbon and oxygen, though, bind with some other elements in a large variety of extended molecules. Among these, some have the particular property, in a suitable chemical environment - such as exists within a cell, for example - of being able to serve as templates for the repeated assemblage of complex chains from shorter chains free in a water solution around them; their synthesis-sequences, acting together, both feed upon and generate the intracellular food in which they live. One in particular - ribonucleic acid - has the striking property of serving as its own template. Within cell nuclei it forms double chains, parallel spirals, with two types of asymmetric "linkages", each of which has two "sides" and consequently two orientations - a total of four possible types of link making up chains of indefinite length and complexity. In cell division the double spiral peels apart, each side immediately replicating the double chain by binding to free bases in the soup.

While all cells employ this basic chemical machinery for replication, the life cycles of daughter cells also depend on the cell's environment. Cells inherit characteristics acquired from the environment; the developmental programs of a genome provide for cells to produce offspring whose traits are determined by the parent's acquired characteristics - that is, for "cultural" transmission between cells to affect cell replication.

Some cells restrict their environment in their favour by cohabiting as groupings of several cells. Some among these further improve things by evolving patterns of specialising differentiation among themselves - organisms. Their coherence and structure derive from their cells' common ancestry and the presence of an identical development program laid down by the precise pattern of DNA replicated throughout each organism's cell nuclei. Because this provides the "atomic" basis of organisms' coherence, replicated in this infinitely repeatable way, all organisms' life cycles, no matter how differentiated and elaborate, involve a single-celled reproductive stage, which has to contain the overall development program for the organism.

The sub-program controlling the cell's own division and replication, however, must be rather sensitively attuned to the mother cell's environment and history if it is to produce a variety of appropriate types of daughter cells at the right times. "Copying errors" occur. In most cases these are lethal to the daughter cell, and excretion disposes of her remains. Occasionally, though, the copying error produces replicatively viable daughters, or even cells with a replicative advantage relative to normal cells, such as those which form tumors. Sometimes errors arise at the single-cell reproductive stage, and very occasionally these are neither lethal nor disadvantageous for the offspring, but advantageous.

We can think of these replication processes in terms of DNA sequences or of "genes" - which interconnect but don't necessarily correspond to contiguous sections of DNA. Although genes are notional in a sense that contiguous DNA sections are not (or less so) biologists find that thinking in gene terms can have considerable advantages. {Footnote 101}

## Memes and the evolution of psyches

"We keep passing unseen through little moments of other people's lives" - Robert Pirsig, Zen and the Art of Motorcycle Maintenance.

Species have distinctive behaviours as well as structural adaptations. Many "extended phenotypes" (Dawkins 1982) have behaviours which use other species and/or external materials external structures as functional parts of their life cycles: parasitism/predation, nests, burrows, caches, colonies. Many have life cycles incorporating substantial "parental (kin, sibling) investment" in their offspring beyond their single cell reproductive stage - transferring resources to the offspring's physical structures or environment, or to form both structures and behaviours. {Footnote 102} In humans artefact making and investment in behaviour combine to produce particularly elaborate results. Lengthy childhood dependency develops humans ready to replicate other humans' intentions in extremely flexible and therefore powerful ways.

Memes and humans use other humans as tools, and these walking, talking "tools" themselves use other human tools, chaining them together (through language, shared artefacts, rituals, law etc) in sequences of connected intentions. {Footnote 103} Division of the meme's labour doesn't require this particular human tool, just one suitably adapted. Theatre, for example, turns on the tension between this and the irreducibility of the tool's own experience: mistaken identity with all its comic possibilities; or a human consumed in tragedy's realisation.

We can think of chainings of intentions in terms of human psychologies and of humans' minds and actions. Or we can think of them as phenomena in their own right - memes. Memes grow and replicate through enculturing, influencing and organising humans, somewhat as genomes do through growing proteins and nucleic acids. The core of the case for the meme's eye view is pragmatic: that it can help us understand culture better. We can think of the white stick, or the aria, in terms of the interactions of molecules, or of the interplay of human individuals. But it can be less cumbersome to think in terms of concepts more tailor made for the purpose. When evolutionary biologists argue through their ideas on aging, seed dispersion, altruism or immune systems in terms of genes none of them dispute that these things could in principle be described in molecular terms. But between the "in principle" and today's specialised humans, organised as they historically are with the artefacts they possess, there is a large gap. The idea of a gene helps reduce it a little - so meme evolution tends to adopt it. An important element of the case for memes about memes is that they can do likewise with culture.

"It is natural for geneticists and evolutionary biologists to hope that their disciplines will throw new light on the human condition, and equally natural for social scientists to resist the threatened takeover" - John Maynard Smith, 1988, p86.

Thinking in terms of memes can complement "evolutionary psychology": the study of human psychological adaptations incorporated in our species in its genetic past. The success of a popular food (or a patent cold cure) may be understood as the outcome of our ancestors' liking for salt or sweetness at a time when these were rarities and hominids' reproductive success was affected by preferences for them. It may be understood through the neuropharmacology of our responses to sugars or salt. But it may also be understood as the evolution of families of foodstuff-memes: domesticable food-grains, or beef burgers (involving the co-domestication of cows, consumers and coins). Or it may be understood in terms of memes linked to the "reproductive" cycles of the foodstuff-memes: advertising, or factory farming. Or (as Dawkins argues in "the dissolution of the organism") it may be seen as part of larger self-replicative cycles by infective organisms. {Footnote 104} The explanations each have their own limits, and they overlap. The important principle is that they should be consistent with each other or - apart from anything else - fuller integration will be impossible.

Thinking in terms of memes can help repair - or, at least, supplement - one of "evolutionary psychology's limitations. This is its tendency to work with a model of human "psychological architecture" in which the most "basic" structural components are humanly universal, "species typical" psychological adaptations to a Pleistocene environment of evolutionary adaptedness. Culturally specific phenomena - language, diet, dress - tend to be explained only in part. Thinking in terms of memes can supplement "evolutionary psychology".

Historically, "evolutionary psychology" has evolved in competition with that important current in nineteenth and twentieth century thought which asserts that cultural phenomena can only have cultural explanations. {Footnote 105} The anthropologist Clifford Geertz, for example, seems almost to de-psychologise individual thought:

"Human thought is basically both social and public - ... its natural habitat is the house yard, the marketplace and the town square. Thinking consists not of "happenings in the head" (though happenings there and elsewhere are necessary for it to occur) but of a traffic in what has been called.... Significant symbols, words for the most part..." (Geertz 1973, p 45) {Footnote 106}.

Others have gone further, arguing that the specifics of human biotic evolution have no relevance to understanding the evolution of cultural entities. {Footnote 107} "Evolutionary psychology" castigates any such "Standard Social Science Model" (SSSM) (at least insofar as it segregates psychology from human evolution) and rejects the idea that newborn human psyches are infinitely malleable by their cultural upbringing. It seeks to restore evolved human psychology as the missing link between natural selection and social action.

Much of anthropology is indeed, as evolutionary psychologists complain, dismissive of psychological explanation. Marvin Harris' Cultural Materialism (a common textbook) roundly declares that "the human intuition concerning the priority of thought over behaviour is worth just about as much as the human intuition that the earth is flat" (Harris, 1979, 60). But there is also truth in Geertz' point. Social scientists could certainly use more effective tools to understand the interpersonal "traffic in significant symbols".

"Evolutionary psychologists'" criticisms of the SSSM and their efforts to reinstate psychology converge with some lines of critique from within social theory. Moscovici (1993), for example, writes of sociology's inhibitions against "psychologism" as symptomatic:

"The segregation of the psychological from the social has become institutionalized in our culture. Though independent of any critical reason, this segregation resists every kind of criticism. Those who risk calling this into question come up against censorship and, to begin with, censorship of themselves" (Moscovici, 1993, p11) {Footnote 108}

The way forward, in Moscovici's opinion, lies in a double movement: to investigate the nature and sources of this disease-taboo, and to rescue from it, and hence develop, the flawed insights of social science. Like natural selection, he starts from where it's at, and tries out variants: on Durkheim's image of religion, Weber's views on Protestantism and charisma, and (his particular favourite) Simmel's account of money. {Footnote 109}

There is, I think, some parallel between the individual's Cheshire-cat like tendency to dissolve among practices and values in social science, and the difficulty in biology of pinning the organism down - of separating it out definitely enough from its environment to take it "for granted". Just as one cannot fully visualise a spider bereft of its web, one cannot visualise an individual human, island-like, in abstraction from the values and activities of the society s/he forms part of.

## **Social order and social surplus**

Having hazarded such large claims for a meme's eye view I should also suggest ways in which they might - one day, perhaps - be realized. I end, therefore, by suggesting a few ways that memes might help with a central problem-area of social theory - understanding social order and social surplus.

This can be approached from various directions. The possibility of viewing the social surplus as a whole - and consequently as accessible to alternative uses, or redistribution - is a relatively modern one. It has come about with the monetization of society - that is, the predominance in social life of a single numeriare - money - which invites and demands humans to compare things in a purely quantitative sense: between things, between people, between societies, between periods in time.

Thus the problem of a social surplus - or of the net social product - crops up in a variety of ways in modern political economy. {Footnote 110} We can spy the problem in paradoxes of measurement and definition. {Footnote 111} The relative scale and growth of national economies are compared using measures of national income: gross national product (or income) per head, computed from information about aggregate money flows. But what counts as economic production and what counts (i.e. does not get counted) as nature's bounty is to a degree arbitrary. And the links between national income and human welfare are complex.

For example: Thais are poorer than Britons. By how much must Thai national income per head grow for them to "catch up" with Britain? National income statistics for Britain include a

substantial element for domestic heating. This increases when the winter is cold or the price of oil rises. Heating expenditures for Thailand are much lower. How far should Thai national income statistics be adjusted upwards to allow for its more agreeable climate?

Or: take state expenditure on pensions in developed economies. Where, previously, old people would have lived their closing years as dependents in an extended family, the state now taxes workers and pays non workers, who in turn spend independently, consequently leaving their trace in national income statistics. The circuits of money are amplified, incorporating what was previously hidden in the family. How far is the redistribution brought about by pensions also an expansion of income?

If, in addition, old people live longer, the total national income must be distributed among more recipients, providing less per head. The notional equality of citizenship injects population increase with similar paradoxes: shall "we" be happier because there are more of us, or because we are older and wiser, or the reverse? What if the price of more protracted isolation and poverty in retirement is more monotonous working lives?

Or take the fast food revolution. Instead of obtaining noodles, water, vegetables etc for a few pence, and cooking them at home, families now get takeaway meals (or have pizzas delivered) at much greater cost. Time thus saved is devoted (in part) to watching television; as a consequence family dining tables and family meals have disappeared in many homes. How do such changes show up in national income statistics? {Footnote 112}

In fact, like biologists, economists took some time to settle down with a model of the economy as a whole seen as the circulation and replication of one common substance: money. So-called neoclassical economic theory took shape in the late nineteenth century from an older tradition which sought to analyse economic life as interactions among qualitatively disparate production factors. {Footnote 113} In contrast, for neoclassical economics, the same basic principle of adjustment governs the markets for all particular commodities (pins, cotton, labour, bank debts, and reputations). {Footnote 114} Price varies negatively with supply (the commodity's price rises as supply falls), but positively with demand (the commodity's price falls as demand falls). The adjustment of prices, and consequently of the patterns of circulation of money, are the work of a "hidden hand" continually pursuing equilibrium.

Neoclassical economics applies the same picture to economies as a whole: aggregate values of "labour" and "capital" combine to produce "output". If the amount of capital per worker increases the rate of profit falls. Profitability serves as the "price" of capital, adjusting its quantity until the average entrepreneur feels no impulse either to invest or to dispose of capital. And - this is the beauty of the model - in doing so, and by the same token, the adjustment of the rate of profit determines distribution: how - in the simplest models - national income is distributed between profits and wages. {Footnote 115}

Think back to the baked bean cannery. The common numeraire, money (the economists' nucleic acid) is essential to this model - not only in theory, but in reality. It is the only thing that allows us - workers, businesspersons, consumers - to calculate aggregate amounts of wage, capital, profit etc. For in reality the capital of an economy - indeed, often that of a single plant - consists of a mixture of equipment doing different things, by differing technical means, worn out to

varying degrees by use and by age, and maintained in varying states of repair and partial renewal. And all that equipment is the product of past activity, using both similar and different equipment, and so on ad infinitum. Some common numeraire is necessary to add the variety together across heterogeneity and time; without it they remain incommensurable.

## Aggregate economic production

Describing the freezing-fossilizations effects of price in production economists are easily driven to metaphor. Trying to capture capital's combination of malleability and fixity they have spoken of "fossils" - and also of butter, toffee, putty, clay, meccano pieces, and so on. What the metaphors have in common is the attempt to simplify, and thus capture, relations across sectors, across space and - especially - across time, the sense in which exchangeable resources carry ("fossilized") within them the history of past resource use, and the part played by past technologies and prices in setting the price of new-made artefacts. {Footnote 116}

Relative rates of profit - that is, relative rates of replication of capital - act to extinguish technically superseded capital and replace it with - selected - capital that incorporates more efficient innovations. That is to say, the rate of profit plays, in economic life, something of the part of "fitness" in population biology.

## Paradox

There is a paradox that can be derived from the technical analyses. On the premise that capital is homogenous, the rate of profit may be a rising function of the aggregate value of capital. {Footnote 117} If an increased quantity of capital is associated with an increased rate of profit, what has become of the mechanism through which profit served as the "price" of capital, rationing its supply? {Footnote 118} The treatment of capital as homogenous is crucial to the paradox. It involves treating all particular capitals, different as they are in their technical natures, their ages, and the degree to which they are worn out, as though they were commensurable in a common numeraire. As time goes by, individual items of capital equipment (say, machine tools) change their relative values. They do so for mixtures of reasons. They wear out with use and with age. The prices of the things they are used to make (pistons), or could be used to make (artillery shells), changes (sometimes rapidly). Technical progress occurs, so that even unused machine tools which were made in the past are devalued relative to up-to-date ones. The average amount of capital used (or the total number of machine tools) per worker in the economy changes - in the course, for example, of industrialization.

Distribution affects prices. The proportion of profits going into savings rather than consumption is, in general, higher than the proportion of wages. Thus an increase in real wages - for example, due to strikes, or elections - can change the overall distribution of income between consumption and savings and thence - through a variety of routes - the relative values of different stocks of capital.

The common numeraire serves, through adjustments of prices, to homogenize all these differences into an aggregate value of capital. The paradox of "double switching" (the fact that an increase in the aggregate value of capital may be consistent with an increase in its rate of

profit) can arise because changes in one segment of the capital stock can affect values in other segments.

## Organic life and memetic life

In organic life, however, segregation sets tighter limits. Finite organisms, reproduction, speciation, and the isolation of germ from soma constrain gene flow into distinct channels, which interact through relative rates of population growth. For much population biology measures of quantity (population) and increase of quantity over time (population growth) are (1) conveniently provided by the counting of organisms and (2) subdivided by speciation into distinct gene pools. Speciation largely spares population biology the paradoxes encountered in economics (which may, perhaps, be defined as that part of social science from which organisms have been most thoroughly removed). Repproductively separated populations of distinct organisms leave population numbers and growth rates - measurable essentially by counting - as the primary objects of analysis. For many purposes the analysis of marginal variations in rates of population change provides a framework which is free of paradox. {Footnote 119} {Footnote 120} In certain senses organic life does produce an overall "surplus". What emerges over evolutionary time is an increasing number of organisms and species, of increasing complexity, and themselves organised in increasingly diverse and complex ecologies. It is hard to state this point so that nature's increasing roundaboutness could be compared in any rigorous (read, purely formal or quantitative) manner across time or lineage (though it is reflected in an interesting literary tradition on nature's unfunctional, irrational bounty, as well as in our scientific knowledge of the history of the biosphere to date). This difficult of stating the process is the obverse of the absence of any universal numeraire - like money or value in economic life - which would allow us to compute a net product or surplus across all the diversity.

The difficulties are not only literary. Biology is concerned with maximization of fitness. But what exactly gets maximised, by what, and over what time scale? Maximising population is not the same as maximising population increase. Both are different from maximising the improbability of extinction. For some life forms the definition of individual organisms, and the borderline separating growth from reproduction, is less distinct than for others. Living longer and getting more massive may be relatively more advantageous than producing more members of the next generation. Speciation emerges from reproductive separation. {Footnote 121} As Van Valen (1989, p3) points out, if there is a consensus definition of fitness in population genetics, it is something like the *expected "relative number of individual offspring in the next generation of the population"*. But, as he goes on to show, all eight of the italicized terms are problematic. {Footnote 122}

If the reproduction of individuals is taken as central perhaps the closest analogues for organisms are found in self-reproducing cultures. But memes transfer far more freely than genes, and, consequently, if cultures resemble organisms, they are organisms among which species are extremely fuzzy, hybridization is rampant, and one can make no clear separation between nutrition and sex.

Or, to put the matter in social terms, the very free interchange of memes between cultures means that the evolution of cultures is a process in which latecomers can frequently and easily reap the advantages of backwardness.

## **Wants and needs, variations and adaptations**

A similar parallel and limitations emerge from the distinction between needs and wants, conditions and preferences. A necessary - not sufficient - condition for us to perceive a social system as producing a net surplus is that it has production or activity over and above that involved in its own simple reproduction. The existence of a surplus is what brings the potential for dynamism into the system; without it it would continue as it is. {Footnote 123} But a fundamental problem in social science is that phenomena that emerge, as variants, in response to wants - forms of decoration, exchanges of gifts, conspicuous consumption, ritual performances, potlatches and puberty rites - then get selected as adaptations, embodied in the people and embodied in the culture. They get promoted from wants to needs, from preferences or tastes occasionally indulged, to parts of the social arrangements which must be systematically reproduced. And they then, in later evolution, come to form part of structures on which further adaptations become - so to speak, by further elongation of the embryo - embodied in their turn. {Footnote 124} If the rituals of social order, for example, are omitted or collapse in modern industrialized societies, the society passes into non-stable or non-reproducing conditions - revolution.

There are, I want to suggest, real advantages in seeing the embodying of preferences as the evolution of meme adaptations, compared with other frameworks, such as social functionalism. It helps us see the limited but real sense in which the social can be treated as *sui generis*. Consider ways in which social institutions - marriage, or money, for example - elicit loyalty or compliance from individuals. Social institutions are memes, but it strains matters somewhat to think of them as being imitated or copied from individual mind (or brain) to individual mind - as Dawkins describes a snatch of tune, or a line of text, as being copied. The point is not whether we are or are not aware of social institutions, but of how we - we, in general, that is - regard them, how we poise ourselves in relation to them.

Successful and enduring social institutions are often ones that appear to us to preexist ourselves as individuals, to exist as things rather than to consist merely in the behaviour of individuals - that is, the institutions succeed by becoming "reified". They may also become successful by exciting our emotions (of fear, shame, attachment etc). One of the ways they may do this is by resembling or mimicking other memes to which we have already developed such responses. Another, closely related, way is by exploiting the already-in-place attitude configurations of typical or common personalities - that is, personalities common within the culture.

As a summary of meme replication, copying or "imitation" is too narrow. It is not that the bachelor or the adulterer have not imitated the marriage meme in their minds, even less that the bank robber's (or the debtor's) brain lacks copies of money memes. Rather the memes in question have failed to elicit the usual compliance in these exceptional cases. And as the exceptions become more common, they exert increased pressure for variants of the social institution which achieve greater compliance to emerge and proliferate (serial monogamy, credit cards). Shifts in

the relative extents to which social institutions command compliance are very often the outcomes of highly complex interactions and adjustments between social institutions: the co-evolution of such memes in the course of social history. The relations between them range from direct competition to dependent symbiosis.

It is not only that the behaviour of individual people constitutes social institutions. People are also socialised so that (most of them) reinforce (most of their) enduring institutions. And people are historically specific. It is not that my image of my marriage or my credit card was or was not present in this or that hunter-gatherer; the possibility just doesn't arise; even to try and visualise it is a "category mistake"; it is incoherent with the rest of her or his personality. And, in turn, the historical (evolutionary) specificity of personality types is formed by the institutions through which they are socialised. Those institutions (family, school, peer group) select and combine personality traits (and some traits of physique) into individuals in restrictive sets of ways; some do end up outside the boundaries of what is "normal", but deviants are the exception, not the rule. And, crucially, socialisation uses the norms as templates of social and psychological reproduction. Deviants crop up in every generation, but by and large they are excluded from or neutralized by the social institutions most directly concerned with social and psychological reproduction.

## **Institutions and personalities**

Social institutions and personalities thus co-evolve. This is, perhaps, what historians of the Annales School are referring to when they argue that history is not only the history of institutions but also that of "mentalities". It also has something in common with Thorsten Veblen's view of social institutions as formed by slowly evolving patterns of habit. {Footnote 125}

We may think, perhaps, of an analogy with the gene pools of two intimately dependent organisms - say, a flower pollinated by a single insect, which pollinates only that flower. Over time the gene pools of the two populations interact, but not directly. The interface (that is, the configuration of the *other* gene pool) changes over time and, at any particular moment, shapes the selection pressures operative in its complementary gene pool.

In fact the matter is more complicated, and it is a metaphor to speak of a gene "pool". Genes are not freely-floating candidates for recombination. Rather possible variants on the genome notionally coexist in a pool of alternative sequences. Certain common structures coded for within genomes and pools (organs and other adaptations, that is features universal to the species) form rules ("dams") restricting the channels in which combinations can form and move. {Footnote 126}

As far as our parallel with social institutions and personalities are concerned, we may think of the interface between the pools as separating the objective social (institutional) world from the subjective (psychological) world of a social order. The precise lie of the membrane is affected by the adaptations on each side of it, and this shaping, in turn, influences reproductive success on each side. The current array of institutions corresponds to currently "normal" types and ranges of personality. Important comparisons of anthropology (and much of utopian thought) revolve around the fact that personality traits and combinations that may be socially integrated within one system of institutions can be ferociously deviant if transplanted into another. All cultures

consist of mixtures of personalities, but the range and the mixture vary enormously from culture to culture. {Footnote 127}

Personality is itself a meme complex, of a particular type - relatively "primitive" in terms of the overall evolution of human-parasitizing and human-domesticating memes. Although the specific personalities of long-dead individuals may leave only very scant traces, we are able to observe remnants of the importance of animist thinking and forms of comprehension in the earlier intellectual and technological history of humankind - that is, generally speaking, in societies where recent meme-evolution has been slower. Today's "scientific", self-conscious separations between the human, the living, and the non-living result from developments which started in animal consciousness. Language allowed generalizations about the world to be exchanged, elaborated, preserved and refined, but these long remained expressed in the vocabulary of human personality.

## **Development of personality**

In fact an individual's personality is not one, but many. It develops over time, in conjunction with physical maturation: Piaget and Freud's work represented attempts to give systematic accounts of, respectively, cognitive and affective aspects of the process. Ontogenetic development combines with the acquisition of features from the social and material environment. Some of these are universal in the culture, or so nearly so that they are the culture-typical analogues of genetic adaptations: language, literacy, infant mutilation, beliefs in enduring personality, the viability of money, or the deity. Entering as they do into the formation of virtually all personalities such features seem to humans to be natural and inevitable features of their world; only from outside the culture can they be perceived as memes which have evolved.

But even at one moment in time individual personality is manifold. People navigate their lives with reference to - indeed by means of - at least three types of personal image: how they see themselves, how others see them, and how they wish others to see them. These overlap but never entirely coincide. The development of social institutions goes hand in hand with that of a generalised Other, personifying the expectations and values embodied in institutions - an Other which elicits normality and gives definition to status. Insofar as status is the outcome of individual action rather than of ascription, the most important, and intimate, selective jostling is not that between corporeal individuals but that which takes place between the self's own self-memes or self-images.

## **Status**

Barkow ( ) describes the outcomes of such processes, from the outside, in terms of status or prestige:

"Our folk wisdom speaks of the young person "discovering" who he or she really is, much as the sculptor may "discover" the figure nascent in the stone. In terms of the present theoretical framework, however, the adolescent is actually seeking to determine which social and categorical physically present and physically absent groups will accept him or her, and whether he or she wishes to be accepted by them... [T]here is a weighing of how high the group or

category itself is, versus how high one is likely to rise within the group or category, a size-of-puddle vs. size-of-frog calculation." (Barkow, Darwin, Sex Status, 198)

In hierarchical, narrative societies memory elongates status as reputation, from the infant gloater's "Naahnaahnahnahnaah!" to the peacock-parade of academic dress. The jostlings between selves, between frogs (or princes), and between meme-pools or -puddles exert various selective pressures: between genes (sociobiology and evolutionary psychology), between people/frogs, and between the categories (sizes and puddles) in terms of which people see themselves and relate themselves to others.

## **Status, surplus, variety**

"What makes equality so difficult is that we only want it with our superiors" - Henri Becque  
Contemporary status goes far beyond local jostling in pools. Baumann defines the scarce resource of postmodern society as public attention. Andy Warhol made much the same point: "In the future, everyone will be famous for fifteen minutes." The point can be re-phrased in our terms, and generalised: for memes to secure attention, they must differentiate themselves from existing memes with which subjects are familiar, they must continually innovate - and they must innovate so as to make an impact relative to the patterns of variety and innovation with which subjects are familiar. {Footnote 128}

There is an obverse effect, arising from the limits of meme's human vehicles: for many things to achieve their quantum of public attention, the attention of most individuals must attend to an increasing variety of things and styles, switching with increasing rapidity among them: postmodernism's well-known fracturing of the psyche.

There are also, however, strong selective pressures on memes favouring rapid recognition. What the attention of the individual's psyche largely gets fractured among are stereotypes: signs, gestures, expressions, styles, motifs, brands, logos, themes, leaders, stars. The popularity of simplified stereotypes is detectable even in areas which make the most strenuous efforts to avoid oversimplification and multiply distinctions - such as academic specialisms. It is felt so widely because it is just one general facet of a pressure which affects all intellectual activity by finite or human beings: the pressure to economize on effort and time, working through differentiation between memes.

The competition among memes for public attention also works through to modify the processes from which they issue. It is a commonplace that life provides materials for art, but the reverse effect is also rather general. Litigation, for example, is a form of spectacle, and this has shaped its development. In Greek city-states amphitheatres were law courts, staged dramas, religious ceremonies and funerals, and served as places of public assembly and debate. The core of the curriculum in the education of a Roman gentleman was rhetoric, with its conventional forms of presentation and persuasion. In adversarial legal systems lawyers' success can be as strongly affected by their theatrical skills as by their legal competences - and these skills are learned from TV as well as fellow-professionals. The judge who condemns the felon to twenty years is in the same moment also pronouncing his "sound bite".

Simple replication of a complex system necessarily involves the reproduction of a "surplus", and part of this surplus gets applied to innovation (if it didn't, the system could not have arrived at its complexity). In organic life the surplus provides the elbow room for "wasteful" innovation, change, and the evolution of functional development (including functions which amplify variation, such as sexual recombination, or, like speciation, "focus" its effects). The same is true for memes, with the difference that variation is not entirely random, but partly intentional.

## **Chainings of intentions: memes in many heads**

Happy the hare at morning, for she cannot read The Hunter's waking thoughts. - WH Auden and Christopher Isherwood, *The Dog Beneath the Skin*.

The impressive achievements of memes arise from chaining us and our intentions together in complex and extremely flexible ways. It is not only humans that use other humans as walking, talking feeling tools - so do memes. This raises the interesting question where "human" intentions actually reside - in individuals, institutions, processes - or (for example) memories or moods.

It also returns us to the question of memes which are too large and/or complex to fit in one mind. This is the obverse of the specialisation and academic division of knowledge, and of functional but incomprehensible artefacts, discussed above. Nowadays this applies to many physical artefacts. A modern airliner is produced in several hundred near-identical versions, with an airframe built on a common template. The "idea" of the airliner - as the basis for a machine which can actually fly - cannot be fitted into less than thousands of minds with specialist skills. It does so by having each of them address a limited, local design problem safe in the knowledge that others, with other skills, will resolve the knock-on problems to which their decisions give rise. This finely-structured bonding of many minds and intentions doesn't occur spontaneously: on the contrary a rising proportion of human labour time is concerned with coordinating others' intentions, coordinating the coordinators, and so on. Much the same is true of a modern motion picture, or of the sophisticated knowledge about the human body distributed through a hospital.

It is over simple to think of complex memes as "leaping" from mind to mind. Bits of them may do so, as one specialist replaces another, but it may often be better to think of large memes slowly migrating or "staining" their way across people, across materials and across institutions. But this image does violence to memes' coherence. We need to allow both for the extended, cellular, character of complex memes, the fact that they replicate themselves bit by bit, and their unities of function or intention. Like the seventeenth century axe, whose blade was replaced in the eighteenth century and whose haft was replaced in the nineteenth, memes can remain one implement though all their parts may be substituted. This is also part of the appeal of organic metaphors.

Conversely, when individual humans relocate, only some of the memes that inhere in them are "portable". The knowledge of a complex organisation's methods, procedures and habits which is "en-gorged" in its management and other staff. It is not simply that much of it is specific to the organisation in which it has grown, in the way that an individual organism's biographical history is specific. It is also that much of it - like the airliner - cannot be competently articulated by one

individual, but exists in the organization's traditions and culture. Collective "management buy-outs" express (among other things) the extent to which such context-bound knowledge is significant.

"Portability" is of the essence in education. Separating the processes of replicating and assembling skills from their application "for real" brings many advantages, especially in permitting the development of new, more elaborate and more powerful skills - analogous to the new horizons for technology opened up by industrialization.

One result is that the development of education itself involves self-referring and thus potentially "runaway" elements. All internally differentiated systems of education build on skills previously acquired. Example of language (learnt as dependent child - ambiguity of "childhood dependency"). Part of education is quality control within production - testing out skills acquired already. Ergo selective pressures exist for the development of intermediate skills per se - and all skills are potentially intermediate. The peacock's train is an adaptation for inseminating peahens with preferences for longer trains. Similarly, education systems develop adaptations promoting the transfer and proliferation of memes. And education systems are not only producers of skills, but major users of them. They thus harbour strong "chreodic" tendencies (Waddington 1972) - to get locked into paths of development once established. These cause continual frictions with the demands upon education to serve life, work, the economy etc.

[One habit education develops is measuring achievement against purposes: Education for what? Vain effort to subordinate ed to life - better to see ed as part of life's self-elaboration. Education of tastes, self-knowledge, how to play in the orchestra of our selves more beautifully, more economically, etc]

[Life as pressure to convert ends into means (i.e. to further elaborate ends). Explanation in terms of function. Problem: potential means include people. Problems of power, control, bureaucracy, progress etc in the runaway evolution of memes. Market as a pseudo-solution of the "coordination" problem.]

## Define your memes?

One of the questions I started with was: Why do people write books? One short answer is: They don't, any more than marriages make babies. A book (or a text, a reading, or a comma) is the expression of a complex meme, produced by humans by means of other humans.

I should have liked to make more generous use in this volume of that ancient (and unfortunately extinct) punctuational meme the *cryphia*, a mark inserted at points in the text where "a hard and obscure question cannot be opened up or solved". {Footnote 129}

Maynard Smith has this to say about evolutionary models of culture:

"The explanatory powers of evolutionary theory rest largely on three assumptions: that mutation is non-adaptive, that acquired characters are not inherited, and that inheritance is Mendelian - that is, it is atomic, and we inherit the atoms, or genes, equally from our two parents, and from

no one else. In the cultural analogy, none of these things is true. This must severely limit the ability of a theory of cultural inheritance to say what can happen and, more importantly, what cannot happen" (1988, 119)

My account of memes has provided no proper definition, nor is it exactly placed within falsifiable propositions. I have tried to explain the idea of memes through numerous examples and analogies: genes, images, games, organs, organisms, organizations, letters, words, texts, performances, melodies, jokes, morals, nations, artefacts, machines, money ... and the list could be far extended. Doesn't the hope of a rigorous definition dissolve in the confusion?

Perhaps it does. But does it matter? Think back to Dawkins' meme meme. If it were a phrase we should describe it as self-referent, and philosophers have long puzzled over the logic of self-referring expressions. But it isn't just words, it's a meme. And, like memes in general, the meme meme hasn't got to be where it is, and as it is, by having a rigorously definable meaning (indeed, it's very hard to say, even approximately, what many types of meme *do* mean). Memes get to be as they are by succeeding relative to other memes. In the process they try out numerous competing - and contradictory - variants. As with advertisements, the memes that succeed aren't necessarily rational, dignified or virtuous. At different phases they domesticate our tastes for different tasks. One week science is all the rage. Next week it will be love or war. To every meme there is a season. Among bees the dancing is all about the best way to flowers with nectar; they have no time for logic, justice or science. {Footnote 130}

Perhaps, then, we should be asking different sorts of questions about the meme meme? Not (by analogy with words) "How may we define it, what does it mean?"; but rather in evolutionary or historical terms "Is it succeeding? Which combination of its variants will do best?" To think in evolutionary terms means not measuring phenomena against definitions, concepts against prior purposes or meanings, but concentrating on the becoming of functions, recognising that the difference between logic and advertisement is only relative. Memes share the farmyard with us, and it is up to us which of them we let push us around. {Footnote 131}

## Meme diversity and biodiversity

This also applies to genes. Is culture the enemy of nature? As the human population and its memes proliferate they exert a downward pressure on the number of species, without - until very recently - compensating by generating new species. This occurs because successful memes do best and thus find it convenient - that is, relatively most efficient - to domesticate or cultivate human beings, together with a limited number of auxiliary species. Successful memes are those able to invade or expand and occupy the human habitat - the raw material of which is one large-brained primate species.

Meme types' "niches" are intricately differentiated, but the soil in which their nutrition chains are all rooted is the amorphous, all-purpose, new-born human creature outlined (and to that extent caricatured) in the Standard Social Science Model. God may have (in JBS Haldane's phrase) an inordinate fondness for beetles, but memes (of which gods are just one sort) all have an inordinate dependence on humans, and some prosper by indirect parasitism - persuading the human hosts *they* parasitise to invade and lay waste other species and their habitats.

It seems to me that the question "Is culture nature's enemy?" cannot be answered theoretically or generally. It depends upon selective pressures among memes, among meta-memes (such as values) and among other types of meme-complexes (such as cultures and lifestyles). One of the ways that some memes may maintain biological diversity is by a sort of weak domestication of species whose population is threatened: wildlife management of various forms. Species are conserved, as breeds are domesticated, by making them into appealing and therefore successful memes. {Footnote 132}

This is an instance of a more general problem. Because memes function largely by motivating humans there are strong pressures tending to make us interpret the world in terms of our own wishes and needs and to anthropomorphize many types of meme. Just as it needed lengthy genetic and memetic evolution before we could see our world as atoms, or cells, or history, we are only beginning to bring our memetic co-inhabitants of the planet into focus.

As we do so ancient questions arise in new forms. How shall we choose among the memes (and among the other biological species) with whom we share our world? Who will they choose? Is the disappearance of the giant panda a more serious matter than the extinction of a non-literate pantheon, or of slave-powered triremes? What influences should we try to bring upon meme-selection? Should we try to cultivate "friendly" memes, and if so how? All cultures consist of more or less cohesive complexes of memes. What is relatively newer is deliberate efforts from within cultures to reshape themselves, as in socialist or utopian movements, or more modestly in the design and subsidy of recyclable artefacts. Some of the most difficult problems will arise from memetic developments in our technologies for breeding humans. Is it wise, for example, to persist with our approximately 50:50 distribution between female and male, when the sexes differ so markedly in their creativity and belligerence? There will be differences over such questions, and selection among the answers.

## Progress and progression

I end with one more speculation. Nowadays the prevalent rates of human memetic evolution have "accelerated" ahead of genetic evolution for large organisms. History may roughly be defined as that phase in the development of the biosphere in which memetic evolution has overtaken genetic evolution. {Footnote 133} Moreover, the velocity of memetic evolution (which we can measure against the more constant-speed "clock" of genetic change) continues to accelerate: where cultures once remained almost unchanged over hundreds of years, they now alter spectacularly between parent and child.

What gave rise to, and sustains, this memetic acceleration? My speculation is that here, too, the cultural proto-sciences may have things to learn from the life sciences. Biologists have paid much, and ingenious, attention to defining, measuring and understanding rates of evolution. I have already tried to appropriate something of their understanding of speciation and sex to appreciate analogous forms of "clustering" among human memes (such as the skill division of labour, or the academic division of knowledge) [cross-ref to passage above].

Another thing that seems to have changed is this: where memetic innovation was once individualized, idiosyncratic and accidental, (some) types of meme have "discovered" the

advantages of routinising innovation. Whereas for a long time most meme lineages were dependent on individuals' effortful but apparently accidental insights - Archimedes' Eureka! or Newton's apple - some more modern memes have developed specialised organs for innovation: R&D departments, creative teaming, sabbaticals, psycho-substance use, postmodern theology, etc. And - partly because the distinction between organisms and adaptations is harder to draw among memes - such organs easily take on momenta of their own. Just as human sexuality, decoupled by contraception from reproduction and morality, proliferates recreational variants, so meme innovation has a tendency to develop as organised addictions among humans. Fortunately for us, in urban society at least, many of them are short lived, and the succession of fashions liberates humans from the absolutism of particular memes. But the "open society" that thus evolves is at least as much a society of memes as it is of humans.

The memetic "take off" can also be seen as exaptations of morality. The initial selective advantages of human morality for memes were mainly stabilizing and conservative: the Ten Commandments, patriotism, diligence, honesty, and so on. But moral memes have increasingly been incorporated in meme exaptations whose advantages lie in their links with innovation - imagination, science, tax avoidance, etc. The clades which vary most eclectically and ingeniously generate many of the memetic winners.

Fashions (which may very loosely be thought of as analogous to cell division and multiplication in the life cycles of multicellular organisms) remind us how very approximately notions - such as those of replicators and interactors - must be transposed from biology to memology. In neither field is it possible to draw one clear division - like means and ends, they interchange. Post-Darwinian biologists have met with large difficulties in their central concepts - yet great progress has been made despite the abstract confusion. Much modern biology developed with assumptions about the relevant "units" of natural selection (the organism, or the group) that may seem unreflective. And defining and distinguishing the lineages on which biological selection acts is still a very live issue. {Footnote 134}

Among memes there is no unit of selection which even appears so clear cut as does the individual organism in biology. That is why, to begin with, I blithely skated over distinctions between more and less complex complexes of memes. Fashions are a very rough analogue of organisms - or at least a relevant and informative level of selection. So, as I suggested above, are (human individuals') personalities. And so, too, are the stuff-in-trade of social scientists: institutions and cultural practices. Durkheim's "autonomy of the social" is relative but very real - Institutions renew themselves through the reassembly of finite personalities - and personalities become material for inclusion in institutions by their formation in (often other) institutions. If we remember this we may be less prone to amplify the dialogue-of-the-deaf between social and natural science, as with the "Standard Social Science Model" and its critics, discussed above.

Most of my suggestions have been directed towards the social sciences. But - as the natural sciences, heavily addicted to the separation of fact and value, are becoming more aware - specialisation and differentiation also involves getting memes crossed - and high levels of juvenile mortality among hybrids.

## References

- Seidman, Steven (1994 Sep) Contested Knowledge: social theory in the postmodern era (textbook), Blackwell, Oxford.
- Turner Bryan (1995 Apr) The Blackwell Companion to Social Theory, Blackwell.
- Kumar Krishan (1995 Feb) New theories of modern society, Blackwell. Textbook, 3 theories: information society; post-Fordism; postmodernism.
- Bauman, Zygmunt (1993) Postmodern ethics, Blackwell. Re-enchantment of the emotional, returning dignity to emotions.
- Sztompka, Piotr (1993) The sociology of social change, Blackwell. Textbook: 4 visions: evolutionary, cyclical, dialectical, post-developmentalism.
- Rose, Gillian (1993) Judaism and modernity: philosophical essays, Blackwell.
- Law, John (1993) Organising modernity: social ordering and social theory, Blackwell. Sociology of verbs, recursive modes.
- Teubner Gunther (1993) Law as an autopoetic system, Blackwell. Hb only.
- Soper, Kate (1995 Feb) What is nature?, Blackwell.
- Gottdiener, Mark (1994 Oct) Postmodernism and semiotics: material culture and the forms of postmodern life, Blackwell. Disneyland, architecture, unisex etc.
- Stivers, Richard (1994 May) The culture of cynicism: American morality in decline, Blackwell.
- Toussaint-Samat, Maguelonne (tr Bell) (1992) A history of food, Blackwell, 801 pp.
- Zukin, Sharon (1995 Mar) The culture of cities, Blackwell.
- Massumi, Brian (ed) (1994?) The Politics of Everyday Fear, Minnesota UP (i-D). On ambient fear etc.
- Wilson DS and Sober E (1993) "Re-introducing group selection to the human behavioural sciences", Behavioural and Brain Sciences (in press Dec 93). See also comments in the same journal?
- Kramer, PD (1993) Listening to Prozac, Viking NY.
- McDowell, Colin (1992) Hats, Thames and Hudson.
- McDowell, Colin (1989?) Directory of Twentieth Century Fashion, 2nd ed
- Bell, Daniel (1993) Communitarianism and its critics, OUP.
- Aida S (ed) (1985) The science and praxis of complexity, UNUP, Tokyo. Inc Prigogine.
- Nozick, Robert (1993) The nature of rationality, Princeton UP. Well reviewed by Ian Hacking, LRB, 27 January 1994.
- Nash, Walter (1993) Jargon: its uses and abuses, Blackwells, Oxford. Well reviewed by Peter Lennard, LRB, 27 January 1994.
- Orbach, S (198?), Fat is a feminist issue
- D'Souza, Christa (1994) "Through thick and thin", Sunday Times Style and Travel Section, 23 January 1994. Oprah Winfrey's diet, roller coaster media attention strategy, "reinventing oneself as if yesterday never happened".
- Christansen, Arthur (19??) Headlines all my life, on headline making, by S Express journalist.
- Kepel, Gilles (1993) The Revenge of God: the resurgence of Islam, Christianity and Judaism in the modern world, Polity Press, Cambridge. Rev in ST Books section 9 January 1994: Christian infant/crucified god an abomination; 25 year rule; birth of nations; televangelism.

- Kilgarriff Micheal (ed) (1987) 3001 Jokes from Kids, Ward Lock, London. First farmer: "Is that new scarecrow of yours any good, Giles?" Second farmer: "Good? I'll say so. The crows are so scared they're bringing back the seeds they stole last week."
- Baker, Houston A (1993) Black studies, rapping and the academy, U Chicago Press.
- Henig, Jeffrey R (1994) Rethinking school choice: limits of the market metaphor, Princeton UP.
- Messer-Davidow, Ellen, Shumway, David R and Sylvan, David J (eds) (1993) Knowledges: historical and critical studies in disciplinarity, UP of Virginia, Charlottesville VA. (Launching new series: Knowledge: Disciplinarity and beyond. Write for details to PO Box 3608, University Station, Charlottesville VA 22903 - NYRB)
- Jamison, Kay R (1993) Touched with Fire: manic-depressive illness and the artistic temperament, The Free Press. Rev ST Books section 2 January 1994. Woolf, Byron, Hemingway, Pound, Plath. Writing as therapy. Manic-depressive genes.
- Allen PM and Lesser M (1991) "Evolutionary human systems: learning, ignorance and subjectivity", in Saviotti PP and Metcalfe JS (eds) Evolutionary theories of economic and technological change: present status and future prospects, Harwood, Reading. Economics as locked into extravagant evolution.
- Allison PD (1992) "Cultural relatedness under oblique and horizontal transmission rules", Ethology and Sociobiology, 13, 3.
- Appleyard, Brian (1993?) Understanding the Present: science and the soul of modern man, (deals with scientific specialization, amorality and separation from life?)
- Baines, J. (1983) "Literacy and ancient Egyptian society", Man, (New Series), Vol 18, pp 572-99.
- Baker, Nicholas (1993) "Survival of the fittest", New York Review of Books, 4 November 1993. Review with evolutionary theme of Parkes (1993). Word spacing, separation, slash and pause, paragraph, smiley. Printing and other simplification of conventions (commash, etc). Inverted C (ancient cryphia) to signal places in the text where "a hard and obscure question cannot be opened up or solved" - dropped because classification of error etc too rich for short symbols? Inverted commas only c1714. Refs to Lennard, John (1991) But I Digress, OUP.
- Bakhtin / G H Mead on language and its place in social psychology
- Ball, John A (1984) "Memes as replicators" Ethology and Sociobiology, 4: 145-161. Excellent, lively on why memes proliferate.
- Barkow JH, Cosmides L and Tooby J (eds) (1991) The Adapted Mind, Oxford UP.
- Barkow JH (1989) Darwin, Sex and Status, U of Toronto Press.
- Barkow JH and Burley N (1980) "Human fertility, evolutionary biology, and the demographic transition", Ethology and Sociobiology, 1:163-80.
- Barkow JH (1989) "The elastic between genes and culture", Ethology and Sociobiology, 10, 1-3.
- Barnes, Timothy D (1993) Athanasius and Constantius: theology and politics in the Constantinian empire, Harvard UP.
- Barnhart Robert K (ed) (1993? repr) Dictionary of Etymology, HW Wilson Co, NY. Evolution of 30,000 English words.
- Bartlett FC (1932) Remembering, Cambridge UP, Cambridge.
- Basalla G (1988) The evolution of technology, Cambridge University Press.

- Pinker, Steven (1993) *The Language Instinct: the new science of language and mind*, A Lane.
- Wills, Christopher (1993) *The Runaway Brain: the evolution of human uniqueness*, Harper Collins.
- Baumann, G (1988?) *The Written Word*, OUP, Oxford. Scrutinizes the social and cultural implications of literacy.
- Baumann, Zygmunt Postmodernism ? On public/social recognition as the limiting resource of postmodernism (funny resource - recognition of what - return via species to differentiation and surplus).
- Beard, Mary and North, John (eds) (1990) *Pagan Priests: religion and power in the ancient world*, Cornell UP, Ithaca NY. hdbk \$39.95
- Beckford, James A. (1985) *Cult Controversies: the Societal Response to New Religious Movements*, Tavistock Publications, London.
- Beckstrom JH (1989) *Evolutionary Jurisprudence*, U of Illinois Press, Urbana. V good on is/ought, illumination sociobiology offers, defining genetic self. Positivist. See on altruism, doubted paternity, incest, knowledge of adoption, self-interest of lawmakers. Plus (with RD Alexander etc) on inner meaning of biocultural evolution, as culture also (Lamark-wise) inherited, but NOT as memes being evolving entity. See also on parental choice, plus Beckstrom's (1985) *Sociobiology and the Law: the biology of altruism in the courtroom of the future*. And see on universalistic altruism, immortality of species etc.
- Benjamin, Walter ( ) *Art in the age of mechanical reproduction*
- Bercovitch FB (1991) "Social stratification, social strategies and social success in primates", *Ethology and Sociobiology*, 12, 4.
- Berlin Isiah (with Henry Hardy) (1993) "The Magus of the North" *New York Review of Books*, 21 October 1993. On Johann Georg Hamann (1720-88), anti-rationalist. "Think less and live more" - Hamann to JG Herder, 1765.
- Bernstein, Alan E (1993) *The formation of hell: death and retribution in the ancient and early Christian worlds*, Cornell UP.
- Bertram BCR (1993) *The ostrich communal nesting system*, Princeton UP.
- Betzig LL, Borgerhoff Mulder M and Turke PW (eds) (1988) *Human reproductive behaviour*, Cambridge UP.
- Betzig LL (1989) "Rethinking human ethology: a response to some recent critiques" *Ethology and Sociobiology*, 10, 5.
- Biederman, Irving (1987) "Recognition-by-Components: a theory of human image understanding", *Psychological Review*, 94, 115-47. Dictionary noun evidence on discriminating 30000 objects (qu Petrovski)
- Blau, Herbert (1993) *The Audience*, Johns Hopkin UP. Audience as common culture, now as fragmented media.
- Bloom, Clive (ed) (1990) *Spy Thrillers: from Buchan to Le Carre*, Macmillan, London. 202 pp.
- Boden, Deirdre (1990) *The Business of Talk*, Polity Press, Cambridge. L25 hdbk
- Bohman, James (1991) *New Philosophy of Social Science: problems of indeterminacy*, Polity, Cambridge, 273pp.
- Bok, Sissela (1978) *Lying: moral choice in public and private life*, Harvester Press.
- Bok, Sissela (1982) *Secrets: on the ethics of concealment nd revelation*, Pantheon.

- Bonner, John Tylor (1993?) Life cycles: reflections of an evolutionary biologist, Princeton UP.
- Bonner, JT (1980) The evolution of culture in animals, Princeton UP, Princeton.
- Bonner, S F (1989?) Education in Ancient Rome, Bristol UP, Bristol. 3 century BC to Trajan.
- Boulding KE (1981) Evolutionary Economics, Sage, Beverley Hills CA.
- Boyd R and Richerson PJ (1985) Culture and the evolutionary process, Chicago UP, Chicago.
- Boyd R and Richerson PJ (1992) "Punishment allows the evolution of cooperation (or anything else) in sizeable groups", Ethology and Sociobiology, 13, 3.
- Bradie M (1986) "Assessing evolutionary epistemology", Biology and Philosophy, 1, 401-60.
- Brandon R and Burian R (eds) (1984) Genes, organisms, populations: controversies over the units of selection. NB selection NOT transmission.
- Brooks, Peter (1993) Body work: objects of desire in modern narrative, Harvard UP.
- Brunvand, Jan Harold (1993) The Baby Train and other lusty urban legends, WW Norton, New York and London.
- Brunvand, Jan Harold (1993) The Baby Train and other lusty urban legends, Norton, 14.95. "The kidney heist", the vanishing hitchhiker, the Mexican pet, and Curses, Broiled again. Urban myths surviving through mutations. Review by Christopher Bray in Sunday Times Book Section 5 December 1993. Levi-Strauss: sketches of the mind, horror of urban myths is expression of our need to believe in an even worse world.
- Burkert, Walter (1987) Ancient Mystery Cults, Harvard UP, Cambridge MA. 176pp.
- Cameron, Averil (ed) (1988ish?) History as Text. LRB review of Canfora (1989) on rethinking of (ancient?) history as transmitted texts - cf Goody etc?
- Campbell D "The two distinct routes beyond kin selection to ultrasociality: implications for the humanities and social sciences", in D Bridgeman (ed) (1983) The nature of prosocial development: theories and strategies. On universalistic, "kantian" etc altruism.
- Campbell D (1974) "Evolutionary epistemology" in Schlippe PA (ed) The Philosophy of Karl Popper, Lasalle IL.
- Candland, Douglas Keith (1993) Feral Children and Clever Animals: reflection on human nature, OUP.
- Canfora, Luciano (1989) The Vanished Library: a wonder of the ancient world, Radius. 205p L14.95 0 09 174049 5. Reviewed LRB - on the library at Alexandria and other ancient depositaries as vulnerable concentration/target in literate culture (cf GC Williams on mass extinctions; Dennett on scholars/libraries). Bias of literary fossil record via copying and "imprvement" - fossil record consists of "Piltdown"s.
- Caplan AL (1982) "Stalking the wild culturegen", Behavioural and Brain Sciences, 5, 1, 8-9.
- Carnoy M, Castelis M, Cohen SS, Cardozo FH (1993) The new global economy in the information age, Penn State UP.
- Carrithers, M, Collins S and Lukes SL (eds) (1985) The category of the person: anthropology, philosophy, history, Cambridge UP, Cambridge.
- Cashdan E (1989) "Hunters and gatherers: economic behaviour in bands", in S Plattner (ed) Economic Anthropology, Stanford U Press. See on Ache food sharing (meat - group / plants - family)

- Catchpole CK (1986) "The biology and evolution of bird songs", Perspectives in Biology and Medicine, 30, 47-62.
- Cavalli-Sforza LL, Feldman MW, Chen KH, and Dornbusch SM (1982) "Theory and observation in cultural transmission", Science, 218, 19-27.
- Cavalli-Sforza LL and Feldman MW (1981) Cultural transmission and evolution, a quantitative approach, Princeton UP, Princeton.
- Chagnon, Napoleon A (1992) Yanomamo: a vanished Eden,
- Chagnon, Napoleon A (1993) "Killed by kindness? The dubious influence of the Salesian missions in Amazonas", TLS 24 December 1993. Miners massacres of Yanamamo. Salesian competition with protestant missions; medical care; shotguns; attract and reduce; adult mortality. Charities "own" survival groups.
- Chargaff, Erwin (1986) Serious Questions: an ABC of Skeptical Reflections, Birkhauser, Boston MA.
- Chargaff, Erwin ( ) Heraclitean Fire
- Chartier, Roger (ed) (1989) Power and the uses of print in early modern Europe, 15-19th centuries, Polity Press, Cambridge.
- Chester, Andrew (1990) The Social Context of Early Christianity, Polity Press, Cambridge. pbk cL7.95.
- Cheyney, Dorothy and Seyfarth, Robert (1990) How Monkeys see the world. Rev TLS 30 Nov 1990. On vervet monkeys communication.
- Clark, Grahame (1992) Space, Time and History: a prehistorian's view, CUP.
- Clutton-Brock, Juliet (1993) Cats ancient and modern, Harvard UP/BM. History from 7000 years ago.
- Clutton-Brock, Juliet (ed) (1989) The walking larder: patterns of domestication, pastoralism, and predation, London: Unwin Hyman, xxii,367p.
- Coulter, Jeff (1989) Mind in Action, Polity Press, Cambridge. pbk L8.50.
- Crawford CB, Salter BE and Jang KL (1989) "Human grief: is its intensity related to the reproductive value of the deceased?", Ethology and Sociobiology, 10, 4.
- Crawley, Micheal J (ed) (1992) Natural enemies: the population biology of predators, parasites and diseases, Blackwell Scientific. Taxonomic body-text, plus theory reviews on coevolution theory, population dynamics, optimal foraging theory. Rev by Simms (1993) TREE, 8,2, "Eaten and being eaten": should have tried to explore 3+ species models, eg impact to predation pressure on foraging behaviour. [Compare with predation pressure on copulation ("foraging" of 2-sexes).]
- Creel S (1993) "Why cooperate? Game theory and kin selection.", TREE (Trends in Evolutionary Ecology), 8, 3. Mutualism, kin selection, r-values. Circumstances for unrelated cooperation (rule of thumb - beat with stick no thicker than thumb).
- Cronk, Lee (1991) "Human behavioural ecology", Annual Review of Anthropology, 20: 25-53. Good overview, especially enculturation, adoption, parenting etc.
- Crook JH (198 ) The Evolution of Human Consciousness, OUP.
- Daly M and Wilson M (1989) "Homicide and cultural evolution" Ethology and Sociobiology, 10, 1-3.
- Dandeker, Christopher (1989) Surveillance, Power and Modernity: bureaucracy and discipline from 1700 to the present day, Polity Press, Cambridge. c250pp cL25

- David PA (1986) "Understanding the economics of QWERTY: the necessity of history", in Parker WN (ed) (1986) Economic History and the Modern Economists, Blackwell, Oxford.
- David PA (1987) "Some new standards for the economics of standardization in the information age", in Dasgupta P and Stoneman PL (eds) Economic policy and technological performance, Cambridge UP, Cambridge.
- Dawkins R (1983) "Universal Darwinism", in DS Bendall (ed) Evolution from molecules to men, Cambridge UP.
- Dawkins, Richard (1993) "Viruses of the mind", in B Dahlbom (ed) Dennet and His Critics, Blackwell, Oxford. An abbreviated version of Dawkins essay appeared in New Statesman, 1933, and a BBC2 TV late show item on 30 Nov 1933.
- Dawkins, Richard (1994) "God's Utility Function", Sunday Times Culture Section, 9 January 1994, pp 14-15. [Extract from 15 March 1994 ST lunchtime lecture] [tel 071-792 9512 for details, access] "The universe that we observe has precisely the properties we should expect if there is, at bottom, no design, no purpose, no evil and no good, nothing but blind pitiless indifference." END
- de Sainte Croix, G E M (1990) The Class Struggle in the Ancient Greek World: from the archaic age to the Arab conquests, Cornell UP, Ithaca NY. Pbk \$19.95
- Dawkins R (1993) "Is religion just a disease?", The Daily Telegraph, 15 December 1993.
- Dawkins, Richard (1982) The extended phenotype: the gene as the unit of selection, Oxford: W.H. Freeman.
- Dawkins, Richard (1976) The Selfish Gene, OUP.
- Dawkins, Richard (19 ) The Blind Watchmaker
- De Landa, Manuel (1992) War in the age of intelligent machines, MIT Press, Cambridge MA.
- Delius, Juan D (1990) "Of mind, memes and brain bugs: a natural history of culture", in W A Koch (ed) (1990) The nature of culture, Bochum.
- Delius, Juan D (1991) "The nature of culture", in Dawkins MS, Dawkins R and Halliday TR (eds) (1991) The Tinbergen Legacy, Chapman and Hall, London, 75-99.
- Dennett, Daniel C (1991?) Consciousness Explained, Penguin.
- Dennett, Daniel C (1987?) The Intentional Stance, MIT Press, Cambridge MA.
- Dennett DC (1991) "Two contrasts: folk craft versus folk science and belief versus opinion", in J Greenwood (ed) The Future of Folk Psychology: Intentionality and Cognitive Science, Cambridge, CUP.
- Dennett, DC (1989) "Cognitive ethology: hunting for bargains or a wild goose chase?", in A Montefiore and D Noble (eds) (1989) Goals, Own Goals and No Goals: a debate on goal-directed and intentional behaviour, Unwin Hyman.
- Dennett DC (1990) "Memes and the exploitation of imagination", Journal of Aesthetics and Art Criticism, 48, 127-35. Fuller than Consciousness Explained.
- Derber, Charles, Schwartz, William A, Magrass Yale (1990) Power in the highest degree: Professionals and the rise of a new mandarin order, OUP, NY.
- Derrida, Counterfeit Money
- Diamond, Jared (1990?) The rise and fall of the third chimpanzee,
- Directory of Intentional Communities (1991?) edited by the Fellowship for Intentional Community (Box T, Deadwood, OR 97430), 328p, \$18.00 ppd, ISBN 0-9602714-1-4.

- Donald, Merlin (1991) *Origins of the modern mind: three stages in the evolution of culture and cognition*, Harvard UP, 413pp, 22.25. TLS: two transitions: episodic to mimetic culture, including language; thence to theoretic culture with External Symbolic Storage; hints at further transition. Also on 1980s+ neuroscience. Ref in Sacks on Edelman (NYRB 8 April 1993)
- Douglas, Mary and Hull, David (ed) (1992) *How classification works: Nelson Goodman among the social sciences*, Edinburgh University Press. Articles by Ian Hacking on evolution of child abuse and classificatory imperialis. David Hull on biological species.
- Dummett, Michael (1989?) *Voting Procedures*, OUP, Oxford.
- Dupuy, Trevor (1984) *The evolution of weapons and warfare*, Fairfax VA.
- Durham, W (1990) "Evolutionary culture theory", *Annual Review of Anthropology*. Full references to 1990. Defines ECT by common continuous descent of cultures. Overview of five models (full refs). Then concentrates on primary means of effective cultural transmission (secondary values).
- Eco, Umberto and Zorzoli, GB (1963) *The picture history of inventions from plough to Polaris*, Macmillan, NY.
- Edelman, Gerald M (1993) *Bright Air, Brilliant Fire*,
- Edelman, Gerald M (1990) *The Remembered Present: a biological theory of consciousness*, Basic Books, New York NY. Draws on recent neuroscience.
- Edie, James M, Scanlan, James P, and Zeldin, Mary-Barbara (eds) (1965) *Russian Philosophy*, 3 vols, Chicago.
- Ehrenreich, Barbara (1989?) *Fear of Falling: the inner life of the middle class*, Pantheon. 292p \$18.95. US new class fear of falling, education as antidote.
- Eldredge, Nils (1992) *The Miner's Canary*, Prentice-Hall.
- Elster, Jon (1989) *Solomonic judgements: studies in the limitations of rationality*, CUP, Cambridge. OUL L 302.13 ELS.
- Elster, Jon (1983) *Explaining technical change*, Cambridge UP, Cambridge. FP - On the appropriate forms of explanation in biology and social science. OUL L 306.46.
- Elster, Jon (1993) *Local justice: how institutions allocate scarce goods and neccessary burdens*, Russell Sage Foundation.
- Enquist, Magnus and Arak, Anthony (1993) "Selection of exaggerated male traits by female aesthetic senses", *Nature*, 361, pp 446-8. Neural networks; artificial recognition systems; hidden female preferences; general biases of nervous recognition systems; signalling in general; ritualization. USE in social surplus section. See comment and interpretation in Howlett (1993).
- Enteman, Willard F (1993) *Managerialism: the formation of a new ideology*, U Wisconsin Press. Irrelevance of voting.
- Ethology and Sociobiology, 11, 4/5. Special discussion issue on "Darwinian psychology"; articles by Alexander RD, Turke PW, Barkow JH, Blurton Jones NG, Irons W, Tooby J and Cosmides L, Silk JB, Symons D.
- Etzioni-Halevy, Eva (1992) *The Elite Connection: problems and potential of western democracy*, [textbook, on plural elites].
- Ewer, RF (1969) "The 'instinct to teach'", *Nature*, 222, 698.
- Fagan, Brian M (1990?) *The Journey from Eden: the peopling of our world*, Thames and Hudson, London. 256pp. 12.95.
- Farrell, Christopher J, a theory of technical progress, unpub but inc in Petroski.

- Ferguson, Eugene S (1992) *Engineering and the mind's eye*, MIT Press, Cambridge MA.
- Finke, Roger and Stark, Rodney (1993) *The Churching of America, 1776-1990: winners and losers in our religious economy*, Rutgers University Press, pb.
- Finley, Moses I (1965) "Technical innovation and economic progress in the Ancient World", *Economic History Review*, 18, pp 29-45.
- Fischer, Ernst ( ) *The Necessity of Art*
- Fisher, Helen (1992) *The anatomy of love*, Simon and Schuster.
- Fiske AP (1991) *Structures of social life: the four elementary forms of human relations*, Free Press, New York. Quoted in *The Adapted Mind: 4 models of sharing*, including: communal sharing (Ache), coningent exchange (barter) (is kin sharing a third).
- Forty, Adrian (1986) *Objects of desire*, Pantheon, NY.
- Fowden, Garth (1993) *The Egyptian Hermes: a historical approach to the late pagan mind*, Princeton UP, new pbk ed. "Sage, scientist and sorcerer, Hermes Trismegistus was the culture-hero of Hellenistic and Roman Egypt."
- Fowden, Garth (1993) *Empire to Commonwealth: consequences of monotheism in late antiquity*, Princeton UP.
- Galbraith, JK (1992) *The Culture of Contentment*, Penguin. Functional underclass, immigrant labour, black economy.
- Galef BG and Zentall T (eds) (1988) *Social learning*, Earlbaum, Hillsdown NJ. Includes article by Galef on "Imitation in animals"
- Gardiner, A.H. (1935) *Hieratic Papyri in the British Museum, Third Series*, British Museum, London. (cited in James, *Ancient Egypt*, 96, on vocation of scribe).
- Gay, Peter (1993) *The Cultivation of Hatred*, Norton. C19 literary and social aggression, status-seeking.
- Gazzaniga, Michael S (1993) *Nature's Mind: the biological roots of thinking, emotions, sexuality, language, and intelligence*, Basic Books, New York.
- Geertz, Clifford (1973) *The interpretation of culture*. "Thinking .. goes on ... not in the head but the market place, yard, in traffic etc ..." (quoted in *The Adapted Mind* (or Moscovici?))
- Geertz, Clifford (1990?) *Works and Lives: the anthropologist as author*, Stanford UP. \$7.95 pbk. Ruth Benedict, Evans-Pritchard, Malinowski, Levi-Strauss.
- Gelb, I.J. (1963) *A Study of Writing*, second edition, Chicago.
- Giblin, James Cross (1987) *From hand to mouth: or, How we invented Knives, Forks, Spoons, and Chopsticks & the table manners to go with them*, Crowell, NY.
- Giddens, A (1984) *The constitution of society: outline of the theory of structuration*, Polity Press, Cambridge. See Ch 5 (pp 227-280) on "Change, evolution and power", with refs pp 274-80 on social/cultural evolution, especially critique of Parsons.
- Gifford, Don (1993) *The Farther Shore: a natural history of perception 1798-1984*, Atlantic Monthly Press.
- Gilfillan, SC (1935, repr 1970) *The Sociology of Invention*, MIT Press (plus illustrated companion volume: *Inventing the Ship*) (FP)
- Gilfillan, SC (1964) *Invention and the Patent System*, US Government Printing Office.(FP)
- Glantz K and Pearce J (1989) *Exiles from Eden: psychotherapy from an evolutionary perspective*, WW Norton, New York.

- Godfrey-Smith, P and Lewontin, R C (1992) "The dimensions of selection", *Philosophy of Science*.
- Golden, Mark (1993) *Children and childhood in classical Athens*, Johns Hopkins UP.
- Goodway, David (1990?) *For Anarchism: History, Theory, Practice*, Routledge History Workshop Series, 12.99. Rev Solidarity 25/26 by Quail (Slow burning fuse). Levy on Italian anarchism to 1926, plus Alan Carter "Outline of an anarchist theory of history": "history driven by the need of the state to develop the "forces of coercion", in turn conditioning the development of the forces and relations of production".
- Goodwin, Malcolm (1993) *Angels: an endangered species*, Boxtree, 16.99. Review Frances Spalding, ST Book Section 5 December 1993. Obscure sources, gen view of angels as our self-knowledge.
- Goody, J. (1986) *The Logic of Writing and the Organization of Society*, Cambridge. {Use 1-2 pp "Specialization: priests and intellectuals", pp16-17]
- Goody, J., ed (1968) *Literacy in Traditional Societies*, Cambridge, CUP.
- Goody, J. (1977) *The Domestication of the Savage Mind*, Cambridge University Press, Cambridge.
- Grafton, Anthony (1993) *Forgers and Critics: creativity and duplicity in Western scholarship*, Princeton UP.
- Bell, Daniel (1993?) *Communitarianism and its critics*, OUP.
- Grassivaro Gallo P and Viviani F (1992) "The origin of infibulation in Somalia: an ethological hypothesis", *Ethology and Sociobiology*, 13, 4, 253-66.
- Green, M. (1981) "The construction and implementation of the cuneiform writing system", *Visible Language*, Vol 15, 345-72.
- Griffin, Donald R (1992) *Animal Minds*, U Chicago Press.
- Guthke, Karl S (1993) *The Last Frontier: imagining other worlds, from the Copernican revolution to modern science fiction*, Cornell UP.
- Haig, D Maternal-fetal conflict in human pregnancy and pregnancy complications
- Halliday MAK and Martin (1993) *Writing Science: literacy and discursive power*, University of Pittsburgh Press. Evolution of scientific discourse.
- Hallpike CR (1982) "The culturegen - science or science fiction?", *Behavioural and Brain Sciences*, 5, 1, 12-3.
- Halperin, David A. (1983) *Psychodynamic Perspectives on Religion, Sect and Cult*, John Wright-PSG Inc, London.
- Hammond, John A (1975) Adaption in natural and artificial systems. [PCW ref]
- Hampson, Norman (1991) *Saint-Just*, Basil Blackwell, Oxford. 245 pp. Rev E Weber TLS 15 Feb 91. SJ as theorist of paper, bureaucracy (article by Guy Thuillier: Le ministere et un monde en papier. Egypt and Rome thought much, wrote little, See other sources on papyrus and empire. SJ: "A people has only one dangerous enemy: its government."
- Haraszti, M (1977) *A Worker in a Workers' State: piece rates in Hungary*, Penguin, 1977.
- Harcourt, GC (1972) *Some Cambridge controversies in the theory of capital*, Cambridge UP. See refs on reswitching etc.
- Hardison OB (1993) *Disappearing through the skylight: culture and technology in the twentieth century*, Viking. Snidely rev in NYRB 26.4.93: eternal, gravity-free, telepathic, androids vs carbon man.

- Harpending H and Rogers R (1990) "Fitness in stratified societies", *Ethology and Sociobiology*, 11, 6.
- Harris, Marvin (1979) *Cultural Materialism: the struggle for a science of culture*, New York, Random House.
- Harris, William V (1989) *Ancient Literacy*, Harvard UP, Cambridge MA. 408pp.
- Harth, Erich (1989) *Dawn of a Millenium: beyond evolution and culture*, Little, Brown and Co, New York NY.
- Harvey PH, Pagel MD (1991) The comparative method in evolutionary biology, OUP.
- Havelock, E.A. (1982) *The Literate Revolution in Greece and its Cultural Consequences*, Princeton UP, Princeton NJ.
- Hawthorn, Geoffrey (1991) *Plausible Worlds: possibility and understanding in history and the social sciences*, CUP, Cambridge, 192pp.
- Healey, Phil and Glanvill, Rick (eds) (1992) *Urban Myths*, Virgin, London.
- Healey, Phil and Glanvill, Rick (eds) (1993) *The Return of Urban Myths*, Virgin, London.
- Hearn, Jeff, Sheppard, Deborah L, Tancred-Sherriff, Peta and Burrell, Gibson (eds) (1989) *The Sexuality of Organization*, Sage, Newbury Park, CA. Reviewed by Sorca O'Connor, *Educational Administration Quarterly*, 28, 1, 125-136 (Feb 1992): all-pervasiveness of sexuality in gendered organizations; make them gynocentric.
- Hewson, Barbara (1993) "Doing justice to women's rights", *Law Society Gazette*, 90/35 (20 September 1993), pp 19-21)
- Heyes CM and Plotkin HC (1989) "Replicators and interators in cultural evolution", in Ruse M (ed), *What the Philosophy of Biology is: essays dedicated to David Hull*, Kluwer, Dordrecht.
- Hill EM and Low BS (1992) "Contemporary abortion patterns: a life history approach", *Ethology and Sociobiology*, 13, 1, 35-47.
- Hirst PQ (1976) *Social evolution and sociological categories*, Allen and Unwin, London.
- Hirst PQ and Wooley P (1982) *Social relations and human attributes*, Tavistock, London.
- Hockney, David (1993) *The Way I See It*, Thames and Hudson, London.
- Hadamard, *Psychology of invention in the Mathematical Field* (FP)
- Hodgson, GM (1993) *Economics and evolution: bringing life back into economics*, Polity, Cambridge.
- Holldobler Bert and Wilson EO (1993) *The Ants: 100,000,000 years in the making*, Harvard UP.
- Hood, Roger (1989) *The Death Penalty: a worldwide perspective*.
- Horton, Robin (1993) *Patterns of thought in Africa and the West: essays on magic, religion and science*, CUP. Rev Appiah TLS 2.7.93. W African religion addresses same questions as western science; "devout opposition" imposing Christian monotheism.
- Howlett, Rory (1993) "Beauty on the brain", *Nature*, v 361, pp398-9 (4 Feb 1993). "Darwin argued that either such [extravagant male] traits had evolved because of the mating preferences of females (intersexual selection - eg peacocks' trains), or because they had a direct role in the competition between males for access to females (intrasexual selection - eg antlers)." Commenting on Enquist and Arak (same issue): the "ability of simple [neural network] recognition mechanisms to generalise from familiar stimuli to similar but otherwise novel stimuli could exert considerable selective pressure on signal form"

- Hoyrup, J. (1985) "Varieties of Mathematical Discourse in Pre- Modern Socio-Cultural Contexts: Mesopotamia, Greece and the Latin Middle Ages", *Science & Society*, vol LXIX, no 1, Spring, pp 4-41.
- Hughes, H Stuart (1988) *Sophisticated Rebels: the political culture of European dissent, 1968-1987*, Harvard UP, Cambridge MA. 192 pp.
- Hughes, Robert (1993) *The culture of complaint: the fraying of America*, OUP.
- Hugman, Richard (1991) *Power in Caring Professions*, Macmillan, 260p, 9.99.
- Hull, David L. (1988) *Science as a process: an evolutionary account of the social and conceptual development of science*, University of Chicago Press, Chicago IL. xiv, 586 pp illus, \$39.95.
- Hull, David L (1988) "Interactors versus vehicles", in Plotkin HC (ed) *The role of behaviour in evolution*, MIT Press.
- Hull, David L (1982) "The Naked Meme", in Plotkin HC (ed) *Learning, development and culture: essays in evolutionary epistemology*, Wiley, Chichester.
- Hull, David L (1978) "Altruism in science: a sociobiological model of cooperative behaviour among scientists", *Animal Behaviour*, 26, 658-97.
- Humphrey, Nicholas (1993?) *A history of the mind: evolution and the birth of consciousness*, Harper Collins.
- Illich, Ivan (1993) *In the Vineyard of the Text: a commentary to Hugh's Didascalicon*, U Chicago Press. C12 transition from monastic to scholarly writing (punctuation; indentation; chapters; indices); more important than printing revolution.
- Ingold, Tim, Riches, David and Woodburn, James (eds) (1988) *Hunters and gatherers; 1: History, evolution and social change*, Oxford: Berg, viii,331p, 2 vols in all.
- Ingold, Tim (1986) *Evolution and social life*, Cambridge: Cambridge University Press, xv,431. Subjects: Social evolution/Human evolution/Culture.
- Ingold, Tim (c 1990) "An anthropologist looks at biology", *Man NS* 25, 208-229. Recommended by Brian Goodwin, tel 12 May 1994, as interested in formation of person: thus considers memes etc (Goodwin regards it as a vague concept)
- Jackson JBC, Buss LW and Cook RE (eds) (1986) *Population biology and evolution of clonal organisms*, Yale UP, New Haven CT.
- Jewkes, J, Sawers D, Stillerman R (1958) *The sources of invention*, Macmillan. Many case studies. FP
- Johnson, Mark (1987) *The Body in the Mind*, U of Chicago Press.
- Johnson JL, McAndrew FT and Harris PB (1991) "Sociobiology and the naming of adopted and natural children", *Ethology and Sociobiology*, 12, 5.
- Kaprow, Allan (1993) *Essays on the blurring of art and life*, U California P. Performance art.
- Kawai M (1965) "Newly acquired pre-cultural behaviour of the natural troop of Japanese monkeys on Koshima Islet", *Primate*, 6, 1-30.
- Keegan, John (1993) *A history of warfare*, Knopf.
- Kennedy, Duncan (1993) *Sexy dressing etc*, Harvard UP. Exponent of Critical Legal Studies.
- Kirk, Raven and Schofield pp 72ff on wholesale meme mutations in Greece between C9th and C6th BC. FP
- Kline, Stephen (1993) *Out of the Garden: toys, TV and children's culture in the age of marketing*, Verso, London. On marketing innovation from late C19, Disney and

"merchandising". Read esp for childhood as sensitive period (Dawkins). Ch 1 good on consumerism and its meanings. Talked to children - v good.

- Knight, Chris (1991) *Blood Relations: menstruation and the origins of culture*, Yale UP, 581pp, 40.
- Koss, Mary (1993) *The Scope of Rape*, Psychologist at Arizona. 70 pc US undergraduates victims of rape or attempted rape (includes sex while intoxicated). "The law punishes the drunk driver who kills a pedestrian. And, likewise, the law needs to be there to protect the drunk woman from the driver of the penis."
- Kosslyn, Stephen (1992?) *Wet Mind* Ref in Sacks on Edelman (NYRB 8 April 1993)
- Kroeber AL (1948, repr 1963) *Anthropology: culture patterns and processes*, New York, Harcourt Brace. Includes 1917 American Anthropologist article "The superorganic". Kroeber AL (1917, 1948, rep 1963) summarised in Barkow (1989), p232.
- Kroeber AL and Kluckhohn C (1952) *Culture: a critical review of concepts and definitions*, New York, Vintage/Random House.
- Kropotkin, Paul (1972) *Mutual Aid: a factor of evolution*, Alan Lane/Penguin, London, ed and intro by Paul Avrich. 1st ed 1902.
- Kumar, K (1989?) *Utopia and Anti-Utopia in modern times*, Blackwell, Oxford.
- Lakoff, G (1987) *Women, Fire and Dangerous Things*, U of Chicago Press. On categories, Byte AI issue.
- Laland KN (1992) "A theoretical investigation of the role of social transmission in evolution", *Ethology and Sociobiology*, 13: 87-113.
- Lennard, John (1991) *But I Digress*, OUP. History of the parenthesis (mentioned in Baker (1993)).
- Levi-Strauss, Claude, *Structural Anthropology*, Ch 1: quotation from Tylor and comments following it (Penguin ed p 4) (FP)
- Levin DA (1993) "Local speciation in plants - the rule not the exception", *Systematic Botany*, 18, 2, 197-208.
- Lloyd, G E R (1989?) *Polarity and analogy: two types of argumentation in early Greek thought*, Bristol UP, Bristol?.
- Lloyd E (1988) *The structure and confirmation of evolutionary theory*, Greenwood Press.
- Lloyd, James E (1986) "Firefly communication and deception: "Oh, what a tangled web""", in Mitchell and Thompson (eds) (1986), pp 12945.
- Lloyd, Tom (1993) *The Charity Business*, John Murray, London. D Telegraph journalist on convergence of charities and corporations. Monetisation of charity.
- Loftus EF (1979) *Eyewitness testimony*, Harvard UP, Cambridge MA.
- Lovenduski, Joni and Randall, Vicki (1993) *Contemporary feminist politics*, CUP.
- Low BS (1989) "Cross-cultural patterns in the training of children: an evolutionary perspective", *Journal of Comparative Psychology*, 103: 311-19.
- Lumsden CJ (1989) "Does culture need genes?" *Ethology and Sociobiology*, 10, 1-3.
- Lumsden CJ and Wilson EO (1981) *Genes, mind and culture*, Harvard UP, Cambridge MA.
- Lynch A and Baker AJ (1993) "A population memetics approach to cultural evolution in chaffinch song - meme diversity within populations", *American Naturalist*, 141, 4, 597-620. [Question: does increased chaffinch density lead to increased meme flow, less meme diversity?]

- MacDonald K (1989) "The plasticity of human social organization and behaviour: contextual variables and proximal mechanisms", *Ethology and Sociobiology*, 10, 1-3.
- Mackenzie, Donald (1990) *Inventing accuracy*, MIT. FP, on pointless accuracy in missile guidance systems.
- MacShane, Denis (1979) *Using the media: how to deal with the press, television and radio*, Pluto Press, London. On natural selection of memes in news, media organizations. Press officers, press releases etc as analogues of sperm competition. Interactive (sexual-selection-type) evolution of exaggeration: star system etc.
- Mamet, David (1992) *Oleanna*, Methuen. Epigraph? "What did it mean? Nothing. It was some jerk thing, some schoolkid told me, that took up room inside my head." (p 32)
- Mansfield E (1968) *Industrial research and technological innovation*, Norton. FP
- Margulis L (1981) *Symbiosis in cell evolution*, Freeman, San Francisco.
- Marx Gary T (1990) *Undercover: police surveillance in America*, University of California Press, Berkeley CA. pbk \$10.95
- Maturana, Humberto R and Varela FJ (1988) *The tree of knowledge: the biological roots of human understanding*, Shambhala (New Science Library), Boston and London.
- Rosenberg A (1982) "Are there culturegens?", *Behavioural and Brain Sciences*, 5, 1, 22-4.
- May, RM (1988) "How many species are there on earth?", *Science*, 231, p 144. (FP)
- Mayr, Ernst (1991) *Towards a new philosophy of biology: observations of an evolutionist*, Harvard UP, 564pp, \$14.95 pbk.
- Mayr, Ernst (1982) *The growth of biological thought*, Belknap/Harvard UP, Cambridge MA. OUL L 574.09. FP - Heavier than Ridley.
- Mayr, Ernst (1992) *One long argument: Charles Darwin and the genesis of modern evolutionary thought*, Allen Lane, London.
- Mayr, Ernst (ed) (1980) *Evolutionary synthesis*, OUL L 575
- Mazlish, Bruce (1993) *The Fourth Discontinuity: the co-evolution of humans and machines*, Yale UP.
- McCagg, William O and Siegelbaum, Lewis (eds) (1989) *The Disabled in the Soviet Union: past and present, theory and practice*, U Pittsburgh Press, 312p. 11 essays on mentally and physically handicapped, inc psychiatry.
- McFarland, David (1989) *Animal behaviour: psychobiology, ethology and evolution*.
- McGeorge Bundy, X and Survival (on evolution of nuclear weapons) (FP)
- McQuail, Denis and Windahl, Sven (1993, 2nd ed) *Communication models for the study of mass communication*, Longman, Harlow.
- Mikhailovsky, English source on division of labour.
- Miller, Daniel (ed) (1993) *Unwrapping Christmas*, OUP. Structuralists on Christmas. Levi Strauss on hanging of FC in Dijon cathedral 1951. Tokyo store: crucifixion of Santa Claus; happy shopping. Xmas as lying, game, fooling children, rites de passage (for the children), game compulsorily reuniting families, forcing intimacy on strangers. Time of suffering. O'Henry: giving, destroying, potlatch. Priapic US Santa Claus'. Analogy with charity.
- Miller, Delbert C (1991) *Handbook of Research Design and Social Measurement*, 5th ed, Sage, London, 720p, 29.00. Part 6: Sociometric scales inc measures of organizational structure.
- Minsky, Marvin (1989?) *The Society of Mind*, New York. Byte AI issue.

- Mitchell, Robert W and Thompson, Nicholas A (eds) (1986) *Deception: perspectives on human and nonhuman deceit*, SUNY Press.
- Modell, Arnold (1993) *The Private Self*, Harvard UP. Materialist view of psychoanalysis, and internal evolution and selection of psychic behaviours, like motor patterns (cd Esther Thelen). See for Bowlby, attachment etc. Ref in Sacks on Edelman (NYRB 8 April 1993)
- Moore J (1984) "Parasites that change the behaviour of their host", *Scientific American*, 250, 82-89.
- Moore, John A (1993) *Science as a way of knowing: the foundations of modern biology*, Harvard UP. NYRB quotes Bruce Alberts "marvellous".
- Moscovici, Serge (1993) *The Invention of Society*, Polity, Cambridge. Psychological explanations, force fields, whole and parts, action at a distance, fluid structures, post/modernism, money, Simmel, social physics, cell biology, microcosm / macrocosm, intellectual legacy as anonymous cash. Simmel, Money.
- Mumford, Lewis (1963) *Technics and civilization*, Harcourt Brace Jovanovich, NY.
- Neeter, Anna(?) "I'd sooner be me", song, arrangement and music by Pauline Bradley on Bradley (1993) audiocassette: Don't give up the fight for freedom.
- Newsweek, 25 October 1993. Features on sexual correctness, arguments in US feminist movements. Cites Rophe
- Ng, Yew Kwang [Yew Kwang-Ng] (Monash) "Welfare biology", 16 June 1993 lecture, OU Dept of Biology. Texts on mesoeconomics, welfare economics (and articles on electronics and memory research).
- Normal, Donald A (1989) *The Design of Everyday Things*, Doubleday, NY. On artefacts etc as memory.
- Nussbaum MC and Sen A (eds) *The quality of life*, OUP, 1992?. Essays by economists and philosophers.
- Nyberg, David (1993) *The Varnished Truth: truth telling and deceiving in ordinary life*, U of Chicago Press.
- Opie, Iona (1993) *The People in the Playground*, OUP. Reviewed Lurie (1993).
- Ornstein, Robert (1993) *The roots of the self*, Harper.
- Oxford surveys in evolutionary biology, edited by R. Dawkins and M. Ridley... [et al.] Vol.1- Oxford: Oxford University Press, 1984- v.; 24cm
- Pacey, Arnold (1990) *Technology in world civilisation*, MIT. FP
- Pacyga, Dominic A (1993) Polish immigrants and industrial Chicago, Ohio State University Press. Pp 48-52 on technology of the stockyards c 1900, Upton Sinclair, *The Jungle*. Gravity, dissassembling machine, replicative efficiency in killing, breeding, preparing,
- Papanek, Victor (1983) *Design for human scale*, Van Nostrand, NY.
- Parker, Geoffrey (1988r) *The Military Revolution: Military Innovation and the Rise of the West, 1500-1800*, Cambridge University Press, Cambridge, 250pp. 0 521 32607 9. Advert in LR of Books 8.12.89.
- Parker, Gillian (1993) *With this Body: caring and disability in marriage*, Open U Press. Based on 21 couples.
- Parkes, MB (1993) *Pause and effect: an introduction to the history of punctuation in the West*, University of California Press. Graceful review by Nicholas Baker (1993) "Survival of the fittest", New York Review of Books, 4 November 1993.

- Pearson L (1989?) *The architectural and social history of cooperative living*, Macmillan, London.
- Penzias, Arno (1989?) *Ideas and Information*, Simon and Schuster, New York NY.
- Petroski, Henri (1993) *The evolution of useful things: how everyday artefacts - from forks and pins to paperclips and zippers - came to be as they are*, Pavilion, pp288. Thingophiles, 3M Post-it notes, chopsticks v knife and fork, zips, tinned veal, sewing machine and safety pin (Hunt). Not form follows function but form follows failure. (Ergo success = reproduction.) Luxury is the mother of invention.
- Pickering, Andrew (ed) (1993) *Science as practice and culture*, . TLS 24.9.93: hostile anti-sociological review by L Wolpert.
- Pickering, Andrew (ed) (1991) *Science as Practice and Culture*, U of Chicago Press, pbk, 482pp.
- Pierre, Jose (ed) (1992) *Investigating Sex: Surrealist Research 1928-1932*, Verso.
- Pinguet, Maurice (1992) *Voluntary Death in Japan*, Polity, Cambridge.
- Pippard AE ( ) "Technological innovation: innovation in physics-based industry" Appendix to Swann report (FP)
- Pitt Rivers, A Lane Fox (1906) (ed JL Myers) *The Evolution of Culture and Other Essays*, Oxford UP.
- Plotkin HC (ed) (1988) *The role of behaviour in evolution*, MIT Press.
- Plotkin HC (ed) (1982) *Learning, development and culture: essays in evolutionary epistemology*, Wiley, Chichester.
- Plotkin HC (1993) *The nature of knowledge*, Allen Lane. Well rev Lewis Wolpert.
- Pomper, Philip (1970) *The Russian Revolutionary Intelligentsia*, Harland Davidon, Arlington Heights, Illinois. Chap 5 v good on populism.
- Popper, Karl (1972) *Objective knowledge: an evolutionary perspective*,
- Powell, C (1989?) *Humour in society*, Macmillan, London.
- Prigogine, Ilya, Complexity
- Pulliam HR and Dunford C (1980) *Programmed to learn: an essay on the evolution of culture*, Columbia UP, New York.
- Rawson B (ed) (1985) *The Family in Ancient Rome*, Cornell UP, Ithaca.
- Reeve HK and Sherman PW (1993) "Adaptation and the goals of evolutionary research", Quarterly Review of Biology, 68, 1, 1-32.
- Riddell, Peter (1993) *Honest opportunists: the rise of the career politician*, Hamish Hamilton. Selection, westminster, min office as f/t career.
- Ridley M (1983) *The explanation of organic diversity*, OUP.
- Ridley, Mark (1993) "The rite of the genes", Times Literary Supplement, 16 July 1993. Review-article on Barkow, Cosmides and Tooby, *The Adapted Mind*, and other works of evolutionary psychology. SSSM=tabula rasa; see also on evolution applied to medicine.
- Ridley, Mark (1985) *The problems of evolution*, OUP, Oxford. FP - introductory, brief. OUL L 575 RID.
- Ridley, Mark (1986 ) *Animal behaviour: a concise introduction*
- Ridley, Mark (1986) *Evolution and classification*, Longmans, London. OUL L 575 RID.
- Rindos, David (1985) "Darwinian selection, symbolic variation, and the evolution of culture", *Current Anthropology*, 26, 65-88. On anthropological v Darwinian concepts of evolution.

- Rogers, AR (1990) "Evolutionary economics of human reproduction", Ethology and Sociobiology, 11, 6.
- Roiphe, Katie (1993) The Morning After: sex, fear and feminism on campus, NY? Cf also Andrea Dworkin on assymetries.
- Ruelle, D?, Chance and Chaos (FP)
- Ruse, Micheal (ed) (1989) What the philosophy of biology is: essays dedicated to David Hull, Dordrecht; London: Kluwer Academic, 337p
- Ryan, Alan (1993) "Foucault's Life and Hard Times", NYRB, 8 April 1993. Rev + sum of three biographies, especially on F's own structuring (s/m) sexuality. C17 on incarceration/ institutionalisation is obverse of modern individualism? Follow up incarceration/death as punishment, deterrence, sterilization (against imitation, transmission of memes).
- Sahlins MD (1977) The use and abuse of biology: an anthropological critique of sociobiology, Tavistock. London.
- Salthe SN (1985) Evolving hierarchical systems, Columbia UP, NY.
- Samuel, Geoffrey (1990) Mind, Body and Culture, Cambridge UP. Physicist-turned-anthropologist. Tries to transcend individual/social and mind/body distinctions by deriving them from MMFs. See on socialisation procedures as stabilising cultures, memes and related ideas (last chapter). Shamanic v clerical societies. Technical v transformational processes. Mood vs external world.
- Sarat, Austin and Kearns, Thomas R (eds) Law in everyday life, U Michigan Press.
- Saviotti PP and Metcalfe JS (eds) (1991) Evolutionary theories of economic and technological change: present status and future prospects, Harwood, Reading.
- Hodgson, GM (1988) Economics and Institutions: a manifesto for a modern institutional economics, Polity, Cambridge.
- Schechner, Richard ( ) The Future of Ritual
- Scientific American (1992) Variety of school types
- Shatz, Marshall S (1989) Jan Waclaw Machajski: a radical critic of the Russian intelligentsia and socialism, U Pittsburgh Press, 272p.
- Sheldrake, Rupert (1987) The new science of life: the hypothesis of formative causation, 2nd ed, Paladin, London.
- Sherman PW (1989) "The clitoris debate and levels of analysis", Animal Behaviour, 37, 697-8.
- Simms N (1992) The Humming Tree: studies in the history of mentalities, U Illinois Press.
- Simon, Herbert A (1981) The Sciences of the Artificial, MIT Press, Cambridge MA.
- Slobodkin, LB (1993) Simplicity and complexity in games of the intellect, Harvard UP.
- Smith, John Maynard (1978) The evolution of sex
- Smith, John Maynard (1980) Genes, sex and evolution, Harvester/Wheatsheaf.
- Sperber, Dan (1985) "Anthropology and psychology: towards an epidemiology of representations", Man (ns), 20, 1, 73-89. Mental and public representation, different "catchabilities", institutions as sets of representations including own differential distribution (religion, judicial system, school? etc).
- Sperber, Dan (1985) "Anthropology and psychology: towards an epidemiology of representations", Man (ns), 20, 1, 73-89.

- Sperber, Dan (1990) The epidemiology of beliefs. In C Frazer and G Geskell (eds) Psychological studies of widespread beliefs, Oxford UP, New York.
- Sperber, Dan et al (1986) Relevance: communication and cognition
- Sperber, Dan (1985) "Anthropology and psychology: towards an epidemiology of representations", *Man* (ns), 20, 1, 73-89. Mental and public representation, different "catchabilities", institutions as sets of representations including own differential distribution (religion, judicial system, school? etc).
- Sperber, Dan (1985) Anthropological knowledge, CUP?
- Steadman, Philip (1979) The Evolution of Design: biological analogy in architecture and the applied arts, CUP.
- Steklis HD and Whiteman CH (1989) "Loss of estrus in human evolution: two many answers, too few questions", *Ethology and Sociobiology*, 10, 6.
- Studd MV and Gattiker UE (1991) "The evolutionary psychology of sexual harassment in organizations", *Ethology and Sociobiology*, 12, 4. Importance of informal evasion, ergo seniority.
- Symons D (1989) "A critique of Darwinian anthropology", *Ethology and Sociobiology*, 10, 1-3.
- Tannhill, Reay (1989) Food in history, new ed, Crown, NY.
- The National Interest, Special issue, Spring 1993, The strange death of Soviet Communism: an autopsy. (Fukuyama, Reddaway, Pipes etc)
- Thompson, D'Arcy Wentworth (1961) (ed JT Bonner) On Growth and Form, Cambridge UP.
- Thorne, Barrie (1993) Gender Play: girls and boys in school, Rutgers UP. Reviewed Lurie (1993). Ethnography of Borderwork. "Cootie" low-status infection, contamination, ritual cleansing.
- Tinbergen, Niko (1990, 2nd ed) Social behaviour in animals: with special reference to vertebrates, with a new foreword by G.P. Baerends, London: Chapman and Hall.
- Tipton, Steven M. (1982) Getting Saved From the Sixties, University of California Press, London.
- Luce, R Duncan, Smelser, Neil J, and Gerstein, Dean R (eds) (1990) Leading edges in social and behavioural science, Russell Sage Foundation, Ithaca NY. \$59.95 720p
- Tooby J and Cosmides L (1989) "Evolutionary psychology and the generation of culture...", Parts I and II, *Ethology and Sociobiology*, 10, 1-3.
- Toulmin, Stephen (1972) Human Understanding, OUP, Oxford.
- Turner Patricia A (1993) I heard it through the grapevine: rumour in African-American culture, U of California Press.
- Twitchell, James B (1993) Carnival Culture: the trashing of taste in America, Columbia UP.
- van Gereld?, Martin (1989) Technology and War, Free Press, NY. (FP)
- van Hooff, Anton J L (1990) From Autothanasia to Suicide: self-killing in classical antiquity, Routledge, London. 306 pp.
- Van Valen (1989) "Three paradigms of evolution", *Evolutionary Theory*, 9: 1-17. 3 paradigms: standard, energy, information. Does evolution max/minimise anything? If so what, and over when? Genes code
- Varela FJ, Thompson E and Rosch E (1992?) The embodied mind: cognitive science and human experience, MIT Press.

- Veblen TB (1899) *The economic theory of the leisure class*, Macmillan, NY.
- Veblen TB (1919) *The place of science in modern civilization and other essays*, NY (reprinted 1990)
- Vining DR (1986) "Social vs reproductive success", *Behaviour and Brain Sciences*, 9, 167-216.
- Voland, E and R (1989) "Evolutionary biology and psychiatry: the case of anorexia nervosa", *Ethology and Sociobiology*, 10, 4. Contraceptive response to stress.
- Ward PJ (1989) "A model of end-directedness", in KT Hahlweg and CA Hooker (eds) (1989) *Issues in evolutionary epistemology*, SUNY Press, Albany.
- Watson, Alan Roman Slave Law, Johns Hopkins UP. Axised on contrast of African-American and Roman slavery. On enslavement as a meme, parallels with wives, concubines, children, animals, objects. Intention of slave and liability of master. (More generally: legal etc memes which inhere in rights, duties, property. If Dawkins' definition of a meme as the smallest distinguishable unit of cultural transmission applies, juridical and moral properties (transmitted between people, things, and relationships) would qualify. Enslavement/manumission by: transfer, debt, capture, purchase, birth/conception, contract, through law (crime). See also on parenthood and children. Plus refs on Roman family, ritual, weighing of bronze etc accompanying monetary sale in primitive roman jurisdiction. See also forward on biological inferiority of American (black) slaves, Virginia law (1848) on blacks uniting for literacy or education.
- Wells, A (1989?) *The United Nations, UNESCO and the politics of knowledge*, Macmillan, London.
- Werren JH and Pulliam HR (1981) "An intergenerational model of the evolution of helping behaviour", *Human Ecology*, 9, 465-83.
- White, Leslie (1975) "Culturology" - summarised in Barkow (1989), p232.
- White, LA (1959) *The Evolution of Culture*, McGraw-Hill, New York.
- White, S, Evans, P, Mihill C, Tysoe M (1993) *Hitting the Headlines: a practical guide to the media*, British Psychological Society Books, Leicester. Getting your journal article, conference etc into the media. Foreword by Wolpert.
- White, LA (1975) *The Concept of Cultural Systems: a key to understanding tribes and nations*, Columbia UP, NY.
- Whittley, R (1985) *The intellectual and social organisation of the sciences*, OUP.
- Williams, Ian (1989) *The Alms Trade: charities past, present and future*, Unwin Hyman, London.
- Williams, George C (1975) *Sex and Evolution*
- Williams, George C (1966) *Adaptation and Natural Selection*
- Williams, George C (1992) *Natural Selection: domains, levels and challenges*, OUP. Rev Mark Ridley Nature 3 June 1993: clades; stability of fossil record; micro/macro-evolution; unobserved adaptations.
- Williams, Bernard (1993) *Shame and necessity*, U of California Press.
- Wilson, James Q (1990?) *Bureaucracy: what government agencies do and why they do it*, Basic Books, New York NY. \$24.95. Explains US state bureaucracy as product of American hostility to red tape and the state.
- Windrow, Martin (1985) *The Greek Hoplite*, 32p
- Wolpert, Lewis (1992) *The triumph of the embryo*, OUP, Oxford.
- Wolpert, Lewis (1993?) *The unnatural nature of science*.

- Wright R (1988) Three Scientists and their Gods, Harper Collins. Rec by Mark, one is EO Wilson. Includes account of gene-culture theory, how EOW crafted other ideas.
- Wynne-Edwards VC (1986) Evolution through group selection, Blackwell, Oxford.
- Wynne-Edwards VC (1962) Animal dispersion in relation to social behaviour, Oliver and Boyd, Edinburgh.
- Yew Kwang-Ng (Monash) "Welfare biology", 16 June 1993 lecture, OU Dept of Biology.

# WHAT'S A MEME: REFLECTIONS FROM THE PERSPECTIVE OF THE HISTORY AND PHILOSOPHY OF EVOLUTIONARY BIOLOGY

By

John S. Wilkins

3 Peel Grove, Mt. Martha 3934, Australia.

## Abstract

This paper is intended as a focal article to raise philosophical issues about the nature of memes and memetic theory. To bring consistency to memetic analysis, researchers need to understand and agree upon the theoretical role of memes and the generalized model of evolution in which it occurs as a theoretical term. To help this, I have traced the source of Dawkins' conception of memes from GC Williams' *evolutionary gene* and through to the Hull-Dawkins Distinction between *replicators* and *interactors* and Hull's notion of *lineages* and the idea of an individual in biology. The complexity of biological modes of evolution suggests that conceptualizing memes as disease pathogens is not an alternative to evolutionary models of memetic development. I argue for a close and strict analogy between biology and memetics. I introduce the idea of a *memetic individual* or *profile* to clarify the ontology of memes and their ecologies. Some promising methods from biology and other disciplines such as Hamming Distance and Wagner ground plan divergence methods are suggested. A glossary of mainly biological technical terms used and introduced neologisms is included.

**Keywords:** Meme, evolution, gene, replicator, interactor, individual, natural selection, term, mnemone, pheme, deme, ecology, epidemiology, mind virus, Lamarckism, Dawkins, Hull, Campbell, Williams

JULIET: 'Tis but thy name that is my enemy;  
Thou art thyself, though not a Montague.  
What's Montague? it is nor hand, nor foot,  
Nor arm, nor face, nor any other part  
Belonging to a man. O, be some other name!  
What's in a name? that which we call a rose  
By any other name would smell as sweet;  
So Romeo would, were he not Romeo call'd,  
Retain that dear perfection which he owes  
Without that title. Romeo, doff thy name,  
And for that name which is no part of thee  
Take all myself.

(William Shakespeare, *Romeo and Juliet*, Act 2, Scene 2)

## 1 Introduction

The fundamental theoretical concept of memetics is the meme<sup>1</sup> - the unit of cultural evolution and selection. The term is unclear in its meaning and what it denotes, and the application of evolution to culture is often based on a partial or even mistaken notion of the general structure of evolutionary explanation. In this paper I give some historical background to the origin and function of the term "meme" from the debates in evolutionary theory out of which it sprang. I argue that there is a close analogy between the entities and processes of biology and culture, although the parameters that describe each domain often differ in their rate and frequency, and perhaps in the dynamics of the evolutionary process they undergo. It is as crucial to understand what the units of evolution in culture are as it is in biology to be able to apply an evolutionary explanation, and for this reason we must be quite clear about what a cultural individual is. We can make better sense of these issues if we look first at the analogous problems of what a biological individual is, especially in the context of what the units of selection are.

Problems that bedevil memetics include classifying cultural entities and reconstructing historical developments, and these also have their analogues in biology. Resolving them depends on determining the level at which selection occurs, for it is selection that defines what is functionally important in evolution, not the contrary. The conceptual tool of most help to us here is the distinction between replicators and interactors. There are three common approaches in memetic writings - most commonly writers picture meme evolution as the spread of disease pathogens, others restrict selectionist accounts to some part of culture that is amenable to this form of explanation such as science, and I argue for a third option. We must more broadly understand the variety of biological evolutionary phenomena, so that we can see the similar patterns in cultural change. In order to assist in this, I introduce the notion of a memetic individual, and discuss the relationship this entity has with the biological organism in which it resides. In the conclusion and prospectus, I suggest some analytical techniques of biology, particularly methods of taxonomy and genetics, which these distinctions permit for memetic research.

I hope both that biologically literate readers will understand my simplifications of biology and that those not familiar with biological terminology its necessity. I append a glossary of various technical terms introduced in this paper and those that are likely to be unfamiliar to non-biologists.

## 2 The background to the problem

*The increase of the conceptual clarity of a theory through careful clarifications and specifications is, as William Whewell observed more than a century ago, one of the most important ways in which science progresses. He called this process "the explication of conceptions" and showed how a number of theories, in the course of their temporal careers, had become increasingly precise - largely as a result of the critics of such theories emphasizing their conceptual unclarities. Many important scientific revolutions ... have depended largely on the recognition, and subsequent reduction, of the terminological ambiguity of theories.* (Laudan 1977:50)

Certain terms and notions in both the sciences and humanities are fated to be misunderstood, either because they are first vaguely formulated, or because they are so evocative they generate such immense enthusiasm and are applied to almost everything, coming to mean almost nothing. A classic example is the term of Thomas Kuhn's (1962): *paradigm*. Originally intended by Kuhn to apply to what changed radically in a scientific revolution, it came to be applied to perceptual and conceptual changes in cases of individual, social, literary, political, economic and even consumer choice. When a term of philosophy of science is used to advertise a new car design, you know it has lost any definite meaning. Eventually its author abandoned it under criticism in favour of notions and terms that were more specific, but "paradigm" is now ensconced in popular parlance, surviving both author and intended theoretical usage. The difficulty now with the term for a specialist in the philosophy and history of science is that calling a theoretical change a "paradigm shift" has become little more than a metaphor. It describes only an impression and implies only a subjective assessment.

The basic and central notion of memetics is, of course, denoted by the term *meme*, Richard Dawkins' (1977) term for what is transmitted in culture that is analogous to the biological gene. "Meme" is in danger of suffering the same fate as "paradigm". It is used to denote, variously, neural structures, cultural artefacts, practices, economic systems, religions, concepts, phenotypic traits, self-awareness, and epigenetic predispositions. Memes are thought by some to control behaviour, by others to be acquired through a choice or act of will. The term gets applied to all levels of social and cultural structure, from minimal semantic entities like phonemes, through more molecular entities like phrases and snatches of music, to entire traditions and world views. In this blooming buzzing confusion, the usefulness of memes as a category is being lost or degraded.

It wouldn't be the first time this sort of confusion has wrecked an emerging discipline - gene, the very term on which "meme" is modelled, has a history of shifting definition and debates in which antagonists argued past each other for more than sixty years. It took the modern evolutionary synthesis and the discovery of the structure of DNA to resolve it fully, and still, "gene" covers a range of theoretical phenomena requiring the more exact terminology of *codons*, *cistrons*, *introns*, *exons*, *operons*, *regulators* and *base pairs*. Even today, biologists use this term differently and inconsistently - what molecular biologists usually mean by it differs greatly from its meaning in population genetics.

A theoretical term is usually generated to denote a causal nexus in the model the theory describes<sup>2</sup>. On at least one recent account (Suppe 1989, van Fraassen 1980) a scientific theory is an attempt to either isolate or idealize a system - usually a physical system - in such a way that its dynamics can be reduced to a manageable number of variables (each of which is usually represented by a theoretical term) related by a mathematical description, so that the model generates a restricted number of likely outcome states. It is often neutral with respect to many of the attributes of the entities and processes it covers: Newton's theory of falling bodies does not mention their colour, for example; the equations of flow dynamics cover water, hydrocarbon liquids, and gases; and information theory covers transmission of bits by electromagnetic radiation no matter whether the medium is electrical, photic or magnetic.

One such generalized scientific theory is the theory of adaptation and evolution by natural selection. Natural selection is a process that occurs over a range of physical substrate entities: viruses, single celled clonal organisms, multicellular plants, animals and fungi, and so forth. Certain variables need to be adjusted to deal with these very different phenomena - life cycle rates, ecological behaviour, reproduction modes and so forth - but the virtue of the theory as a model is that it is highly generalizable and widely applicable. Recently, for example, natural selection models have been applied through the use of computers to a range of real world problems, and have been developed into a class of formal algorithms called Genetic Algorithms (cf. Holland 1995). Problems to which Genetic Algorithms have been applied include political and social problems, through to hard engineering of complex systems like engines and flow dynamics.

To understand the meaning of *meme*, we need to understand the history of some debates in evolutionary biology, and the ways in which evolution and culture have been thought to be related. Even before Darwin published the *Origin* in 1859, evolutionary reconceptualisations of human society and culture were common. The Lamarckian evolutionary tradition through Geoffroy, Grant and Chambers had drawn radical social conclusions based on their view of evolution as progressivist, universal and hierarchical (Desmond 1989). Not long before the *Origin*, Herbert Spencer developed his form of Hobbesian rugged individualism based on a Malthusian struggle for existence, ideas that found a parallel in Darwin's natural history work. But of interest to a memeticist is the application of natural selection models to ideas, by writers like Alexander Bain, Charles S Peirce, William James, Ernst Mach, and John Dewey, very soon after Darwinism was first promulgated through European intellectual society. In many ways, Dewey is one of the first true memeticists, declaring that we didn't so much solve problems as recover from them, appropriating both the Darwinian evolutionary metaphor and the epidemiological metaphor so popular today (Dewey 1909). Over the years, various others have used natural selection as a metaphor, or as an analogy (the distinction is significant<sup>3</sup>), of conceptual and social change, including Medawar, Popper, Toulmin, and even some leading Darwinians (cf. Cziko 1995 for a review).

Eugenics was a constant source of confusion and debate in the early synthesis, with Fisher and the "Oxford School" along with leading geneticists like Muller proposing that the primary role of natural selection was to trim away deleterious alleles and that the best way to organize society was to permit that process to act without impediment. Dobzhansky and others, largely in North America, argued instead that the role of natural selection was to maintain a pool of variety in populations, and that equality of opportunity would permit whatever alleles had relative advantage to be expressed (Beatty 1987). This was more significant than a clash of political or moral doctrines. Eugenicists (Darwinian eugenics was usually positive - encourage the fitter - than negative - cull the less fit - as in Spencerian social "Darwinism") held that society, or at any rate "civilized" society, acted to lessen the effects of natural selection, while those who like Dobzhansky opposed eugenics tended to see society as either neutral or ancillary to natural selection in biology. The very role of natural selection itself was the subject of intense debate, culminating during the final years of the synthesis and the early 50s. It centred on an extended debate between Fisherians, who saw it as a universal and often all-sufficient mechanism of evolutionary change, and those who like Dobzhansky followed Sewall Wright in thinking that in small populations (demes) random drift and sampling error played a greater or lesser role in the

evolution of traits. Even today this is still an issue. Dobzhansky, who through his early education in the tradition of Russian Darwinism which emphasized co-operation and began with Kropotkin, was well placed to argue for Sewall Wright's notions and by the end of the 1950s, some admixture of drift and selection was accepted by nearly every evolutionary biologist.

The point of this potted history is to show the context in which the meme-analogue "gene" arose and was debated, and to set the scene for the conditions in which memes were developed as theoretical entities; this will go to the question of the nature of memes and their theoretical role in explanation. Throughout the period of the development of Darwinism, it was agreed that there had to be hereditary factors of some kind because natural selection can only act if the traits of organisms that are selected are transmitted to future generations in a process of differential reproductive success based on what we now call differential ecological success. But it took a long time for this to be clarified. Evolutionists often spoke of new traits being "for the good of the species" even though Fisherian genetics modelled fitness as an *individual* property of heritable traits in a population. Indeed, Fisherian genetics treated genes in such an unlinked atomistic way that later writers like Mayr spoke derisively of "bean-bag" genetics, which treated genes as if any single one of them could be separated out from the others and inspected.

Although Mendelian genetics had already realized that genes were often linked in their effects, and that many genes affected each trait (epistasis) and some genes could affect many traits (pleiotropy), it was unclear what natural selection acted *on* - individual genes or the traits they affected. In his 1937, Dobzhansky adopted Darwin's term "co-adaptation" to denote parcels of genes that were selected as a unit because they had evolved together to work well as a team, but there was a real epistemological difficulty in determining what genes selection actually selected.

In 1962, the modern "units of selection" debate was inadvertently kicked off by a naturalist, Wynne-Edwards, who explained the regulation by some gulls of clutch size (the number of eggs laid) during lean times as an adaptation that could only be for the benefit of the group, because individual fitness had to be lower if fewer progeny were produced. Wynne-Edwards used the notion of "group selection" which Darwin had used in 1872 to explain the existence of human moral behaviours: moral groups did better than immoral groups, and so their genes would be better represented later on than those of the others (to reinterpret Darwin's formulation in modern terms). According to Wynne-Edwards, populations that regulated clutch size survived ecological catastrophes better than those that didn't. It's an intuitively seductive idea, but this sort of reasoning was thought to be woolly headed and wrong by those evolutionary theorists raised in the more reductive panselectionist tradition. David Lack argued against the biological specifics of Wynne-Edwards' thesis, but a pivotal outcome of this book was that it spurred George Williams to publish, in 1966, *Adaptation and Natural Selection*.

Williams took an operationally reductive approach; that is, he refused to use selectionist explanations at any level higher in the biological hierarchy than was necessary. It is onerous, he thought, to explain group level phenomena in *sui generis* terms when explanation in terms of the components of the group and their properties will suffice. For selection of a group to occur (and contrary to popular reportage, Williams did not rule out the in-principle possibility of group selection, even identifying a case study), two criteria would need to be satisfied. One, the properties that were favourably selected would need to be the properties of the *group*, not the

individuals that comprised it. Williams illustrated the difference by distinguishing between a group of fleet deer (which as a group may move slowly, though all its parts are fast moving) and a fleet group of deer (which may move quickly, as groups go, although its parts are slower than parts of slower groups). Fleet deer are individuals with properties of fleetness. Fleet groups of deer are *groups* with properties of fleetness. The second criterion that must be satisfied is that group properties had to be heritable: copies must be made if selection is to operate. Therein lay the rub for Williams: groups typically do not reproduce. But Williams went even further: neither are individual organisms usually copied. You are not a clone of your same sex parent. You may have your father's eyes, your mother's hair colour and wave, and be intermediate between them in your genetic predisposition to height. Mendelian factors are the basic unit of evolution, according to Williams, and not all of them, either. "I use the term 'gene' to mean that which separates and recombines with appreciable frequency", he said (1966: 20) and he further defined an evolutionary gene (that is, one that is selectively important) as

"...any inherited information for which there is a favorable or unfavorable selection bias equal to several or many times its rate of endogenous change" (1966: 25).

In set theoretic terms: Mendelian factors (M) are a proper subset of heritable separating and recombining units of information (G), and include as a subset selectively biased - that is, evolutionary (E) – factors.

Ten years later, Richard Dawkins (1977) introduced the "selfish gene" metaphor based on Williams' "evolutionary gene". But the way he did so obscured the point Williams was making (not deliberately, for Dawkins had other fish to fry) - not all genes are evolutionary genes. Given that only genes can be evolutionarily relevant, what portion of those genes actually *are*?

Williams is a bit of a panselectionist<sup>4</sup> and he tends here to brush past questions of random drift and founder effects. His answer to the question is that a gene is evolutionary just insofar as it is subject to selection that exceeds mutation. This is a definition, not a discovery. If it's favourably or unfavourably selected and heritable, it's an evolutionary gene. We are now on the threshold of memes. Williams himself realized this, noting that gene was a cybernetic abstraction, and later describing the notion of a "codical domain" of evolution, wherein the transmitted structure, no matter what its *physical* substrate, is a codex (Williams 1992). In other words, it's a message that is transmitted, and which is subjected to the same theoretical models as any other kind of message transmission, i.e., Shannon-Weaver information constraints. This means that genetic replication is one instance of a class of phenomena in Message World - and memes, like genes, are another.

Dawkins' original introduction of the term "meme" in *The Selfish Gene* mentioned in passing snatches of tunes, crazes and fads, but the paradigmatic example he gave, no doubt due to his personal experience of it, was a scientific notion passed from scientist to scientist. I shall return to this point later, but some properties of scientific notions are of immediate interest. Typically, scientific ideas are either evocative metaphors, like de Candolle's "struggle for existence", that inspired Darwin, or more or less formal models. It is the latter that concerns us here, for when metaphors reach the end of their evocation, they must be turned into formal models anyway in order to be tested against quantifiable phenomena. A formal model like Boltzmann's

thermodynamic entropy is a far cry from Heraclitus' notion of flux, and it does a great deal more conceptual work. The significance of Dawkins' example is that one can, to a relative degree of exactness, determine whether and how far apart or whole of a model has spread to another scientist or textbook, or whatever one takes to be the cultural equivalent of the phenotype - for reasons I hope become clear, I shall refer to this as the phemotype, and the total distribution of co-adapted memetically constituted traits within a lineage the phemorph of that lineage. The neologisms are strained, barbarous and ugly, but I hope they will add some clarity to what is being discussed. In short, scientific examples can be quantified both in terms of their frequency in a lineage of scientists, and their relative rates of increase or decrease. This susceptibility to analysis is essential for modelling change in terms of natural selection, and evolution in general.

### 3 The Hull-Dawkins Distinction and the evolutionary gene

*Do I contradict myself?*

*Very well, I contradict myself,*

*(I am large, I contain multitudes.)*

Walt Whitman, *Song of the Open Road*

To apply evolutionary models outside biology, you need to know what evolution as such is, no matter what the domain it is occurring in. There have been many attempts within biology to generalize evolutionary modelling, for it is a point worth remembering that even biological evolution is not all of a kind. Evolution occurs over single and multicellular organisms, over animals and plants and fungi, and over various kinds of bacteria and viruses. It applies to sexual, clonal and parthenogenic reproduction, to heredity using DNA, RNA and even cytological structure. Generalist species (*eurytopes*) and specialist species (*stenotypes*) both evolve in widespread and continuous ecologies, and in localized and isolated ecologies, and so on.

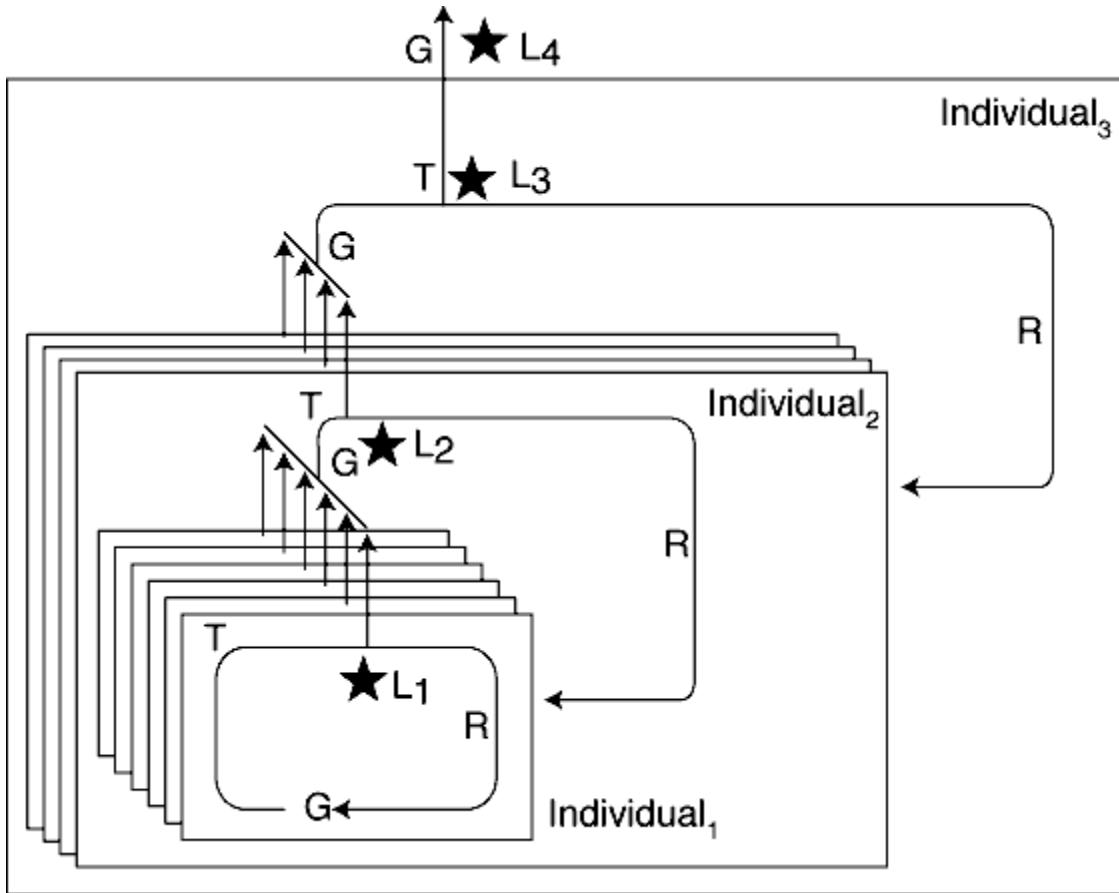
Evolution is not simple in biology, and so biologists have worked to make it as general a model as possible to cover all cases. There is clearly some common thread in all these phenomena, and most find some or all of the explanation in natural selection<sup>5</sup>. As Weismann's and Wallace's heirs perceive natural selection, it is the differential success of the genotype due to the ecological success of the phenotype it creates. The causal arrow is essentially one way: genes cause phenes but not the reverse. As Cziko (1995), following Donald Campbell, puts it, genes are selected but not instructed. Despite all the progress since Weismann on the nature of biological heredity, his Central Dogma has stood almost without modification<sup>6</sup>. It is therefore important to determine whether there is an analogue in culture for these central distinctions of genotype and phenotype and whether memes are ever instructed, or only ever selected, as genes are. And if memes are instructed, does this "Lamarckism"<sup>7</sup> of culture obviate the need for a memetic analysis?

Although it's worth remembering that Darwin was a Lamarckian on heredity, that is, he accepted that the units of heredity ("pangenes") are instructed, this does not really help us answer the question, for we already have non-evolutionary Lamarckian models of culture and the justification for taking an evolutionary perspective rests on the efficacy of Darwinian, perhaps Weismannian, models. We need to understand the ontology of a Darwinian process to come to any resolution.

Dawkins (1982) distinguished between genes as replicators - things that are duplicated with a high degree of fidelity - and bodies of organisms as the *vehicles* of genes. Vehicles reproduce,

but only genes replicate. Some writers, including David Hull (1980, 1987, 1988a, 1988b, 1988c) took exception to Dawkins' characterization of bodies as mere vehicles of genes - they were, in his opinion, much more than passive and controlled robots<sup>8</sup>. Hull had earlier (1974) attacked explanatory reductionism in genetics, arguing that reducing evolutionarily significant traits to genes was unsatisfactory both in terms of a failure of isomorphic relations (mapping) and also of underdetermination - things happen at metagenic levels that are not, and never could be, the results of the sum of the properties of the genes<sup>9</sup>. He therefore rejected Dawkins' passive notion of bodies and the phenotype, and substituted instead the more ecologically and economically active notion of an interactor. Hull's general view of evolution is of a cycle of replicators coding for interactive traits, which through their interactive success acquire (or fail to acquire) the resources needed for further replication. On Hull's account, ecology makes a strong comeback, and the reproduction of organisms is a necessarily *interactive*, rather than a *replicative*, process. This cycle generates lineages of descent, which exhibit themselves in biology as species in the short term and as phylogenies and other taxonomic classes in the longer term<sup>10</sup>.

Hull and Ghiselin, as part of their *individuality thesis* that species are historical individuals, that happen to be made up of many organisms, have also addressed the general problem of what constitutes an individual in biology and evolution in general (Hull 1987, 1988c, 1992, Ghiselin 1997). This also raises its head in memetics, and has not so far been addressed. Their view is that a biological individual, so far from being intuitively obvious (as it seems to be if we only consider obligately sexual vertebrates) is in fact a very fluid category, and their solution is that an individual is *defined* by its functional role in evolution, that is, insofar as it is a unit of selection (this applies to gene sequences as well as to organisms). So, the resultant evolutionary ontology is that replicators (replicanda in Ghiselin's terminology - structures that are replicated, Ghiselin 1987) generate interactors, which are the evolutionary individuals that are subjected to selection and whose economic success or failure biases the regeneration of the replicanda. Plotkin (1994) characterized generalized selective processes as *generate-test-regenerate cycles*, and this helps to clarify the matter (figure 1). An individual is formed at the level of testing, which feeds back into the generation stage. An inclusive hierarchy of individuals can be formed if generation to regeneration cycles occur at higher levels<sup>11</sup>. In biological terms, this accounts for why colonies of organisms like eusocial insects behave as individuals with respect to selection, and also why clonal lineages (genets) behave evolutionarily as individuals even if they are spatially distinct (ramets).



**Figure 1.** A generalized hierarchy of g-t-r (generate-test-regenerate) cycles in an evolutionary process. Star = Replicators relative to testing/selection bias at that level or higher

The Hull-Dawkins Distinction, as it came to be called, gained almost immediate and bipartisan support across the selectionist debate - both Eldredge (1989) and Williams (1992) accepting that replicators and interactors are the active and general entities in Darwinian evolution (interestingly, both made the now almost obligatory passing comments about these terms also applying to culture.) Dawkins' "memes" (like Campbell's mnemonics, 1974, 1988) are *cultural* replicators, which, if they are to function in cultural evolution analogously to genes, must be transmitted with fidelity, and must cause some interactive traits that in turn will cause a differential replication of the memes. This also helps us with the question whether memes are instructed or selected; that is, whether they arise in response to environmental needs (are "learnt") or are generated randomly with respect to the prevailing social ecology ("random trial and error"). Unexpressed memes (memes not "visible" to the environment through their products) cannot be selected, and so the likelihood of them being prescient or anticipatory is reduced, since we would need to have *already* selected some memes as likely candidates for success in order to predict which ones will work in practice<sup>12</sup>. Before a meme can be assessed as "likely to succeed", it must already have passed some tests. We therefore get a regress - any "instruction" of a meme is either a case of transmission followed by selection, or it is a case of transmission of an already selectively tested meme<sup>13</sup>. The Central Dogma remains unshaken for memes, even if some mechanisms of instruction are shown to occur, for even learning is a

selection process (Cziko 1995, chapters 11-12) at some level.

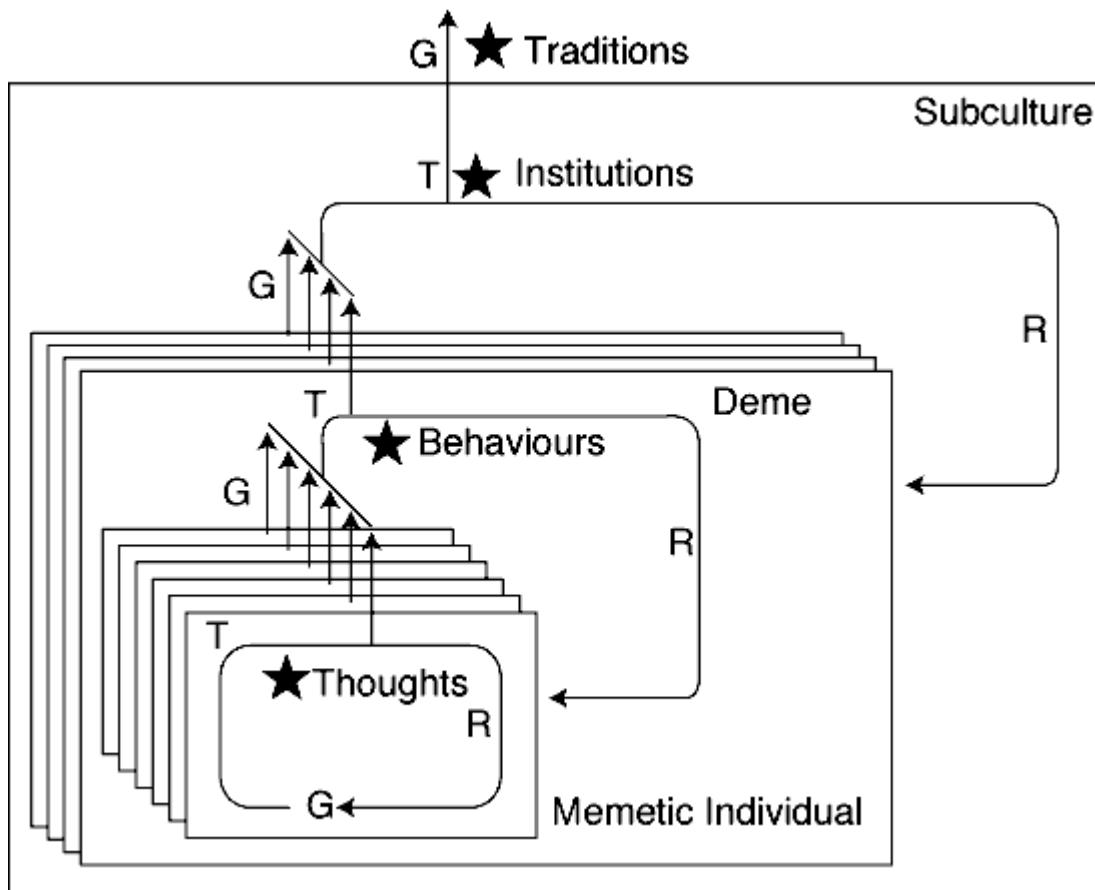
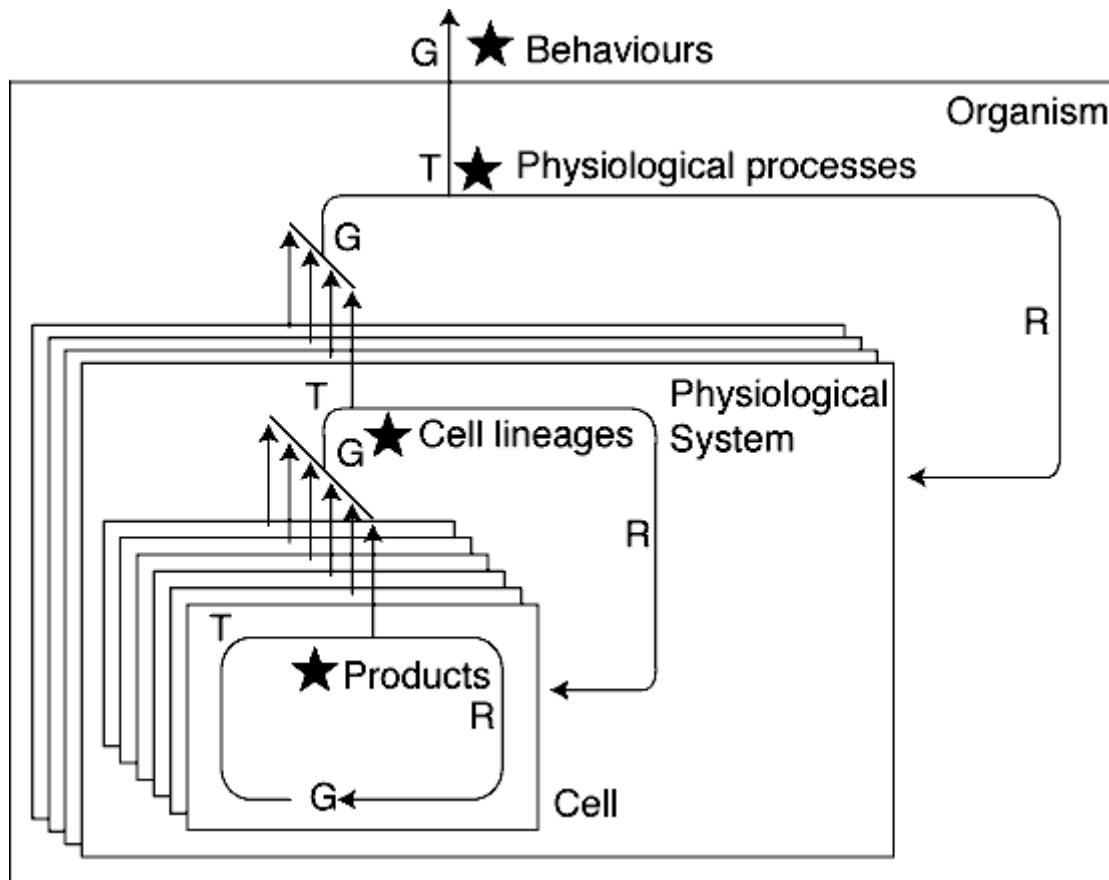


Figure 2. Memes arise as effects of a g-t-r cycle. Star = Meme relative to testing/selection bias at that level or higher

Some may object that memes must be *stored* in neural patterns in order to be considered memes, and that replication through texts, electronic media, and performances, etc., are only secondarily memetic if they affect and are stored in a central nervous system. In figure 2, it can be seen that some selectively biased cultural phenomena can exist as the emergent properties of social systems larger than the individual, and persist without being replicated at time frames far longer than an individual life span. Methodological individualism, as it is called, rejects this in favour of a more parsimonious view, just as genic selectionists like Dawkins reject any form of group selection (Williams' views notwithstanding). But on the Ghiselin-Hull view of individuality in selection, group selection is not only a possibility but an established reality, for viewed appropriately, any organism is a group of related lineages (Buss 1987) and any gene sequence a related group of molecules, and so on. Consider an historical event like the Thirty Years War or an institution like a religion. Political and religious patterns of behaviour in that war need not be stored as neural patterns, and each of the memes of a religion need not reside in at least one brain, for some of these memes are emergent properties of the entire system of acting individuals, and of which they may indeed all be entirely unaware<sup>14</sup>. Were Protestants aware of the linkage between capitalism and their theology during the rise of modern capitalist economies, for example? If so, why did it take a sociologist to point it out explicitly? You pays your money and you takes your choice, for the methodological individualist debate has a long history of

controversy in sociological disciplines that continues till this day. However, the view I am advocating here is neither individualist, nor holist, but a view known as "emergentism" (Nagel 1961): the doctrine that the properties of a collective whole arise from the *relationships* between the properties of the components. Simply understanding the componential properties, without understanding the connections between them does not enable us to model the higher level thing they comprise.

So, what are phenotypic cultural traits, the *phemotypes* as I have called them? One answer, and the right one I think, has been given variously by Campbell (1960, 1974, 1987, 1988), Hull (1988c), Toulmin (1972) and Plotkin (1994), to name a few, is that selection acts on behaviour, or, in the language of the logician, on interpretation of the information contained in a meme. Just as a gene must be expressed in order to be selectively biased, a meme must also find expression in some way. While we often use a verbal shorthand when we say that a gene is selected at a certain coefficient, the biologist should never forget that genes are just the starting point of a biochemical process which undergoes selection at a range of hierarchical levels, from the processes of transcription in the cell to the processes of ecological interaction in an ecosystem (figure 3).



**Figure 3.** Individuals and replicators in a biological hierarchy

## Sidebar: interactors and replicators

The terms interactor and replicator, being substantive nouns, give the impression that there are natural kinds of *things* that interact and replicate. Ghiselin (1987, 1997: 147) has argued that a better term for the Hull-Dawkins Distinction would have been between *replicanda* and interactors, that is, between those things that *get* replicated and those things that are economically biased for a range of reasons. When we understand that replication is the passing on of a message, it doesn't really matter what the medium, or substrate, of the message is, and the notion that there has to be some privileged replicator, like nucleotides or neural columns, becomes unnecessary. It is less obvious that interactors are also a functionally-defined class, but if anything interacts (and correlates closely with replicators) then it is an interactor. It could be the organism, but it could be some subsystem of an organism like a class of immunological peptide, or the immune system itself, or a visual system, or it could even be a group of organisms like a hive. There are no privileged interactors *or* replicators, in biology or culture. We would do well to remember Mill's caution:

*"The tendency has always been strong to believe that whatever received a name must be an entity or being, having an independent existence of its own. And if no real entity answering to the name could be found, men did not for that reason suppose that none existed, but that it was something peculiarly abstruse and mysterious."*

(Cited in Gould 1981: 320)

This modifies Hull's notion of interactor somewhat, for Hull defines them as things, that is, as entities. Plotkin, on the other hand, resists the "entification" of interactors (personal communication). I do not think that Hull's own analysis requires that interactors be a unique class of entity, such as an organism, just as long as it is an *individual* (Hull 1992) with interactive (ecological, or economic) properties. Individuals are entities, but there is no need to restrict individuality to one level of entitativity, so long as the relevant interactive properties are properties of that entity and not of the system of which it is a component. However, interactors are entities even if they are not a specific *kind* of entity.

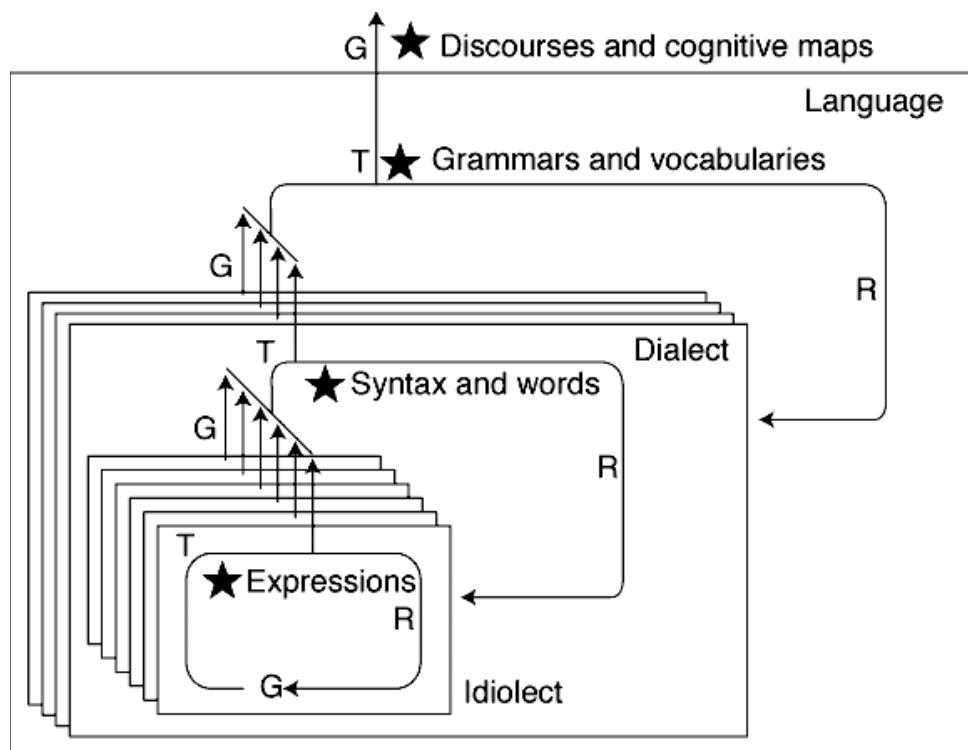
Similarly, memes are selected at many levels through their expression in behaviour, including verbal, practical, instrumental and intellectual, and this behaviour need not be the behaviour of individuals; it can be the behaviour of languages, institutions, societies and even traditions. Consider Campbell's (1974) hierarchy of selectionist knowledge processes<sup>15</sup>:

1. Genetic adaptation
2. Nonmnemonic problem solving
3. Vicarious locomotor devices
4. Instinct
5. Habit
6. Visually supported thought
7. Mnemonically supported thought; Observational learning and imitation
8. Socially vicarious exploration
9. Language

## 10. Cultural cumulation

## 11. Science

At level 7, minds construct models to test and select alternatives - and at this level memes are now active and give rise to trial and error without the selective costs to genetically- and biologically based interaction. Memes must result in some outcome that can achieve success or not, on the basis of which they persist over other memes, to be selected at all. The details of Campbell's schema are not important here, but memes *must* arise at some level of information processing complexity, and *must* affect outcomes at levels higher than that. And the converse is also true. Of any cultural process that involves transmission and selection, there *must* be memes that "constitute" phenotypes at that level. Take two examples - a personal attempt to try out a hitherto unlearnt linguistic form and a scientific theory. When one is learning a first language, syntax and semantics are underdetermined. A child cannot know, except through trying out combinations and seeing the results, whether a certain "hypothesis" about the use, structure and referents of a word, phrase or syntactical structure is correct. For example, consider how children first learn regular declensions of noun plurals and regular conjugations of verb tenses, and are then corrected by more competent speakers when they use a regular form of an irregular noun or verb - "I swummed", "he runned", and so forth. Not only is there an *internal* selection process going on through increasing experience of success, but in a formal sense, these irregular abstract constructs of language are *defined* through the role that they play in the trial and error. Of course, language in itself is not a solitary pursuit, but in *individual* learning, the trial and error selection process is (figure 4, cf. Cziko 1995, chapter 11, Ghiselin 1997, chapter 9).



**Figure 4.** A selective hierarchy of language, following Ghiselin 1997

Back to Dawkins' scientific notions: whatever Mendel, or Darwin, or Newton, internally intended their theories to mean or cover, until the elements of their theories - supporting data, formalizations, explanatory rationales and so forth - were *published* (literally, made public), discussed, tested, and above all in further scientific work, they weren't scientific theories as such. What's more, their theories contained memes of which they were not aware, because they only developed when selection pressures were applied to them. "Action at a distance", or "survival of the fittest", became memes so far as they needed to be defended and supported in the scientific debate when challenged. It is the hallmark and defining characteristic of science that it is presented and tested in the public domain of fellow professionals; and this process modifies theories. Elements of them are quietly or noisily rejected, abandoned or simply never followed up (Laudan 1977). What are the memes of a theory, then? They are those elements of it that are subjected to selection through testing and co-adaptation with the rest of the theory and of the wider field of the scientific culture (Wilkins forthcoming). Theories, like Whitman, contain multitudes (and can therefore contradict themselves), but scientific work consists in large part in exploring the ramifications of a model, in order to eliminate contradictions, either in terms of self consistency and coherence, or in terms of agreement with method and observation.

In the special case of memes in scientific theories, whether or not Dawkins intended it, it becomes clear that in a general evolutionary model a replicator is defined by (1) being transmitted intact, and (2) being subjected to selection. Many things are transmitted in culture that are not selected, and many things are selected that do not get transmitted<sup>16</sup>. I conclude that a meme is something (the "smallest" something you can identify) that gets replicated and selected in culture as a unit and therefore offer the following short definition of "meme".

*A meme is the least unit of sociocultural information relative to a selection process that has favourable or unfavourable selection bias that exceeds its endogenous tendency to change.*

Exceeding endogenous change simply means that it gets transmitted intact more than it gets transmitted mutated (Dawkins' condition of *fidelity*); in other words, it is more information than noise. Compare this with Williams' *evolutionary gene* above.

A single memetic interactive trait, I shall call it the pheme by analogy with a *phene* or Mendelian character, is the expression through some behavioural regularity of a meme at the level of selection. "Behavioural" here refers to interactive activity at the level of expression, and so includes mental behaviour, individual behaviour, group behaviour, and so forth (figure 2). However, the behaviour must be both *causally* effective in terms of acquiring resources and must be in some manner empirically *quantifiable*. One question that memeticists should ask of any analysis is "what are the resources"? A scientific theory may gather resources of researchers' time, grants and lab space, publication space in the appropriate journals, and from what I can gather, postgraduate students appear to be a prized resource. Hull's (1988a, 1988c) mechanism is that individual scientists seek to acquire credit, either through innovation, or more probably (since innovation is rare) by inclusion in a successful research project through citation and extension of results. An idea in an individual mind will need to gather attention space, memory resources, and time, in order to out compete other ideas. Without an answer to the question of *what resources?*, and they must be measurable, any memetic account is merely anecdotal.

## Sidebar: phenotypes, phemotypes and classification

A phenotype in biology, and especially in taxonomy, is the whole array of organismic characters of an organism, and is distinct from its genotype. In the taxonomic debate between cladism and pheneticism described in Hull's book, the resulting consensus was that there were no privileged features that are phenetic characters (Hull 1988c, Ridley 1992) - anything that can be measured might be used to classify organisms and reconstruct phylogenetic relationships. This argument applies just as well to reconstructions and classifications using genetic sequences as it does to skull shape and the presence or absence of feathers. Some taxonomists think that the best way to identify species is to make use of *non-adaptive* characters on the grounds that these are less likely to be the result of convergent evolution. The popular belief that it is the entire *phenotype* that is the interactor in evolution is, in my view, wrong. Interaction occurs at the level of the trait and the "engineering" fitness of an organism - how well it makes a living overall - is the product of its various interactive traits. The analogy I wish to make is that memes result in behavioural regularities that have interactive properties, that is, which result in social successes and failures.

The *phemotype* is the array of the phemes that are the causal outcome of memes in an individual or group. Phemotypes are the basis for grouping social phenomena, that is, for classifying memetic ensembles into religions, communities, traditions and programs.

Once we have identified a meme in a particular process, though, that is not the end of the explanatory story. Memetics covers what is *common* to the spread of strategies, musical phrases, linguistic practices, ideas and theories, but it does not exhaust the research. Again the parallel with genes is instructive. Having identified that there are factors that are inherited which cause ecologically significant traits, geneticists moved on to identifying the physical processes and their effects, which turned out to be extremely complex, with genes that code for traits ranging from the sorts of proteins that a cell expresses, to genes that regulate other genes, to genes that contribute to a range of gross morphological traits at every level of the organismic structure. Identifying memes will assist in identifying the sorts of cultural phenomena that are functionally transmitted, but we should expect that these will in turn become causes of phenomena in another level of analysis.

## 4 The lineages and ecologies of culture

Any chain of descent will create a lineage over time and space if ancestral entities terminate at some stage. Given that the obverse of a replicator is an interactor, an economic entity that competes in a formal sense with other interactors for the same resources, we need to identify the resources, the ecologies, of memes. On the face of it, memetic resources include neurological and behavioural time and space, much like a program can tie up the CPU cycles and screen space of a computer, but they must also extend to the more highly derived necessities of social life, including such things as credit, currency, and so forth. Selection processes define not only the replicators (the memes) and the individuals, but also the active and relevant resources of the ecology and economy of social life. The interplay between these facets of memetic evolution leads to successful and persistent lineages, as well as transitory and unsuccessful ones.

The nature of and reasons for lineages in biological evolution is a complex matter. Debates about the "Species Problem" and the relationship between "horizontal", or "non-dimensional" conceptualizations of species - known generally as biological species concepts - and "vertical" or "phyletic" conceptualizations of lineages over evolutionary time - paleontological, evolutionary and phylogenetic species concepts - have been ongoing since at least the modern synthesis, and indeed is sometimes held as the defining debate of the modern synthesis of evolution (Dobzhansky 1937, Mayr 1970, Mayr and Provine 1980). One thing that is generally agreed upon is that ancestor descendant lineages that do *not* regularly recombine ( reticulate ) at a scale on the order of one generation are not species, but are within-species lineages or something else; for example, what Eigen (1993) calls quasispecies in the case of viruses. In other words, species are those largest collections of biological lineages that do regularly recombine, and which do not themselves recombine with other lineages<sup>17</sup>. However, socio-cultural lineages appear to exchange memes on an almost unconstrained basis. Practices, metaphors, and ideas transfer from one cultural lineage, culture or tradition to another with consistent ubiquity. Autochthonous cultures the world over have adopted a range of cultural artefacts from the dominant western culture, for example, and this pattern of dominant cultural invasion recurs through recorded and archaeological history. This reticulate evolutionary pattern is touted as one of the major disanalogies between cultural and biological evolution. If cultural transmission is so fluid and free, it is thought, what use is the transfer of biological to cultural categories and modes of analysis? I shall argue that this objection rests on an incorrect assumption about biology, founded on our metazoan, vertebrate, and mammalian prejudices, and that even if it were true, it would not undermine the memetic enterprise.

Other than accepting the conclusion of this argument from disanalogy and abandoning memes as useful theoretical entities, there are three alternative responses to the difficulty it raises for memetics. The first is to shift ground, and to treat memes epidemiologically; that is, using a metaphor of disease rather than of evolution. Treating memes as pathogens, and trying to analyse the vectors and pathogenic dynamics of memes, is a strategy adopted by Goodenough and Dawkins (1994), Lynch (1996) and others (cf. Dennett 1995:364-368). "Mind virus" is itself a tolerably good meme. The second approach is to restrict memetic analysis to domains like the sciences, where there is a strong selective pressure and a good record of the spread of scientific ideas, and where lineages tend, within certain limits and until recently, to be relatively isolated and *sui generis*. The third is one that I have not seen explicitly presented but which is implicit in the work of Hull (1988a, 1988b, 1988c). This is to see that evolution produces a continuum of results from entirely isolated lineages with the "bridgeless gaps" so beloved of Mayr resulting from obligate sexual recombination through to moderately hybridizing lineages like those of flowering plants and ferns, to regularly recombining lineages - rarely if ever present in zoological biology but frequent in bacterial phylogeny and some forms of culture. Each of these approaches has its merits.

The first response, the *epidemiological model* of memes, takes into account the rapidity of transmission and mutation (relative to the observers) of certain varieties of cultural items like vernacular speech and fashions, and the varieties of strategies they evolve to invade cultural agents and populations. This view does not, however, obviate the need for an evolutionary, selectionist, account of the pathogens (cf. Ewald 1994). It's a matter of perspective. Being hosts to so many pathogens, parasites and symbiots, humans tend to conceptualize disease as sub-

generational and populational events in a relatively short time scale. Nevertheless, epidemics and pandemics like malaria have evolutionary effects on humans, resulting in evolutionarily stable genetic equilibria like the maintenance of sickle cell heterozygosity in regions affected by the transmission of the malarial parasite *Plasmodium falciparum* by some species of the *Anopheles maculipennis* mosquito group<sup>18</sup>. But we also often overlook the fact that, from the perspective of the *pathogens*, humans and other animal hosts are just another part of their ecosystem, and our immune systems act as selective environments affecting *their* evolution. It's also a matter of degree: viruses, for example, mutate much more rapidly than the gametes of a metazoan, but this is a difference in rate rather than kind. There is a bias imposed on us by virtue of being multicellular, animals, and living at the time scale we do. Evolution has no such bias. The equations governing (or describing) epidemiology are re-castings of the Fisherian and Wrightean equations covering (or describing) natural selection (Cavalli-Sforza and Feldman 1981), which is what we should expect if we see epidemiological phenomena from the pathogenic perspective and not from the host's<sup>19</sup>. What this signifies for memetics is that meme epidemiology needs to clearly distinguish between human minds and memes as entities that generate their own separate lineages. More on this shortly. For now, it is worth observing that meme lineages may not be so obvious as we think, and that cross-lineage borrowing, which occurs within biology as well as culture, may be less frequent than seems the case when we view memetic evolution from the perspective of the host minds.

The second response, the *narrow conception*, is to restrict the application of memetic analysis to those cultural phenomena which are phenomenologically isolated, well recorded, or in which there is a clear selective pressure such as in technological, scientific or economic change. There is some merit to this: but it is a heuristic merit rather than a theoretical limitation. It is a bit like the situation facing early genetics. Genes were, in the heyday of the Mendelian revolution, theoretical entities that some of a positivistic bent considered mere instrumental constructs. Positivists held (and hold) that theoretical constructs have no ontological import. "Genes" were useful in calculation and modelling, but it was illegitimate to infer that any such biological entity existed. In retrospect, this made a metaphysical virtue of a methodological necessity. Once the science, and ancillary sciences and techniques such as x-ray crystallography, had developed further, physical genetic entities were discovered and described, but this was not open to the early Mendelian researchers, and in fact attempts to describe the physical properties of genes by Weismann and others resulted in failure and open speculation. Had this positivistic operationalist approach been taken seriously, the molecular structure of genes might never have been investigated and discovered<sup>20</sup>. It *might* be a good strategy to begin with the more accessible areas of research like the history of a science, and to develop methods and models that can be extended later as more problematic cases come within theoretical reach. In the same manner, biological concepts like "species" are often developed in relatively clear-cut cases like sexually recombining animals and plants, and then extended into cases of parthenogenetic, polyploidal, clonal, viral and other "eccentric" lineages. We must, of course, beware of biting off more than we can chew, but also of not closing off any further avenues of research. Memetics, like any science, must evolve gradually.

The third response I'll call the *generalist conception*. This is to more broadly characterize biological evolution. As Hull has said, there is nothing so outlandish that an example of it cannot be found in biology somewhere. For example, take the cases of reproduction modes and

speciation: memetic evolution appears to be more like the evolution of plants, which have an estimated 30-40% rate of reticulation, and fungi and single celled organisms, with monoparental or clonal reproduction, than it is like zoological evolution with which most non biologists are familiar. Nearly half of all flowering plants are the result of joined lineages rather than split lineages, and nearly all fern species. Amoeba undergo a process called conjugation where genes are exchanged every few generations, although the rest of the time they split apparently.

Amazingly, even viral RNA can be exchanged through crossover of transcriptase in superinfected cells<sup>21</sup>. Evolution is still cleverer than we are. These idiosyncratic forms of individual, phenotype, lineage and evolutionary behaviour are nevertheless covered by the various theories of evolution<sup>22</sup>. The generalization of evolutionary theory accommodates these "eccentric" cases just as well as it does the paradigmatic case of animal evolution on which it was first founded. Natural selection is an emergent process acting on ecologically interacting replicators, and obligately sexually reproducing individual organisms are just a special case, even though they are the most obvious and easily investigated form of evolution *for us*.

Therefore, it would seem that natural selection can model cultural lineages even if they do not happen to form tightly integrated populations of recombining entities. But we need to investigate *a posteriori* whether any particular kind of cultural lineage is tightly integrated like an animal species or not, and not to assume that there is, in fact, such a disanalogy. Some are, some are not.

Of those that are not, selection may be a major cause of change or it may not. This is an empirical argument, not an axiomatic derivation. In other words, we should challenge the major premise of the disanalogy argument in each case. The issue of "quasi-species" (Eigen 1993) in obligately clonal bacteria and viruses illustrates this. In principle, and intuitively, every time a clone or virus generates a mutant, one gets an entirely new strain. We might expect that strains would indefinitely diverge to fill the entire space of ecological niches available. In fact what happens is that viruses tend to cluster about the "wild-type" phenotypic mean (which may not even exist in a single genome), forming analogues to recombinant species. Natural selection accounts for this as well as it does for sexual, biparental species.

While on the topic of cultural lineages and species, let us consider the matter of speciation mode. There is a raging debate about whether sexual species speciate through allopatry (geographical isolation) or sympatry (geographical and ecological coexistence) or some intermediate state (parapatry) (Mayr 1970, Gibbons and Morell 1996)<sup>23</sup>. This resolves to the question whether natural selection, as well as being responsible for allele frequencies *within* a breeding population and for adaptation overall, is also responsible for *creating* lineages. Those who, like Mayr, opt for the allopatric and parapatric modes alone, consider that isolation followed by selectively neutral or decoupled drift is responsible for speciation, and they tend to see selection as an ancillary process to speciation rather than a reason for it. Sympatric speciation is a litmus test of whether one thinks selection is the main or even the only mechanism of evolution. If species can diverge in the same ecological territory, it must be because selection causes multiple modes in the one population, leading eventually to more than one species as the modes chase different adaptive peaks. If correct, this means that selection can cause speciation, and the traits of species that differentiate them are adaptive. If one denies this happens, or that it happens very much, then what differentiates species can be adaptively neutral or even maladaptive. Nobody presently

thinks that selection can be dispensed with even in cases of speciation by hybridization of plants, which occurs in a single generation or even repeatedly (polyploid speciation), but the differences in the degree of the attributed importance of selection points to a deep division in modern evolutionary thought, and this is very important for memetics if memetics is based on selectively significant cultural transmits.

Hull (1988c) has addressed similar issues in scientific evolution. He uses Wright's notion of structured and relatively insulated breeding populations - called demes - to characterize research groups. According to him, scientific research arises first within relatively co-operative demes of scientists - groups within which more variation is tolerated and selection pressures are less severe than in the wider discipline. This permits new theories to be developed and articulated to the stage where they may be able to survive rigorous criticism by rival groups when published. Intriguingly, Hull does not require unanimity of opinion within the deme, but merely a desire to increase one's credit by sharing in the conceptual inclusive fitness of the deme: reputation is everything in science. Hull's view of the social and conceptual development of science is a well developed model of conceptual change, one of the most articulated memetic models yet presented, and it can be extended into other domains. For example, although Hull does not characterize demonic structure itself in this way, it can be thought of as an evolutionarily stable strategy for science to adopt - with a balance between public debate and private support, conceptual transmission in science is able to develop a great deal more novelty and yet retain more strongly tested ("more fit") ideas than was possible in the late medieval guild tradition from which it sprang. Neither a selectively moderate social ecology nor a naked selective environment where new ideas are never given a chance could provide the sort of adaptive growth science has achieved. Obviously, periods of stagnation in a discipline may be due to changes in these parameters (for example, transmission rates) and connecting them in a model based on a generalized evolutionary and ecological process is an important first step to a full memetic analysis of science. *Mutatis mutandis*, the same is true of memetic theory in general. If we are to understand how memes diversify into relatively stable lineages like religious traditions, the structures within which novelties arise and the selective pressures to which they are exposed are crucial<sup>24</sup>.

*Ecological succession* is another analogue of value. It is well recognised in biogeography that niche occupiers in an ecosystem can often fend off adaptively superior invaders; the dominance of a given (say) grazer species in a locale may be due to it being there first rather than its overall adaptive excellence. We should expect that ecological succession, which is complex in biology, will be equally or more complex in culture, but there may be fruitful avenues to explore here. To understand and employ these analogues, we also need an analogue for "ecology" and "ecosystem" for a given socio-cultural domain. The ecology of science, for example, is not nearly as simple as the falsifying crucial experiments that Popperian cycles propose. The resources for which a social or conceptual structure competes will usually be fairly broad even if, as in the case of a scientific idea, storage and processing space and time in an active central nervous system is the basic resource.

## 5 Memetic individuals

With the clarifications and mental tools introduced above, we are now able to ask Hull's question from biology - what is an individual? - in the memetic context. What is a memetic individual? What is subjected to selection in culture? What gets "coded for"?

When Juliet bitterly but eloquently complained how Romeo's social relationships were messing up their love life, she made the interesting observation that being a Montague, and being the person she loved, were two distinct states (apparently she had more of an interest in some of Romeo's biological aspects). The instantiation of the cultural relationship "is a Montague" in a particular biological organism, denoted by the name Romeo, is a case where an individual is something other than the sum of his own memes. The converse argument was given by the idealist philosopher FH Bradley in 1876, in a landmark essay in ethics, "My Station and its Duties". Here Bradley wishes to establish that one's social location and relations determine one's moral responsibilities. To do this, Bradley argues that what we are as social beings, as moral agents, is determined by the community of which we are part. An Englishman is not his biology<sup>25</sup>. In each case, Juliet and Bradley recognize the distinction, often overlooked by meme enthusiasts, between the biological and the memetic. Memes don't necessarily make you more biologically fit, nor are they necessarily going to make you less fit. Memes aren't fit themselves simply because they make you live healthier lives. Memes are fit only insofar as they are propagated successfully; forget the effects they have on biology. In Toulmin's (1972) underappreciated book, he says of biological and cultural evolutionary processes that they are *decoupled*. This is not to say that the two realms do not meet and affect each other, for clearly they do; it is to say that no matter how you might be able to conceptualize cultural phenomena in biological terms (socio-biology), you can independently conceptualize them in social terms (memetics). To lift a phrase of Williams' (1992) there is a dearth of shared descriptors between the entities of biology and the entities of culture. Sometimes they may, indeed, be the same objects or processes, but you have to describe them differently in each analytic realm.

No end of confusion has been caused by ambiguous specification of the socio-cultural analogues of phenotype and genotype. An early criticism of Hull's model of models (Heyes 1988, Tennant 1988) was that scientific memes failed to *constitute* scientists in the same way that genes constituted organic individuals. By the same token, critics argue that selection against memes fails to result in the death or restricted fertility of scientists, and so natural selection is an inappropriate way to model scientific theory change. Another recurring criticism of evolutionary pictures of culture is that unlike organisms and their genes, cultural agents choose which memes they adopt, and that they modify their memes in the light of experience. Cultural change would be Lamarckian, say critics, if it evolved at all. These two attacks on memetics - the *aphenotypic* disanalogy and the *cultural Lamarckism* criticism -- are closely intertwined, and have to be discussed together.

Using the relatively clear case of science as an example, we can see that any one scientist can adopt differing memetic stances over time. When Einstein was learning physics, he was probably taught the axiomatic observer independent Newtonian theories of the day. He later changed his mind about that. When the physicist Schrödinger got interested in biology, he entered into the canons and data of a new discipline (for him). Here are two exemplary cases of memetic

acquisition - the replacement or supplementation of one set of memes with a whole new set. Yet, neither Einstein nor Schrödinger became different organisms with these memes. Moreover, had either of these scientists undertaken a non-scientific hobby requiring much technical knowledge and experience, such as angling, oil painting or cabinet making, the likelihood is that they would be able to keep them relatively distinct from the professional standards and methods of physics. So, how can a "physicist" be *caused* to come into being by memes, in the same way that a blackbird is *caused* by a blackbird genome?

I tacitly answered this question by inserting the word "professional" above (cf. Wilkins forthcoming). A human organism who is a member of a culture, language group, class and so forth, only achieves professional status or competency upon completing some portion of a "developmental" process of acquiring and exercising the relevant memes. A scientist, to return to our example, must acquire the standards, knowledge base and experience of the particular discipline (physics) and the sub disciplines (for example, high energy physics). There is an *aspect* or *profile* to every scientist that is (a) constituted by memes, and (b) within the distribution curve of traits exhibited by the lineage (the *phemorph*). As a scientist, an organism is distinct from other cultural profiles he or she instantiates, and the profile is a separate class of entity from any biological organism of species *H. sapiens* or any set of traits that the organism exhibits by virtue of being a member of the species. Bradley considers the thought experiment of raising the same biological organism now an Englishman in another society, or on a desert island. Clearly, says Bradley, not in these words, the *properties* of Englishness are derived from the relational and (non-biological) developmental processes of living in English society and having a locus there. In our terms, "being English" is a memetic profile derived from the acquisition of Englishness memes. We therefore have to be careful of the denotation of the terms "scientist", "Englishman", "capitalist" and so on. The scientist-as-organism is a distinctly *biological* entity, distinct from the *cultural* entity scientist-as-Englishman, or scientist-as-capitalist, which depend for their specification on the memetic profile of the scientist proper, the scientist-as-professional. In the context of a professional lineage such as a science, or other intellectual pursuits or social practices such as accounting or rap dancing, the memetic individual is the competent member of the lineage, which is developed by the lineages' professional or cultural properties to create a profile within the human (biological/neurological) individual organism. With appropriate adjustments, the notion of a memetic individual can be generalized from the case study of science into other discernible cultural institutions and traditions. In the limiting case, where the cohesion of a lineage is very loose, memetic individuality may be partial or fragmentary (like being good at using a yoyo), but if memes *constitute* anything at all, they constitute the profile of a memetic individual.

Memes, like genes, can only "code for" a norm of reaction. All cultural as well as biological traits are distributed over a population curve, with the mean and the mode correlating to the memetic selection bias. As Hull observes, no two scientists even with identical theoretical commitments interpret their views exactly the same way, and it is an oft repeated half joke, half complaint, that there are as many views in a research program as there are practitioners, sometimes even more. Neither memes nor genes determine all aspects of the properties of the entities they constitute. What they do determine are the degrees of freedom and they bias and constrain the outcomes of the so-constituted system. In thermodynamic phrasing, they specify the field of states such memetic systems have a propensity to attain. Do they do this through a

process analogous to the Weismannian Central Dogma, with a one way germ lineage, or through something analogous to a Lamarckian pangenesis (Hull 1988c, chapter 12) with the environment instructing the memes? Is the question even relevant? Cultural and biological evolution *are* going to be different in the frequencies of the kinds of processes they undergo. Perhaps culture does exhibit Lamarckian-style inheritance through the sort of environmental instruction that has never, or rarely, occurs in biological evolution. One outcome of this would be that variation would come more frequently and more intensively than in a purely Weismannian process, which has to await random mutation or use stored variation from earlier mutation in order for selection to have something on which to operate (Fisher 1930). It would, however, still be Darwinian *evolution* even if *inheritance* were Lamarckian. Lamarckian *evolution*, however, is a different sort of process altogether, driven by perceived need to achieve foreseen outcomes; (see<sup>7</sup>). Lamarckian *inheritance* is not inconsistent with a Darwinian model of memetic evolution. Lamarckian *evolution* would totally demolish the foundations of memetic theory, and leave us with more traditional forms of cultural analysis. However, I do not think that the case has been made that memes are even acquired through instruction in the Lamarckian sense, and refer readers to Cziko's (1995) discussion for further consideration.

It is my opinion that cultural inheritance is not particularly creative, and that most "novelty" is in fact the recombination of pre-existing memes in novel ways: there is little that is new under the sun. Neither do I think that the rate of creativity rises in times of stress or great change. I speculate that what changes, and gives the appearance of memetic novelty, are such factors as selective pressure coefficients, "migration" rates, and density dependent recombination rates that vary from the "usual" background rates. Problems can persist for long periods even in the face of urgent need, and unless the appropriate combination of memes occurs, more or less at random<sup>26</sup>, they do not get solved. This takes us back to the question of the demonic structure of culture - too great an isolation and a more optimal memetic combination is unlikely; too little isolation, and combinations are unlikely to be stable enough to undergo selection, and will be swamped.

The one major disanalogy Hull accepts between biological and cultural evolution, and with which I agree, is in the rate, but not the kind, of reticulate phylogeny; that is, in the merging and crossover of memes between lineages. Again this is not so much a difference of kind as of degree. As noted above, cultural evolution resembles that of bacteria, plants and fungi more than that of animals - Marvell's "vegetable love" that grows "vaster than empires" is perhaps the better analogy, although on an evolutionary time scale, culture is anything but "more slow".

It might help to visualize the analogy I am proposing by giving examples in a table of the analogues between biological and cultural entities and processes, mapped to the Hull-Dawkins and Ghiselin-Hull ontologies (Table 1, cf. Ghiselin 1997 for a review).

Entity (e) or process (p)	Biological	Memetic	Science	Language	Economics
Interactor (e)	Phenotype or trait	Phemotype or Pheme	Experiment or Observation	Linguistic behaviour	Enterprise or transaction
Replicator (e)	Gene	Meme	Theory or Hypothesis	Language element <sup>27</sup>	Account, resource
Character (e)	Mendelian Trait (phene)	Pheme	Method, Result	Linguistic practice	Business practice
Lineage (e)	Species, phylum	Tradition, institution	Research program	Regional dialect, language group	Business, industry
Economy (e)	Ecological system	Culture	Disciplinary community <sup>28</sup>	Language community	Market sector, fiscal system
Reproduction (p)	Organismic reproduction	Constitution of new profile	Degree, training	Language acquisition	Establishment of new enterprise
Individual (e)	Organism, kin group, colony	Memetic individual	Scientist	Language speaker	Economic agent
Substrate <sup>29</sup> (e)	Organic molecules	Neural networks			
Code (p)	DNA alphabet	Semantic	Semantic and mathematical	Grammar and vocabulary	Currency
Encoding medium (e)	Amino acids, proteins	Neural structures, practices	Neural structures, journals, books, institutions, disciplinary associations	Neural structures, written material, recordings, etc.	Neural structures, computers, books, ledgers, receipts, bank files, etc.

**Table 1** Examples of general evolutionary analogues (e=entity, p=process)

It must again be stressed that these are only some of a number of possible sets of analogues, and what functions as an interactor at one level may function as a replicator at another, so that a pheme for one level may be a meme at another, and vice versa.

## 6 Conclusion and prospectus

**6.1 Conclusions:** My conclusions are as follows. Memes are those units of transmitted information that are subject to selection biases at a given level of hierarchical organization of culture. Unlike genes<sup>30</sup>, they are not instantiated in any exclusive kind of physical array or system, although at base they happen to be stored in and expressed from neurological structures<sup>31</sup>. Many memes reside as neural net structures in the central nervous systems of humans, but many also emerge at a higher cultural level. All memes have neural substrates, but not all are encoded in those substrates. Memes are also situated in a variety of larger semantic structures, behavioural regularities, and cultural substrates. They are identified in virtue of their selective roles. Memes must be expressed in a cultural ecology in order to be selected, but it is the *class* of behaviours rather than the behaviours themselves that are memes. Memes do not control behaviour (including mental behaviour) rigidly, but bias and constrain it to a norm of reaction. Memes are the replicators of cultural evolution and the structures that bear the cultural properties they express as are the interactors, in the language of the Hull-Dawkins Distinction. They are, as Hull once entitled a paper (1987), genealogical actors in ecological roles. Packages of memetic interactive properties - phemes - constitute the phemotype of memetic individuals, or memetic profiles, that are not coextensive with the descriptors of the biological individuals in

which they are instantiated, and cultural evolution is neither identical to nor derived from biological evolution. Memetic inheritance may be, but probably isn't, analogous to Lamarckian inheritance, but in any event, memetic evolution is Darwinian. Memes form ancestor descendant chains of populations that ramify and reticulate with frequencies differing from biological phylogeny, but the differences appear to be within the extremes of the parameters of biology. The models developed for biological evolution and ecology need to be understood more broadly than just vertebrate animal evolution and applied as they are suited to culture, in order to determine the general evolutionary properties of both domains.

**6.2 Methods:** The methodological future of memetics lies in the use of techniques drawn from information theory, modern taxonomy and computer science. By interpreting memes as messages, we can make use of Shannon-Weaver entropy (cf. Brooks and Wiley 1988), cladistics (Wiley, et al 1991), and connectionist mathematics as implemented in the Artificial Life program (Holland 1995), but as each of these requires quantification of the data or input, it is important to be able to commensurably map the memetic elements and to measure the selection biases they undergo. If, as I have argued, a meme exists in virtue of biased transmission rates, then there is no smooth reduction of memetic structures from cultural behaviour to atomic memes, just as there is no smooth reduction from phenotypic traits to single genes. The researcher seeking to explain a singular historical shift of memetic frequencies must iteratively refine the data until it becomes clear what is being transmitted at a level and how it is being expressed. The problem of classification lies at two ends of the scale - identifying cultural traditions as they exist now and over time, and identifying elements of those traditions as they persist and recombine. Although this sounds subjective, it need not be. Behavioural regularities indicate that something objective has been spread, and even if the underlying memes cannot be formulated in some universal logical language, the structure of memes can still be identified in the same way as Mendelian genes and molecular sequences, through the use of consensus maps and by noting when their absence or presence makes a difference.

The analytic tools available to us are legion, having been developed in biology over a century or more. They include Wagner ground plan methods (unrooted character similarity trees based on presence-absence matrices), Hamming Distance measures (the sum of the number of simple differences between two memes)<sup>32</sup>, cladistic reconstruction using parsimony methods,<sup>33</sup> and pattern recognition methods that use neural networks and other connectionist models. Eventually, these and other methods will no doubt be incorporated into a body of canonical techniques within the memetic enterprise, and I expect they will be generalized as methods applicable in a range of social disciplines. The methods being developed for complex adaptive systems theory at the Santa Fe Institute are also likely to become important for memetics as well as other research dealing with complex adaptive phenomena (Casti 1994). However, if we lack clarity on the core ontology of a selection process in culture, the data to which we could apply these methods will be subjective and the explanations that we derive from them are in danger of being entirely vacuous.

## Acknowledgements

I am indebted to Mario Vaneechoutte and Mark Mills for critical comments on an earlier draft of this paper, and the referees for subjecting the submission to strong selective pressures on style

and content. I must also thank the late David Rindos for encouragement. Major stylistic improvements are due to the attention of Vladimir Brusic, though he is not to blame for any remaining infelicities.

## Glossary of technical terms used

### Allele

An alternative gene at a particular locus.

### Allopatry

The state of populations living apart from each other (literally "other homeland"). Hence, *allopatric speciation*.

### Central Dogma

Weismann's hypothesis that gametes ("germ cells") are passed on independently to what happens to the organism's body.

### Cladistics

Taxonomic classification that reconstructs the order of the appearance of evolutionary novelties from their present distribution. Also known as *cladism*.

### Cladogenesis

The origination of a new species by the splitting of a single lineage into two. See *speciation*.

### Co-adaptation

The process of local adaptation of genes and their effects to each other, so that they function well as a unit.

### Conjugation

In some single celled organisms, the occasional exchange of genes in otherwise apparently reproducing organisms.

### Deme

A small population that is relatively isolated from the larger species or tradition and has its own distinct genetic or memetic characteristics.

### Emergentism

The metaphysical position that properties arise from the relations of objects that are not properties of the objects themselves.

### Epistasis

Linkages in the effects of genes, such that a single gene may affect many traits or a single trait may be affected by many genes.

### Evolutionary gene

"[A]ny inherited information for which there is a favorable or unfavorable selection bias equal to several or many times its rate of endogenous change" (Williams 1966: 25)

#### Evolutionarily stable strategy

A strategy coded for by genes or memes that does better when interacting with copies of itself than alternative strategies do, and which will tend to become the sole or dominant strategy of a population that will not then be susceptible to invasion by other strategies.

#### Facultative

Of an organism, the possibility to adopt variant lifestyles, one of which is the norm.

#### Founder effect

The evolution of a new lineage based on the sampling errors of the small starting population, which may be of different proportions to the original populations.

#### Gamete

The sex cell of each parent, which recombines to produce the zygote, such as sperm and egg, or spore.

#### Gene

The fundamental physical unit of heredity that transmits information from one cell to another and thus to successive generations.

*Mendelian gene*: a unit of heredity that causes a single phenotypic character or trait.

*Molecular gene*: a sequence of nucleotides (DNA, tRNA or rRNA) that functions as a unit during transcription and which is transmitted whole.

*Evolutionary gene*: see evolutionary gene.

#### Genet

All the clonal entities that share a genotype.

#### Genome

The complete complement of the genetic material in a cell or carried by an individual.

#### Genotype

The genetic constitution of an individual, often referring to genetic basis of some particular characters.

#### Heterozygosity

Carrying two different alleles from each parent at a locus.

#### Individual

A relatively well bounded and functionally coherent system comprised of components and their relationships. The formal opposite of a class or universal type.

### Interactor

"[A]n entity that interacts as a cohesive whole with its environment in such a way that this interaction causes replication to be differential" (Hull 1988c: 408).

### Lineage

"[A]n entity that persists indefinitely through time either in the same or an altered state as a result of replication" (Hull 1988c: 409).

### Meme

The least unit of socio-cultural information relative to a selection process that has favourable or unfavourable selection bias that exceeds its endogenous tendency to change.

### Memetic profile

The array of phemes that constitutes a memetic individual.

### Memetic individual

A competent member of a memetic cultural lineage, which is developed by the lineages' professional or cultural properties to create a memetic profile within the human (biological/neurological) individual organism.

### Methodological individualism

The philosophical belief that collectives and their properties are just the sum of the individuals and their properties that comprise them, especially in social and historical explanation.

### Mnemone

Donald Campbell's term for a conceptual replicator, roughly equivalent to meme.

### Norm of reaction

The distribution curve of the phenotypic effects of a gene in a population.

### Obligate

Of a parasite, the forced mode of living or the necessary host. In general, the lifestyle that an organism is forced to adopt.

### Pangenesis

Darwin's theory of heredity, a use inheritance view, in which parts of the body were supposed to throw off "gemmules" that were carried to the reproductive organs, and which carried information about the body's experience to the next generation. This view was discredited by August Weismann in the 1880s.

### Panselectionism

The view that all characters of an organism have an adaptive reason for evolving.

### Parapatry

The state of living in adjacent regions with or without some overlap ("bordering homeland"). Hence, *parapatric speciation*.

**Parthenogeny**

Asexual reproduction through unfertilized eggs of a lineage that evolved from a sexually reproducing ancestral state. Adj. *parthenogenic*.

**Pheme**

A single memetic interactive trait which is the expression through some behavioural regularity of a meme at the level of selection. It is the least type of selectively biased behaviour relative to a culture.

**Phemorph**

The normal distribution curve of traits exhibited by a cultural lineage.

**Phemotype**

The array of the phemes that are the causal outcome of memes in an individual or group.

**Phene**

A Mendelian character or trait. Adj. *phenetic*.

**Pheneticism**

Taxonomy based on the groupings of phenes, without respect to evolutionary lineages.

**Phenotype**

The observable features of an organism, which develop according to its genetic code (genotype). Adj. *phenotypic*.

**Pleiotropy**

The state in which one gene affects two or more phenotypic traits not otherwise directly related.

**Polyplody**

The fusion of three or more complete sets of chromosomes, sometimes from distinct species, usually in plants.

**Quasi-species**

Manfred Eigen's term for clusters of related clonal organisms that mimic species in the way they remain similar.

**Ramet**

An entity that is one of a number of genetically identical organisms.

**Random drift**

Sewall Wright's model of random allele frequency changes in small populations without the operation of natural selection.

### Reductionism

*Explanatory reductionism*: the philosophical doctrine that a complete explanation of a complex whole is given by enumeration of the components of that whole and their properties, especially in scientific explanation.

*Genetic reductionism*: a form of explanatory reductionism that has developed into a research program, which holds that all phenotypic properties of organisms and their evolution can be understood in terms of genes and their fitness levels.

### Replicator

"[A]n entity that passes on its structure largely intact through successive replications" (Hull 1988c: 408).

### Reticulation

The recombination of distinct phylogenetic lineages. See *speciation*.

### Selection

"[A] process in which the differential extinction and proliferation of interactors cause the differential perpetuation of the relevant replicators" (Hull 1988c: 409).

### Spandrel

Gould's and Lewontin's (1979) term for a trait or structure that is a necessary by-product of some other adaptive feature, and which is not therefore explained in terms of selection in favour of it.

### Speciation

The process of the evolution of a new species, through splitting from an existing species (cladogenesis). Also used of hybridization of existing species to form a third (reticulation).

### Species

The largest collections of (sexually reproducing) biological lineages that do regularly recombine, and which do not themselves recombine with other lineages.

### Sympatry

The state of coexisting in the same region ("same homeland"). Hence, *sympatric speciation*.

### Use-inheritance

Also called *soft inheritance* (Mayr 1982: 691). A view held by Darwin (*Origin* chapter V) that "use ... strengthens and enlarges certain parts and disuse diminishes them, and that such modifications are inherited". Sometimes called "Lamarckian inheritance", although Lamarck was not the first to propose this view, which is a folk belief of long standing. In Darwin's theory of pangenesis, use-inheritance is responsible for modifying the frequency and novelty of variation, and therefore evolution. Fisher (1930) discusses the problems of use-inheritance for Darwinian evolution, and the arguments against it.

## Notes

1. Technical terms are linked to definitions in the glossary and discussed, where relevant, in the body of the paper.
2. That is, it represents something causally significant in the theory, like an electron.
3. An analogy between two domains depends on a common etiology. A metaphor can only suggest limited and vague similarities. For instance, the term "analogy" in phylogeny refers to convergently evolved traits through adaptation to similar conditions of life. If selection of culture is an analogy, it has theoretical weight; if a metaphor, we can abandon or modify it when we find it difficult to apply.
4. And Dawkins is a *lot* of a panselectionist. Panselectionists tend to find selective explanations for every feature of organisms, for instance asking what the adaptive significance of the human chin is, when chins are just the result of two growth fields interacting.
5. By "evolution" I mean the generation of variations and adaptation. I am not committing myself to the claim that natural selection is, or is not, responsible for new species through cladogenesis.
6. Weismann was the first researcher to show that gametes (he called them "germ cells") are passed on independently of what happens to the organism's body. His Central Dogma, as it became known, was an attack on use-inheritance, the time honoured belief (shared by Darwin among others, cf. Mayr 1982: 687-694, Richards, 1992: 172) that those features of an organism that are most needed and most used will be more frequently inherited. Even after the period of the development of Mendelian genetics, the Central Dogma was not universally accepted by biologists until the middle of this century and the molecular revolution in genetics.
7. The term "Lamarckism" is very ambiguous in the way it is used in biology. Much historical injustice to Lamarck was done by those who called themselves "neo-Lamarckians" in the late nineteenth century. Only two sense of "Lamarckism" are relevant to this paper: *use-inheritance* (see previous note), also known as the inheritance of acquired characteristics; and the belief that novelties evolve to meet the *needs* (sometimes interpreted as the desires) of the organisms when their environment changes. A third form often and more historically accurately referred to is that evolution is progressing to some form of perfection. The charge that evolution of culture is Lamarckian (cf. Gould 1993: 216, 1997: 222) generally refers to use-inheritance (for rebuttals, see Hull 1988: 452-457, Cziko 1995).
8. More recently, Ghiselin, with whom Hull collaborated on the so-called individuality thesis, has reviewed the arguments both have made against the "entification" of replicators, including noting that genes, in the population (or Mendelian) genetics sense, can include *deletions*, that is, the loss of a sequence can have population level effects. In his view, this is a *reductio ad absurdum* of Dawkins' genetic reductionism (Ghiselin 1997: 143-148).

9. But it is worth noting that they can be the *product* of the properties of genes. In algebra, sums are linear and additive, while products are non-linear and multiplicative. Non-linear products can be extremely complex, and are so-called because their functions do not result in a straight line on log log graphs.

The Kolmogorov theorem shows that between any two arbitrarily large sets there can be at least one mapping so long as there is an intermediate set of links, an important result for connectionist systems research. These Kolmogorov mappings are non-linear products, and applied to genes and memes the theorem means that while there is no simple reduction from interactive characters to genes/memes a mapping relation does exist, even if it is extremely complex.

10. See Rosenberg's 1994 discussion and clarification of the Hullian entities as they apply to biological processes, and also his discussion of Hull's problem of reductionism. See also Ghiselin 1997.

11. To avoid unnecessary confusion, it should be noted that any number of hierarchies can occur in a single domain like biology or culture. There may be a genetic hierarchy leading to species, but there may also be a hierarchy leading to ecological patterns, and these need not be the same hierarchy. (Eldredge 1989) This figure and the following figures are not intended to represent *the* hierarchy of that domain, but only one of a possible many.

12. However, this does not deny that a selection process based on a mental model of the states of affairs can occur in an individual mind. In terms of figure 2, the memes here are the entities that are selected in a mental environment, that is, what makes sense to the thinker, and which when expressed will become memes in a higher level process [cf. Dennett 1995, chapter 12].

13. One possible exception to this bifurcation is when a meme is a spandrel, that is, a by-product of some *other* meme that is selectively favoured, and with which it inevitably must be transmitted (Gould and Lewontin 1979, Gould 1997).

14. This resembles the Marxist-Leninist doctrine of false consciousness in some ways, but with no presumption that any part of society will instantiate a "true consciousness" of social reality.

15. Discussion in Cziko 1995: 140-149.

16. For example, a practice may spread just *because* it is selectively neutral. Junk DNA can be transmitted because it does not generate products that are "visible" to selection. Kimura's theory of neutral DNA shows that biased transmission without selective advantage is a biological reality. And any trait at the interactor level can be selected for or against, in a sense, even if it is not heritable. A blacksmith may do well biologically because of his strength, but that will not (necessarily) be bequeathed to his progeny.

17. As we shall see below, this is not a complete picture; but it captures the initial intuitions of taxonomists.

18. Interestingly, identifying the five or six distinct, but hitherto unrecognized (*cryptic*), species of *A. maculipennis* was accomplished through analysis of their being vectors for *P. falciparum* or not (Cain 1954).
19. I am of the philosophical school of thought that laws and equations merely *describe* dynamic patterns rather than *govern* phenomena (instrumentalism), but that is not germane to this essay. See Rosenberg 1994 for a discussion of instrumentalism and biology.
20. Because the resources required are rather significant, which is an interesting memetic point.
21. This is not a mistake, despite its contradiction to textbook biology. The result was announced in a recent paper (Boerlijst, Bonhoeffer and Nowak 1996).
22. It is worth remembering that Darwinian, or rather synthetic, evolution, is modelled by a number of theories (common descent, natural selection, sexual selection, biogeographic distribution, Mendelian genetics), and should not be considered a single theory on its own.
23. Allopatry is the state of populations living apart from each other (literally "other homeland"), sympatry that of coexisting in the same region ("same homeland"), and parapatry of living in adjacent regions with or without some overlap ("bordering homeland"). When speciation occurs in these states, then it is called allopatric speciation, and so forth. Many now think that parapatric speciation occurs, but the frequency is still at issue.
24. Dobzhansky's concept [1937] of "co-adaptation" (a term initially used by Darwin, I believe) is significant at this point of the argument. If memes at first arise within restricted domains, all they need at this early stage to adapt *to* are the other memes within the deme, or some part of it. Co-adapted memes may play a role in storing memetic variation until the fitness of variants rises to significant levels. Moreover, it would explain why apparently maladaptive memes can persist in a tradition. In the end, the final selection event is extinction, and maladaptive memes may persist because they are fitter within a tradition than on their own, even if they make the lineage phenotype less adaptive overall.
- As an example, consider the meme of common property in the fideist traditions such as the Mennonites and the Shakers. This meme was a core belief and tradition in these communities, and was well adapted to the rest of the memes shared by the communities, but it was maladaptive during the most aggressive period of the rise of modern industrial capitalism, and these communities eventually faded away.
25. An early shot at proto-sociobiological notions, although I think Bradley underestimates how much biology affects one's *place* in a culture. Imagine the role of a Chinese child growing up in Victorian England.
26. Random with respect to the selection process, that is.
27. Phonemes, words, syntactical structures, etc.

28. Funding for research is peer reviewed, even if it derives from governments, and so the primary economic resource is standing in that community, cf. Hull 1988c.
29. Supporting physical entities of the code.
30. Genes also include and require RNA of various kinds, so even this generalization is not absolute.
31. At least, until artificial intelligences become a reality.
32. Interestingly, Hamming Distance was first formulated in the context of information theory, and adapted for use in genetics after the molecular revolution. See Gabora forthcoming for a treatment of Hamming Distance in memetics, but one that I think is overly reductionist.
33. Cladistic reconstructions rely on an absence of reticulation, and so should be used with care in the memetic context. Cf. Wiley et al 1991.

## References

- Beatty, J: 1987. Dobzhansky and Drift: Facts, Values and Chance in Evolutionary Biology. In *The Probabilistic Revolution: Volume 2: Ideas in the Sciences*, L Krüger, G Gigerenzer and MS Morgan, eds., MIT Press, Cambridge MA, 1987
- Boerlijst, MC, S Bonhoeffer, and MA Nowak: 1996. Viral Quasispecies and Recombination. *Proceedings of the Royal Society of London, Series B*, 263: 1577-1584
- Bradley, FH: 1876. *Ethical Studies*, Cambridge University Press, Cambridge UK
- Brooks, DR and EO Wiley: 1988. *Evolution as Entropy: Toward a Unified Theory of Biology*, University of Chicago Press, Chicago IL, second edition
- Buss, L: 1987. *The Evolution of Individuality*, Princeton University Press, Princeton
- Cain, AJ: 1954. *Animal Species and Their Evolution*, Hutchinson's University Library, second edition 1963
- Campbell, DT: 1960. Blind Variation and Selective Retention in Creative Thought as in Other Knowledge Processes. In Radnitzky, G and WW Bartley III (eds.), *Evolutionary Epistemology, Rationality, and the Sociology of Knowledge*, Open Court, La Salle IL, 1987
- Campbell, DT: 1974. Evolutionary Epistemology. In Radnitzky, G and WW Bartley III (eds.), *Evolutionary Epistemology, Rationality, and the Sociology of Knowledge*, Open Court, La Salle IL, 1987
- Campbell, DT: 1987. Selection Theory and the Sociology of Scientific Validity. In Callebaut, W and R Pinxten eds, *Evolutionary Epistemology, a Multiparadigm Approach*, D Reidel, 1987

Campbell, DT: 1988. A General 'Selection Theory' as Implemented in Biological Evolution and in Social Belief-transmission-with-modification in Science. *Biology and Philosophy* 3: 413-463

Casti, JL: 1994. *Complexification: Explaining a Paradoxical World Through the Science of Surprise*, Abacus/Little, Brown, London

Cavalli-Sforza, LL and MW Feldman: 1981. *Cultural Transmission and Evolution: A Quantitative Approach*, Princeton University Press, Princeton NJ

Cziko, G: 1995. *Without Miracles: Universal Selection Theory and the Second Darwinian Revolution*, MIT Press, Cambridge MA

Darwin, CR: 1872. *The Descent of Man, and Selection in Relation to Sex*, John Murray, London

Dawkins, R: 1977. *The Selfish Gene*, Oxford University Press, Oxford UK (1989 edition)

Dawkins, R: 1982. *The Extended Phenotype: The Long Reach of the Gene*, Oxford University Press, Oxford UK, revised 1989

Dennett, D: 1995. *Darwin's Dangerous Idea: Evolution and the Meanings of Life*, Allen Lane Press, London

Desmond, A: 1989. *The Politics of Evolution*, University of Chicago Press, Chicago IL

Dewey, J: 1909. The Influence of Darwin on Philosophy. In Appleton P ed., *Darwin: A Norton Critical Edition*, Norton, New York, 1970

Dobzhansky, T: 1937. *Genetics and the Origin of Species*, 3rd edition 1951, Columbia University Press, New York

Eigen, M: 1993. Viral Quasispecies. *Scientific American*, July 1993: 32-39

Eldredge, N: 1989. *Macroevolutionary Dynamics: Species, Niches, and Adaptive Peaks*, McGraw-Hill, New York

Ewald, PW: 1994. *Evolution of Infectious Disease*, Oxford University Press, Oxford UK

Fisher, RA: 1930. *The Genetical Theory of Natural Selection*, Clarendon Press, Oxford UK (rev. ed. Dover, New York, 1958)

Gabora, L: forthcoming. A Day in the Life of a Meme. *Philosophica*

Ghiselin, MT: 1987. *Bioeconomics and the Metaphysics of Selection*. *Journal of Social and Biological Structures* 10: 361-369

Ghiselin, MT: 1997. *Metaphysics and the Origin of Species*, State University of New York Press, Albany NY

Gibbons, A and V Morell: 1996. On the Many Origins of Species. Special News Report, *Science* 273: 1496-1502

Goodenough, OR and R Dawkins: 1994. The "St Jude" Mind Virus. *Nature* 371: 24

Gould, SJ: 1981. *The Mismeasure of Man*, Norton, New York

Gould, SJ: 1993. More Light on Leaves. In *Eight Little Piggies: More Reflections on Natural History*, Norton, New York

Gould, SJ: 1997. The Exaptive Excellence of Spandrels as a Term and Prototype. *Proc Natl Acad Sci, USA* 94: 10750-10755

Gould, SJ and RC Lewontin: 1979. The Spandrels of San Marco and the Panglossian Paradigm: A Critique of the Adaptationist Programme, *Proc R Soc Lond B* 205: 581-598

Heyes, C: 1988. Are Scientists Agents in Scientific Change? *Biology and Philosophy* 3: 194-199

Holland J: 1995. *Hidden Order: How Adaptation Builds Complexity*, Helix Books/Addison-Wesley, Reading MA

Hull, DL: 1974. *The Philosophy of Biological Science*, Prentice-Hall, Reading MA

Hull, DL: 1987. Genealogical Actors in Ecological Roles. *Biology and Philosophy* 2: 168-184, reprinted in *The Metaphysics of Evolution* by DL Hull, State University of New York Press, Albany NY, 1989

Hull, DL: 1980. Individuality and Selection. *Annual Review of Ecology and Systematics* 11: 311-332

Hull, DL: 1988a. A Mechanism and its Metaphysics: An Evolutionary Account of the Social and Conceptual Development of Science. *Biology and Philosophy* 3:125-155

Hull, DL: 1988b. A Period of Development: A Response. *Biology and Philosophy* 3: 241-263

Hull, DL: 1988c. *Science as a Process: An Evolutionary Account of the Social and Conceptual Development of Science*, University of Chicago Press, Chicago IL

Hull, DL: 1992. Individual. In *Keywords in Evolutionary Biology*, ed. EF Keller, Harvard University Press, Cambridge MA

Kuhn, TS: 1962. *The Structure of Scientific Revolutions*, second edition 1970, University of Chicago Press, Chicago IL

- Laudan, L: 1977. *Progress and Its Problems*, University of California Press, Berkeley CA
- Lynch, A: 1996. *Thought Contagion: How Belief Spreads Through Society*, Basic Books
- Mayr, E: 1970. *Populations, Species and Evolution*, Harvard University Press, Cambridge MA
- Mayr, E and W Provine eds.: 1980. *The Evolutionary Synthesis: Perspectives on the Unification of Biology*, Harvard University Press, Cambridge MA
- Nagel, E: 1961. *The Structure of Science: Problems in the Logic of Scientific Explanation*, Routledge and Kegan Paul, London
- Plotkin, HC: 1994. *The Nature of Knowledge: Concerning Adaptations, Instinct and the Evolution of Intelligence*, Penguin, London
- Ridley, M: 1986. *Evolution and Classification: The Reformation of Cladism*, Longman, New York.
- Rosenberg, A: 1994. *Instrumental Biology, or the Disunity of Science*, University of Chicago Press, Chicago IL
- Suppe, F: 1989. *The Semantic Conception of Theories and Scientific Realism*, University of Illinois Press, Chicago IL
- Tennant, N: 1988. Theories, Concepts and Rationality in an Evolutionary Account of Science. *Biology and Philosophy*, 3: 224-231
- Toulmin, S: 1972. *Human Understanding*, Cambridge University Press, Cambridge UK
- van Fraassen, B: 1980. *The Scientific Image*, Clarendon, London UK
- Wiley, EO, D Seigel-Causey, DR Brooks, and VA Funk: 1991. *The Compleat Cladist: A Primer of Phylogenetic Procedures*, University of Kansas, Museum of Natural History, Lawrence, KA. Special Publication Number 19.
- Wilkins, J: forthcoming. The Evolutionary Structure of Scientific Theories. *Biology and Philosophy*
- Williams, GC: 1966. *Adaptation and Natural Selection: A Critique of Some Current Evolutionary Thought*, Princeton University Press, Princeton NJ
- Williams, GC: 1992. *Natural Selection: Domains, Levels, and Challenges*, Oxford University Press, Oxford UK
- Wynne-Edwards, VO: 1962. *Animal Dispersion in Relation to Social Behaviour*, Oliver and Boyd, London

# MEMES AIN'T (JUST) IN THE HEAD

## A Commentary on Gatherer's paper: Why the 'Thought Contagion' Metaphor is Retarding the Progress of Memetics"

By

John Wilkins

wilkins@wehi.edu.au

PO Box 542, Somerville 3912, Australia.

In a famous quip, Hilary Putnam once stated that "meanings just ain't in the head". It seems that the field of memetics and memetic theory is developing a couple of deep divisions: one between those who think of memes as genes and those who think of them as germs and the other between those who think memes are in the head, and those who think they ain't. In this essay I shall deal with these questions, but not attempt to further define what I think a meme is, which I have already done (Wilkins 1998a), based upon the cybernetic notion of an evolutionary gene of G. C. Williams (1966) and the Shannon-Weaver concept of a message:

A meme is the least unit of sociocultural information relative to a selection process that has favourable or unfavourable selection bias that exceeds its endogenous tendency to change. Derek Gatherer (1998) has written a provocative and interesting paper attacking the "thought contagion" notion of memetics and claims that no population model of memes is possible. He follows Benzon's (1996) earlier approach of locating memes in the overt behaviour and artifacts of culture. He and Benzon are both anti-head memeticists, and I suspect that they prefer to see memes as Mendelian genes, observed by their effects. His target is Aaron Lynch's (1998) pro-germs and pro-head view (overlooking here some of Lynch's subtleties), derived from the later views of Dawkins (1982, 1993).

Gatherer notices a shift in the meme definitions provided by Richard Dawkins, and calls them *Dawkins A*, and *Dawkins B*:

**Dawkins A:** '...a unit of cultural transmission, or a unit of imitation.' (Dawkins 1976, p.206); 'Examples of memes are tunes, ideas, catch-phrases, clothes fashions, ways of making pots or of building arches' (Dawkins 1976, p.206); 'Popular songs and stiletto heels are examples. Others, such as the Jewish religious laws...' (Dawkins 1976, p.209); 'Perhaps we could regard an organised church, with its architecture, rituals, laws, music, art and written tradition, as a co-adapted stable set of mutually-assisting memes.' (Dawkins 1976, p.212); 'Memes for blind faith have their own ruthless ways of propagating themselves.' (Dawkins 1976, p.213).

**Dawkins B:** (referring to the original Dawkins A definition, above) '...I was insufficiently clear about the distinction between the meme itself, as replicator, on the one hand, and its 'phenotypic effects' or 'meme products' on the other. A meme should be regarded as a unit of information residing in a brain..... It has a definite structure, realized in whatever physical medium the brain uses for storing information....I would want to regard it as physically residing in the brain.' (Dawkins 1982, p.109); 'The phenotypic effects of a meme may be in the form of words, music, visual images, styles of clothes, facial or hand gestures.....' (Dawkins 1982, p.109).

The different approaches I call the "genes" and the "germs" view of memes are based respectively on Dawkins B and Dawkins A. The genes view seeks to understand memes as analogous to the information units of selection and the causes of (in this case, sociocultural) evolution. Memes are units of cultural heredity, just as genes are units of biological heredity on this approach. The germs view seeks to explain the epidemiological spread of memes, like viruses through a population. It does not explicitly concern itself with selection or evolution. Gatherer rejects Dawkins B in favour of the following definition of a meme, which he is revising from William Benzon's (1996) definition:

*Meme: an observable cultural phenomenon, such as a behaviour, artefact or an objective piece of information, which is copied, imitated or learned, and thus may replicate within a cultural system. Objective information includes instructions, norms, rules, institutions and social practices provided they are observable.*

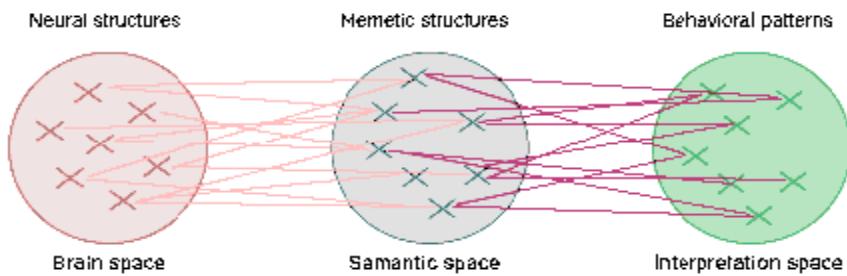
While I agree with much of Gatherer's criticism and many of his conclusions, especially of the thought contagion approach to memetics, I am moved to argue against his major claim that memes aren't to be located in the heads of individuals and to assert the possibility and even immediate attainability of a populational memetics. I do this from a philosophical perspective, and philosophers are notoriously poor at prescribing the best routes for science to take, particularly in the area of methodology, so I won't do that. Instead I'll do the traditional philosophical job of armchair criticism, and I expect to be taken in that light. Mostly, I will reject the inference from a specific denial to a general one, the move I call "simply because". My view could be seen as a compatibilism between both the germs/genes division and between the in-head/out-of-head division. Such is the wishy-washyness of being a philosopher.

*Objection 1 - Over generalisation.* My first objection is that he has overgeneralised. Simply because we can now have no population memetics, or because it fails in Gatherer's aptly chosen example of the Windsor knot, it does not follow that we cannot have it in any case. Gatherer wishes to extend the fact that we cannot easily, if at all, identify the knowing of "how to tie a Windsor knot" with the act of "tying a Windsor knot", and we can only say with certainty how many actual Windsor knots there are in a population. Again, just because in some cases we cannot say who "has" the Windsor knot meme; it is not thereby true that we can never know who has memes. Some memes are identifiable, and in quite specific terms. I have instanced the memes of scientific theories (Wilkins 1998a, 1998b) but I also expect that there will be a number of other examples of the attitudinal kind usually found in social surveys. These can be operationally defined by the process of iterative refinement exemplified by Grounded Theory in health research (cf Wilkins 1998b).

*Objection 2 - Over reliance on instrumentalism.* Simply because we cannot assay memes except through behaviour, which is quite true, it does not follow that memes of a mental kind are simply hypothetically postulated entities (or HYPEs). To argue this is to confuse information with its expression, the cause with the effect. It is true that we reason backwards from phenomenon to mechanism in science, but HYPEs have this habit of becoming observable, measurable, entities in their own right. Electrons are a case in point (Hacking 1983). Once the very model of a HYPE, and treated as an instrumental theoretical notion, useful for calculating but not much else, electrons are now routinely manipulated; sprayed, bathed and no doubt washed over objects for

our own purposes. Hacking suggested the Higgs Boson as an echt HYPE, but almost as soon as he did, it was observed. The process of natural selection itself has moved from being a hypothetical process to being an observed and repeatably manipulated phenomenon (Bell 1992). So the "simply because" here fails to support the blanket conclusion of the impossibility of mental memes.

However, the engram proposal is, I agree, quite wrong. The relation between neural net structures and dynamics on the one hand and cultural structures and dynamics on the other is many:many, and represents a Kolmogorov mapping between domains of over determination in both directions (figure 1). It is unhelpful to try to identify memes in terms of the neural structures that encodes them, because memes are multiply realized and neural nodes are involved in realising more than one meme, at least most of the time. While a single region of a brain may be involved in the storage and activation of a given concept or word, it does not follow that every brain does it in quite this way. Similar issues apply to genes and traits. While there happens to be a strong conservation of many structural and regulatory genes, such as Hox which is implicated in the expression of appendages, the variety of appendages so generated are very different. And not all appendages involve Hox or its orthologs. The "leg" genes are a heterogeneous set across lineages, what in taxonomy is called a polythetic set.



**Figure 1.** A Kolmogorov mapping of memes

*Objection 3 - Memes are units of information.* Although I agree with Gatherer that memes are not beliefs (in philosophy this is known as the distinction between propositions and doxastic attitudes to propositions), I disagree with him that memes are not sequences of information (and with Lynch 1998 that memes are single units; it is naive in the extreme to think that propositions are simple entities, in defiance of all logical and linguistic analysis since Frege). Memes are meaningful, semantic, information transmitted as a message from a source to a receiver. Note that semantic information is a notion that differs from the Shannon-Weaver concept of information, which is a signal-relative notion that may, or may not, have semantic content. The most "information-bearing" signal is one that is entirely random under Shannon-Weaver, because it takes a minimum message equal to the actual message to transmit it. Any semantic message is still a Shannon-Weaver message, but not all Shannon-Weaver messages are semantic messages. It follows that memes are messages, of the semantic variety.

Any message may, relative to an encoding protocol, be represented as a bit string, and so once we have such a conventional protocol (a formal "language" of the minimum kind) we can: Identify empirically, through census techniques, what the relative sizes and durations of memes are;

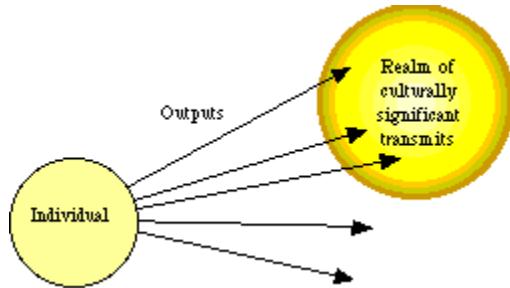
Deliver a populational and cluster analysis and measure changes in both instances and message structure over time

Relate differing homologous versions of memes; in other words, classify memes according to common ancestry (subject to the difficulties of such phyletic reconstructions and taxonomies caused by lack of information and loss of information due to homoplasies and reticulations in the evolutionary tree).

I shall say more on the structure of memetic information below. However, it must be noted that the observation of "objective" behaviours suffers exactly the same problems as semantic memes: we have no easy way to identify them, classify and relate them (i.e., distinguish homologies from analogies) and measure them purely on the external, observable properties and structures alone. This is because we need to have semantic information about them before we can do any of these things. Again, I shall say more about this shortly.

Gatherer rightly makes much of the notion of variations in memes in response to Lynch's naive conception of "mnemonics" (which, in the end, are just the "propositions" of late medieval Platonic philosophy). This point is well-made: anything that can be expressed as a sentence cannot be a unitary semantic entity. The sentence, "There is(are) <n> god(s)" is a complex structure with place holders for variables, as is the sentence, "Napoleon died in <year date>", and an infinite number of sentence tokens can be generated from the primitive relations (predicates) used to construct it. Using such primitive operators of the conventional natural language as "... died in ...", "some definite number of ...", along with the natural quantifiers "there exists" and cognates, and proper names and the natural numbers, etc, one can regenerate a literal infinity of statements from a very few, indeed quite small, number of stored and retrievable linguistic structures. If Chomsky's Transformational Grammar or something like it has any validity as a schematic for the neurology of language (cf Pinker 1994) then there is no real objection to memes being in the head in the sense that we can store a certain number of these constructions, and yet generate as many as needed on demand. Gatherer's objection would hold if we had to store all possible memes instead of only these variables and connectives. Instead, these semantic entities and their relations can be represented as a hypercube of however many dimensions there are particles. A given meme is therefore a well-formed coordinate in that resulting space. We can store some number of coordinates without having to store all of them, and yet be able to traverse this semantic space if it becomes important.

It's not up to philosophers to conjecture about mechanisms - that should be left to working researchers like Gatherer who use experiment and empirical data collection. However, I can make one suggestion: where memes might be located. If memes are messages, then they are entirely relative to considerations of encoding protocols, language (and logic) forms, and evaluative attitudes that determine recognition and responses. The meme therefore consists in the relationship between the neurological states and resulting behaviours of individuals and these contextual abstractions. This is illustrated in figure 2.



**Figure 2.** Culturally significant transmits as a subset of an individual's outputs

The output of the memes in an individual might be behaviours (Windsor knot tying), or artifacts (Windsor knots), or both. But a one-year old (let's call him Charlie) playing with Daddy's ties might randomly create the topology we call a Windsor knot, exhibiting neither the meme in thought, behaviour or in artifact. What counts here is the mental state - I hesitate to call it the intention, but at any rate the knowledge and understanding - its openness to inspection, the output, and the context... and above all the transmissibility of the result to others.

We have two entirely different issues here. What kind of entity, on the first hand, is a meme? On the other hand, what method and heuristic can we use to locate, measure, differentiate and model memes? Gatherer wishes to follow Benzon and locate them in physical structures, not of neurons and glial cells, but of artifacts and the actual behaviour patterns of cultures. Only these are measurable and objective. We can see a Windsor knot, and we can watch it being tied.

I have two objections to the Benzon-Gatherer "External Meme" thesis, but in so doing I do not wish to reassert the truth of Dawkins B - the thesis that memes are just units of information residing in brains. They are, I aver, informational structures, and they do reside in brains a lot of the time, but as I have argued before (Wilkins 1998a), they also often exist in the cultural relations between brains, and may not even exist in any brains at all at a time or ever.

Moreover, I expect that what goes on inside brains will sometimes be, not memes, but meme-like; call them "memoids" after Dawkins' term for design-like objects, "designoids" (Dawkins 1996). They will resemble memes in that they appear to be semantic structures encoded in neural networks, but their selection within the brain will often be hit-or-miss and inconsistent, despite Dennett's thesis of a strict internal selection (Dennett 1991, chapter 8). What counts isn't so much the selection process inside the head as the one outside, and that requires transmission.

The first objection to the exclusive External Meme thesis is again a "simply because" objection. Just because we can locate, measure, differentiate and model something is not reason to think that what we (now, or even in principle) cannot is nonexistent or nonsensical. This was the deep error of Skinnerian behaviourism, a holdover of old positivism, and it made of an epistemic limitation an ontological virtue. Although I am sure that neither Gatherer nor Benzon is a positivist, this is still an all-too-common error in science. While we may narrow down the search space of alternative HYPEs in this way, and Occam's Razor may seem to give weight to that approach, the history of science has shown (not least with the notion of a "gene") that it often leads to error and failure of research. We had better not kid ourselves that if we can't see it, it can't affect us, like kids hiding from monsters under the blanket at night.

My second objection is related also to the faults of Skinnerian behaviourism - the denial of what we are pretty sure is real and significant. Skinner denied the reality of psychological states - the External Meme thesis seems to be predicated on the denial of concepts. Questions raised by Gatherer of what in philosophy is called the K-K principle (if we know, we must know that we know, ad infinitum) aside, surely if I know how to tie a Windsor knot, some causally relevant structure exists inside my brain? And it must be one that satisfies the relevant semantic relation to the culture that defines Windsor knots and which therefore makes it "about" Windsor knots.

Now I am not here trying to revivify the extreme reductionism of engrams, the one meme-one engram theory, if ever it was seriously proposed. Memes are, in my perspective, multiply realized and multiply instantiated by neural net structures (as neural nets are wont to do). The view proposed by Delius (1991) that a meme will have a single locatable and observable structure ("a constellation of activated neuronal synapses") is not viable. In philosophical jargon, memes are not supervenient properties, for a property that supervenes on a physical system must be identical if the physical system is. The same meme might be realized in a number of physical systems, consistently with supervenience, but identical or very closely similar neural nets might encode quite distinct memes. This is because memes are relations in semantic space, not neurological space.

Let me explain by returning to Charlie's Windsor knot accident. If his knot matches exactly that of the very metre of Windsor knots, Prince Phillip's butler, when none of the "normal" intentions or concepts are possible, then the "Windsor knot" that results is only an accident. We would not call it a cultural artifact if the wind "tied" a bullrush in that knot, and we wouldn't here, except in an honorary sense. There is a continuum of intentionality and "aboutness" from the bullrush to Prince Phillip's butler. There is no absolute point at which mere topology makes that meme, and the resemblance the bullrush has for an observer is merely subjective - it looks like one so the observer, who has learned about Windsor knots, "reads" that topology that way.

But what makes that knot a Windsor knot is the fact that it results from the transmission from one agent to another (with, I suppose, some Urbutler in the House of Windsor as the Last Common Ancestor) of the "how-to" of Windsor knot tying. The more-or-less formal algorithm is what gets transmitted. This constitutes a message, an abstract array of symbols that are interpreted in the context of a consensual, or conventional, language and set of behaviours and standards. So, we cannot really say that it is the physical knot, nor even the tying behaviour, that is the meme, at least not without the interpretive context and semantic structures of a culture.

So, can we make memetics an operational science at all? I believe we can, and the reason is that in many cases we already have just the semantic apparatus needed to identify memes, and that resolving memes is a matter of resolving them enough to identify them. Goodenough and Dawkins (1994) had no trouble identifying instances of the St Jude Chain letter, even across cultural divides, and with a suitable knowledge, say, of Spanish, one could identify instances of it across languages. For most situations being investigated, seeing a Windsor knot being tied is evidence enough that there is a Windsor knot meme about. Eccentric examples of multiple personalities and other exotic exemplars are not enough to shake our confidence that there are memes, that usually people have them, and that we can count them well enough to model them in a populational, statistical, manner.

A general observation: Gatherer rightly worries about the academic respectability of memetics, just as serious researchers have worried about, say, the Gaia hypothesis - which is only now coming into serious contention, even by such respectable academics as William Hamilton (Hamilton and Lenton 1998) (of Hamiltonian inclusive fitness) - or complexity theory. The debasement of memetics by quick and easy metaphors and popularised science to serve metaphysical agendas and political ideologies, with which we are all too familiar, is just the latest instance of serious evolutionary theory being perverted in that way, beginning with Spencer and the Edinburgh radicals of the 1840s, through social "Darwinism", "cultural evolution" theories, eugenics positive and negative, and so forth.

I'm of the view that the term memetics is now a liability, at least until it can be rehabilitated with some actual research instead of conjecture and anecdote. I tend to think of this whole field as cultural Darwinism, stressing the essentially amoral and nonevaluative dynamics of culture, and drawing no conclusions about what should be transmitted, instead seeking to know why things are transmitted, and adopting the generally pluralistic nature of Darwinism.

In Darwinism proper not everything must be explained in terms of its adaptation - Sewall Wright is centrally installed in modern synthetic Darwinism. Memetics is analogous to genetics not selection, and it is a subdiscipline of cultural Darwinism dealing with the evidence for causal factors that create the behaviours and artifacts we see and their rates of transmission.

As soon as we use a metaphor like "thought contagion" we are insert all kinds of socially-weighted evaluations associated with the terms "pathogen", "parasite", "disease" and the like, none of which play a useful role in a Darwinian model of any kind of change. The arbitrary difference between a predator and a parasite is the rate at which it kills its host, but we tend to think of parasites as, well, parasitic and therefore nasty, and predators as having a more noble ecological role. But in the end there is only selection and drift, and I refer readers to Paul Ewald's excellent treatment of this topic (Ewald 1994) for a fuller discussion.

## Acknowledgements

Bill Benzon and Derek Gatherer have engaged me in electronic conversation on this topic over a period of time. Benzon was good enough to send me a selection of his papers, and Prof Juan Delius a copy of his paper, which makes my present disagreement with them an act of ingratitude. Despite my conversations with them and many others, including Mark Mills, for which I am continually grateful, nobody is responsible for the errors contained herein but me.

## References

- Bell, G. (1997). *Selection: The Mechanism of Evolution*. London: Chapman and Hall.
- Benzon, W. (1996). Culture as an evolutionary arena. *Journal of Social and Evolutionary Systems*, 19:321-362. <http://www.newsavanna.com/wlb/CE/Arena/Arena00.shtml>
- Dawkins, R. (1976). *The Selfish Gene*. Paperback edition 1978. Granada Paladin, London

Dawkins, R. (1982). *The Extended Phenotype*. Paperback edition 1983. Oxford University Press, Oxford UK.

Dawkins, R. (1993). Viruses of the mind. In: *Dennett and his Critics*. Dahlbom, B. (Ed.) Blackwell, 13-27.

Dawkins, R. (1996). *Climbing Mount Improbable*. NY: Viking.

Delius, J.D. (1991). The nature of culture. In: *The Tinbergen Legacy*, MS Dawkins, TR Halliday and R Dawkins (eds.). London: Chapman and Hall.

Dennett, D.C. (1991). *Consciousness Explained*. NY: Little Brown.

Ewald, P. (1994). *Evolution of Infectious Disease*. Cambridge: Cambridge University Press.

Gatherer, D. (1998). Why the 'Thought Contagion' metaphor is retarding the progress of memetics. *Journal of Memetics - Evolutionary Models of Information Transmission*, **2**. [http://jom-emit.cfpm.org/1998/vol2/gatherer\\_d.html](http://jom-emit.cfpm.org/1998/vol2/gatherer_d.html)

Goodenough, O.R. and R. Dawkins (1994). The 'St Jude' mind virus. *Nature*, **371**:23-24.

Hacking, I. (1983). *Representing and Intervening*. Cambridge: Cambridge University Press.

Hamilton, W.D. and Lenton T.M. (1998). Spora and Gaia: How microbes fly with their clouds. *Ethology Ecology and Evolution*, **10**:1-16.

Lynch, A. (1998) Units, Events and dynamics in memetic evolution. *Journal of Memetics - Evolutionary Models of Information Transmission*, **2**. [http://jom-emit.cfpm.org/1998/vol2/lynch\\_a.html](http://jom-emit.cfpm.org/1998/vol2/lynch_a.html)

Pinker, S. (1994). *The Language Instinct*. London: Penguin.

Wilkins, J. S. (1998a). What's in a meme? Reflections from the perspective of the history and philosophy of evolutionary biology. *Journal of Memetics - Evolutionary Models of Information Transmission*, **2**:2-33. [http://jom-emit.cfpm.org/1998/vol2/wilkins\\_js.html](http://jom-emit.cfpm.org/1998/vol2/wilkins_js.html)

Wilkins, J.S. (1998b). The evolutionary structure of scientific theories. *Biology and Philosophy*, **13**: 479-504. <http://www.wehi.edu.au/~wilkins/theories/evoltheory.html>

Williams, G.C. (1966) *Adaptation and Natural Selection: A Critique of Some Current Evolutionary Thought*. Princeton NJ: Princeton University Press.

# **ON CHOOSING TO EVOLVE: STRATEGIES WITHOUT A STRATEGIST**

## **A Commentary on Rose's Paper: "Controversies in Meme Theory"**

By  
**John Wilkins**  
wilkins@wehi.edu.au  
PO Box 542, Somerville 3912, Australia.

In his 1974 paper Michael Ghiselin wrote that while there were evolutionary strategies, the only evolutionary strategists were the biologists studying evolution. Nick Rose (1998) has raised several interesting problems not only for memetics but evolutionary thinking in general. To avoid the Scylla of Lamarckism, he has chosen instead the Charybdis of denying the bleeding obvious - that intentions play a part in evolution. Before anyone accuses me of deserting the Darwinian fold for Lamarck, let me explain.

It seems that the reason Rose takes this tack and follows Dennett's attack on the Cartesian Theatre of the Self (Dennett 1992, 1995) is that if we admit the role of intentions into evolution, we end up with some kind of Piagetian Lamarckism (Piaget 1979) and lose the strength of the blind watchmaker schema - evolution through natural selection (plus sundry other theoretical bits like drift). Must it be so? Are intentions really that dangerous?

Usually when the charge of Lamarckism is raised, the spectre of clairvoyance lies not too far beneath it (Hull 1988). We choose to solve problems, we seek solutions, we adapt by learning, not by random variation and selective retention (Campbell 1960, Popper 1972). If we take this approach, then evolutionary explanations of memes are redundant and indeed otiose. Better to deny that we can choose and control memes at all. If folk psychology admits a causal role for intentions, so much the worse for folk psychology. I think this is wrong. We can have our memetic cake and eat intentions too, if we recast matters just a bit.

The lion, contra Wittgenstein (1968) can surely be understood to have intentions to catch the gazelle. Just as surely, if a gazelle thinks and intends at all, it intends not to be caught by the lion. Yet, the outcome is an evolutionary one, and purely Darwinian. How can this be? Didn't we just stipulate that intentions make Darwinism otiose? Well, there is an obvious difference between intending to solve a problem and actually achieving the solution. The problem of induction, and its heirs like Nelson's 'Grue' Paradox (Goodman 1973), shows that intentions are not endowed with magical clairvoyance. After all, we humans have intentions that pave many roads with varying destinations. Selection operates on outcomes not the provenance of variations.

Having intentions and being able to learn are at once both biasing effects on the selective landscape and also adaptations in themselves - one reason why we do have the cognitive faculties we do is that it broadens the range of adaptation. If we can learn, we can deal with challenges without waiting for mutations, but once we have the novel lifestyle, selection can catch up as soon as more cost-effective instinctual behaviours appear genetically. This is called the Baldwin Effect (Belew and Mitchell 1996, Turney et al. 1996) and is perfectly Darwinian. Learning deforms the fitness landscape.

Ignore, for a moment, the supposed psychic powers of intentionality and consider instead trait U. This trait permits an organism to deal more effectively with its environment than non-U organisms can. It is not perfect, because the environment can always bowl a spin ball, but it does bias the odds in favour of U-bearers as opposed to non-Us. It will obviously spread through its deme so long as selection is stronger than drift. There's nothing too problematic about this. Now, substitute 'intention' for U, and see if it makes the slightest difference to the Darwinian nature of the model. I submit that this does not. So, how does this biological argument bear on memes? If memes are selected for or against intentionally on the basis of prior experience and propensities to estimate the likely success of a given strategy, that will deform the fitness landscape for those memes and meme-bearers. We pick those memes that we estimate will succeed. But no matter how good we are at picking strategies, our intentions merely bias, they do not determine the outcomes. Our social ecology is not totally plastic any more than the biological ecology is. We cannot, on pain of sinning against St. David Hume, predict with complete accuracy what comes our way next.

The strategies we inherit are not good because we rationally analysed them. Indeed, we inherit them because they are good ones, for whatever reason. No grand Cartesian General planned them, and the presence or absence of a minor strategist does not impugn evolutionary explanation. Quine once wrote (Quine 1969) that "*creatures inveterately wrong in their inductions have a pathetic but praiseworthy tendency to die before reproducing their kind*" (p 126). Social agents that are ordinarily wrong about good ways to proceed tend not to have much influence or leave many intellectual descendants (the praiseworthiness is left as a judgment for the reader). So, we had better be good at choosing our memes if we want to have any influence.

However, come the next revolution, many of our strategies may find themselves first against the wall. Selection reigns in social games as much as in the struggle for existence. If wishes were horses, then beggars would ride. No clairvoyance, just good old trial and error. I believe that any Lamarckian process of evolution can be reformulated as a Darwinian process. The difference is one of rate, degree of novelty, and fidelity of transmission (Fisher 1930). Dennett (1995) notes that memetic selection occurs first in the head, then locally and finally more broadly (Cziko 1995, Wilkins 1998). But although the selective filters of our mind are not a Cartesian Self, that there exists an intentional social agent is silly to deny. It is bounded by the extent of direct control of our body by the central nervous system. To agree with Wittgenstein this time, the best model of the human soul is the human face.

An intending self is the social agent, and it has some control over the kinds of memes it hosts, not unlike the limited control our immune system has over the sorts of microorganisms that a body hosts. Cognitive structures and processes behave something like the immune system. Good immune systems, and good memetic selection patterns, get more copies of themselves passed on. It may make it harder to model things, but we have to live with intentions in memetics.

## References

- Belew R.K. and Mitchell, M. (eds.). (1996). *Adaptive Individuals in Evolving Populations: Models and Algorithms*. Massachusetts: Addison-Wesley.

Campbell, D. T. (1960). Blind Variation and Selective Retention in Creative Thought as in Other Knowledge Processes. *Psychological Review*, 67:380-400.

Cziko, G. (1995). *Without Miracles: Universal Selection Theory and the Second Darwinian Revolution*. Cambridge MA: MIT Press. [http://www.ed.uiuc.edu/facstaff/g-cziko/without\\_miracles/](http://www.ed.uiuc.edu/facstaff/g-cziko/without_miracles/).

Dennett, D. C. (1992). *Consciousness explained*. London: Allen Lane, The Penguin Press..

Dennett, D. C. (1995). *Darwin's dangerous idea: evolution and the meanings of life*. New York: Simon & Schuster.

Fisher, R. A. (1930). *The Genetical Theory of Natural Selection*. Oxford: Clarendon Press.

Ghiselin, M. T. (1974). *The economy of nature and the evolution of sex*. Berkeley: University of California Press.

Goodman, N. (1973). *Fact, fiction, and forecast*. Indianapolis: Bobbs-Merrill.

Hull, D. L. (1988). *Science as a process: an evolutionary account of the social and conceptual development of science*. Chicago: University of Chicago Press, 1988..

Piaget, J. (1979). *Behaviour and evolution* Trans. Nicholson-Smith, D., London: Routledge & Kegan Paul.

Popper, K. R. (1972). *Objective knowledge; an evolutionary approach*. Oxford: Clarendon Press..

Quine, W.V. (1969). *Ontological Relativity and Other Essays*. New York: Columbia University Press.

Rose, N. (1998). Controversies in Meme Theory. *Journal of Memetics - Evolutionary Models of Information Transmission*, 2. [http://jom-emit.cfpm.org/1998/vol2/rose\\_n.html](http://jom-emit.cfpm.org/1998/vol2/rose_n.html)

Turney, P. Whitley, D. and Anderson, R. (1996). Evolution, Learning, and Instinct: 100 Years of the Baldwin Effect. *Evolutionary Computation*, 4:213-234. <http://ai.iit.nrc.ca/baldwin/toc.html>.

Wilkins, J. S. (1998). What's in a Meme? Reflections from the perspective of the history and philosophy of evolutionary biology. *Journal of Memetics - Evolutionary Models of Information Transmission*, 2. [http://jom-emit.cfpm.org/1998/vol2/wilkins\\_js.html](http://jom-emit.cfpm.org/1998/vol2/wilkins_js.html).

Wittgenstein, L. (1968). *Philosophical investigations*. Trans. Anscombe, G.E.M., Oxford: Basil Blackwell.