Towards a Cognitive Memetics: Socio-Cognitive Mechanisms for Memes Selection and Spreading

Article · January 2001

1 author:

Cristiano Castelfranchi
Italian National Research Council
422 PUBLICATIONS **14,497** CITATIONS

SEE PROFILE

Towards a Cognitive Memetics:
Socio-Cognitive Mechanisms for Memes Selection and Spreading

Cristiano Castelfranchi
University of Siena - Dep. of Communication Sciences
Siena - Italy
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 behaviour; they differ in the interpretation of the observed behaviour 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 behaviour: 'instrumental adoption' based on means-end reasoning and on practical utility (section 2.); normative adoption, where certain behaviours 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 behaviour 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 behaviour 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 modelling.

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 flavour, 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 behaviours, 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 (behaviour, idea, artefact), 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 behaviour, an artefact 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 a 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 behaviour 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 behaviours in Cognitive Agents. As stressed by Conte [17] "behaviour does not spread automatically but through the agent' mind":

*agents have to observe a behaviour; to interpret/understand it; and then to have their own **motives** and **reasons** for performing (and store and repeat) such a behaviour.*

Only on the basis of their understanding and of their motives and reasons they can replicate the same behaviour or some accidental or adjusted variation of it.
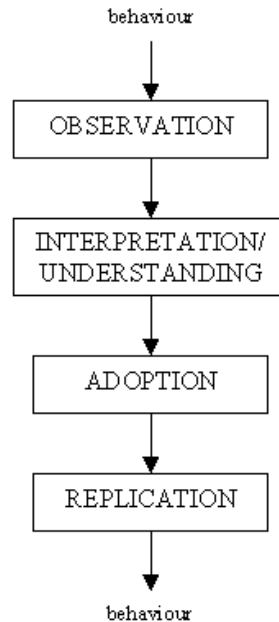

Figure 1.

It is important to notice that

a) the behaviour 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 realisations 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 behaviour 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 behaviour, 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 behaviour, an artefact 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 characterise this mechanism as follows:

Given an Adopter A, a trait x and a Model M,
<u>conditions for 'ADOPT A x from M'</u> 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 that 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 behaviour.

In sum, *individuals accept new behaviours, plans or tools as better solutions for their own problems, as good means for their goals*; the group diffuse and preserve (memorise) 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 behaviour 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 behaviour 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 behaviours: 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 behaviour 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 behaviour, 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 recognised 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 behaviours, goals, ideas; they provide Models and prescribe who should or can do what.
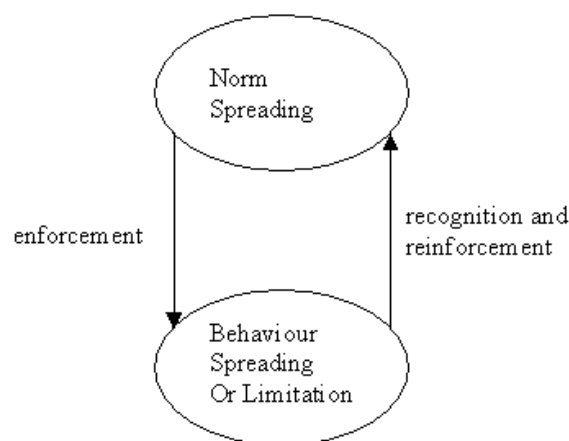
Figure 2.

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 recognise 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 behaviour 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 behaviours 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 'civilisation' 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 (behaviours, attitudes, use of artefacts, etc.) are *symbols* of this membership (and status) and are used for recognising a person as "one of them", [note 12]

     and

I want to send such a message and to be recognised 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 recognised as "one of them"
feel I am one of them and recognise 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, behaviours (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 againthe *selection process* as a *mental process*, and as deciding in favour or against of the trait adoption and determining the macro-spreading or not of a given behaviour.

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 behaviours 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 behaviour. 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 behaviour adoption and its replication by the individual are influenced by its *perceived diffusion*, while actual diffusion -in turn- depends on individual agent adoption:
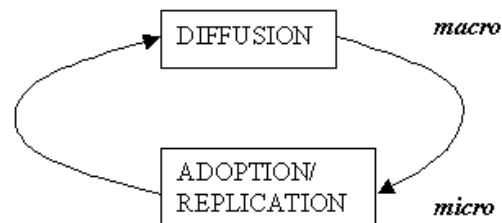

Figure 3.

As we said, the individual mind is not the *only* selecting environment; for example, diffusion in population plays a role by itself:

<div align="center">

DIFFUSION ==> DIFFUSION

</div>

For example the probability of survival of a given belief or behaviour 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*:

<div align="center">

Actual DIFFUSION==>(perceived-diffusion/ interpretation/ adoption) ==> Actual DIFFUSION

</div>

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 '**immergence**' 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 behavioural 'imitation' (that in a cognitive agent - whose 'behaviour' 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" modelling 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 means 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 it 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 in 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 behaviour, 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 defence 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 that 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 organisation (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 likeabeality... and even its moral appropriateness, but also the mere memorisation [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 behaviours 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.
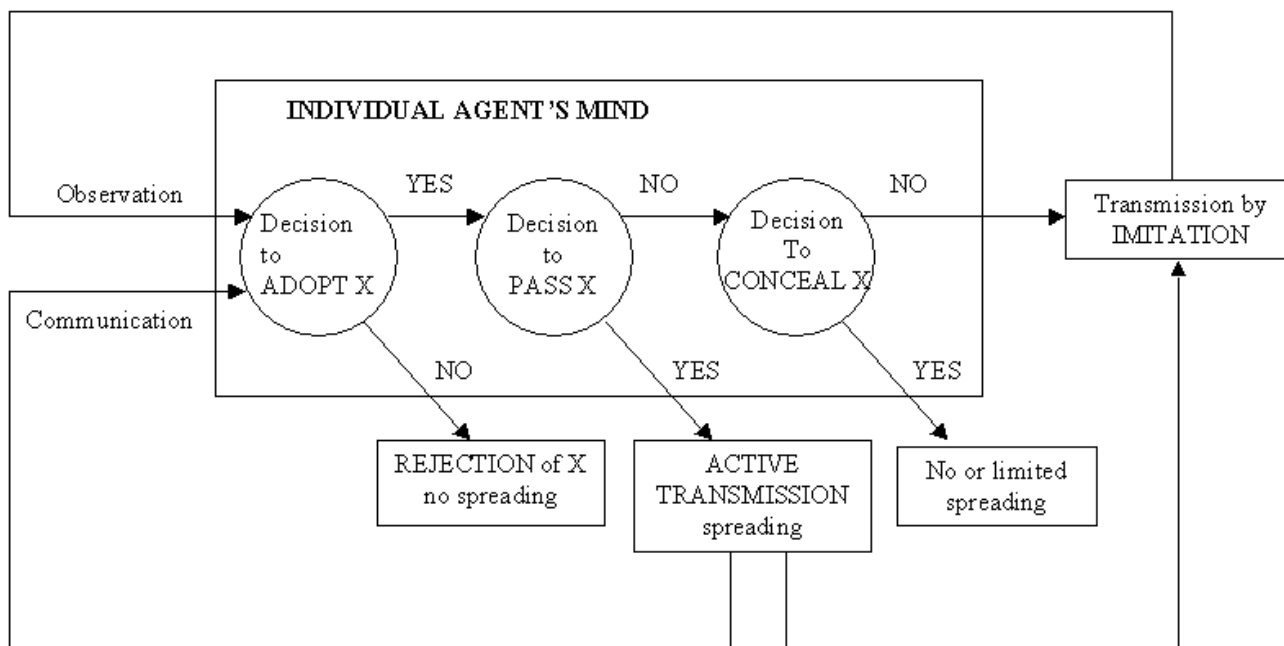

Figure 4.

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, behaviour, 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 behaviour ('behavioural 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 behaviour 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 behaviour (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 behaviour where the Model is passive and indifferent and adoption is a mere individual choice. The same holds for prescribed behaviours and values diffused either through the decentralised and diffuse social control or through enforcing authorities (or both). Analogously, social monitoring and moral or legal prohibition of given behaviours 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 modelling 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 analyse when and why sharing knowledge or diffusing a behaviour is a conscious goal of the agent or is just an unintended functional effect of its behaviour [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 modelled, 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 modelling 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 explicitily discuss here- see Sperber [35]. [back to note mark]

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 behaviours. 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. [back to note mark]

3. For example to its "causal attribution" of the event; for example to and within the activated schemata, frames, or scripts [33]. [back to note mark]

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? [back to note mark]

5. For a nice example of different interpretations of the same bahaviour -with different consequences on spreading- see also [17]. [back to note mark]

6. To be read as "A Believes that M uses x for the goal p". [back to note mark]

7. To be read as "A Believes that M has the goal p" and "A Believes that x is good/useful for p". [back to note mark]

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. [back to note mark]

9. In this work, arguments and distinctions are very good, but the very defintion of "rational imitation" is rather unsatisfactory. [back to note mark]

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 techincal (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.). [back to note mark]

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. [back to note mark]

12. Another very nice example of this goal of "being like them; being recognised as one of them" (intentional imitation), but for completely different reasons and motives, is camouflage. Consider a spy who does not like to be recognised 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. [back to note mark]

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. [back to note mark]

14. Otherwise it would be better to assume as a principle: *Believe something if you do not have reasons for rejecting it*. [back to note mark]

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. [back to note mark]

16. Consider that the notion of 'decision' can be interpreted in more or less 'deliberated' and conscious way (section 1.1). [back to note mark]

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! [back to note mark]

Acknowledgments

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 Algoritms 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://cfpm.org/jom-emit/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://cfpm.org/jom-emit/1998/vol2/wilkins_js.html
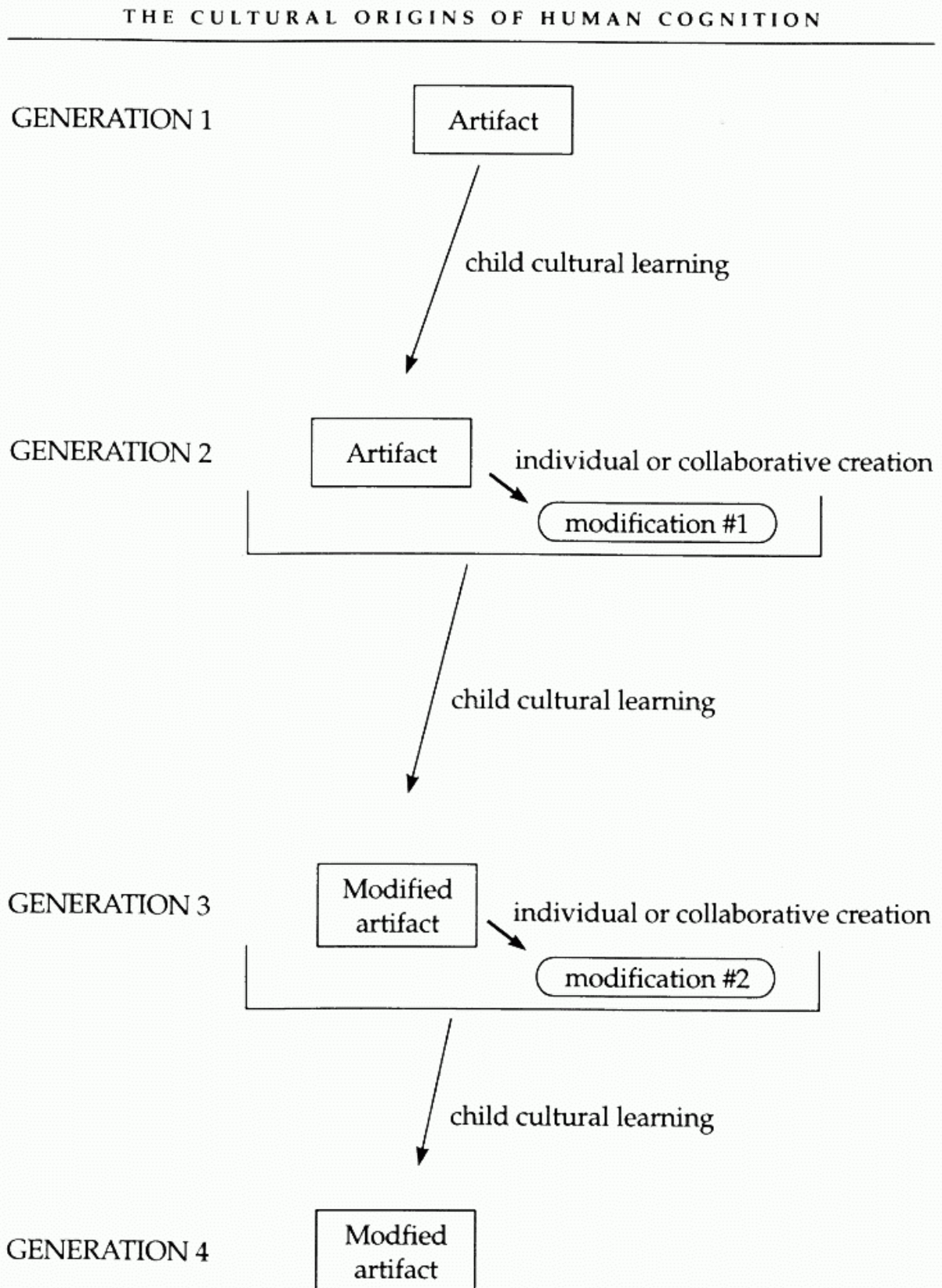
Appendix

THE CULTURAL ORIGINS OF HUMAN COGNITION

GENERATION 1

Artifact

child cultural learning

GENERATION 2

Artifact   individual or collaborative creation

modification #1

child cultural learning

GENERATION 3

Modified
artifact   individual or collaborative creation

modification #2

child cultural learning

GENERATION 4

Modfied
artifact

Figure 2.2 A simplified depiction of the ratchet effect working to produce an artifact with accumulating modifications.

Back to Issue 1 Volume 5