Reprinted from

PROCEEDINGS OF THE BRITISH ACADEMY · 88

EVOLUTION OF SOCIAL BEHAVIOUR PATTERNS IN PRIMATES AND MAN

W.G. Runciman, J. Maynard Smith, and R.I.M. Dunbar (editors) Oxford University Press

Proceedings of the British Academy, **88**, 119–143

Friendship and the Banker's Paradox: Other Pathways to the Evolution of Adaptations for Altruism

JOHN TOOBY & LEDA COSMIDES
Center for Evolutionary Psychology,
University of California, Santa Barbara, CA 93106, USA

Keywords: reciprocity; altruism; co-operation; social exchange; reciprocal altruism; evolutionary psychology.

Summary. The classical definition of altruism in evolutionary biology requires that an organism incur a fitness cost in the course of providing others with a fitness benefit. New insights are gained, however, by exploring the implications of an adaptationist version of the 'problem of altruism', as the existence of machinery designed to deliver benefits to others. Alternative pathways for the evolution of altruism are discussed, which avoid barriers thought to limit the emergence of reciprocation across species. We define the Banker's Paradox, and show how its solution can select for cognitive machinery designed to deliver benefits to others, even in the absence of traditional reciprocation. These models allow one to understand aspects of the design and social dynamics of human friendship that are otherwise mysterious.

FROM A SELECTIONIST TO AN ADAPTATIONIST ANALYSIS OF ALTRUISM

THE ANALYSIS OF THE EVOLUTION OF ALTRUISM has been a central focus of modern evolutionary biology for almost four decades, ever since Williams, Hamilton, and Maynard Smith caused researchers to appreciate its significance (Williams & Williams 1957; Hamilton 1963, 1964; and Maynard

CRISIS MANAGEMENT

The Banker's paradox

> If thou wouldst get a friend, prove him first, and be not hasty to credit him. For some man is a friend for his own occasion, and will not abide in the day of thy trouble... Again, some friend is a companion at the table, and will not continue in the day of thy affliction... If thou be brought low, he will be against thee, and will hide himself from thy face... A faithful friend is a strong defence: and he that hath found such a one hath found himself a treasure. Nothing doth countervail a faithful friend... *From Ecclesiastes 6*

Many people become angry when they first hear the evolutionary claim that the phenomenon of friendship is solely based on the reciprocal exchange of favours, and deny that their friendships are founded on such a basis. Similarly, many people report experiencing a spontaneous pleasure when they can help others without any expectation or anticipation of reward. Their memory of the pleasure is not diminished by not ever having received a reward in return. Indeed, explicit linkage between favours or insistence by a recipient that she be allowed to immediately 'repay' are generally taken as signs of a lack of friendship. What is going on? One widely accepted interpretation is that these denials are simply the deceptive surface of human social manipulation. We think, however, that narrow exchange contingency does not capture the phenomenology or indeed the phenomenon of friendship. We propose that the altruistic adaptations that underlie friendship do not map onto the structure of tit for tat or any other standard model of reciprocal altruism based on alternating sequences of contingent favours.

One dimension of difference is illustrated by what we will call the *Banker's Paradox*. Bankers have a limited amount of money, and must choose who to invest it in. Each choice is a gamble: taken together, they must ultimately yield a net profit, or the banker will go out of business. This set of incentives leads to a common complaint about the banking system: that bankers will only loan money to individuals who do not need it. The harsh irony of the Banker's Paradox is this: just when individuals need money most desperately, they are also the poorest credit risks and, therefore, the least likely to be selected to receive a loan.

This situation is analogous to a serious adaptive problem faced by our hominid ancestors: exactly when an ancestral hunter-gatherer is in most dire need of assistance, she becomes a bad 'credit risk' and, for this reason, is less attractive as a potential recipient of assistance. If we conceptualize contingent benefit-benefit interactions as social exchange (rather than more narrowly as reciprocation), then individuals rendering assistance can be seen as facing a series of choices about when to extend credit and to whom.

Assisting one individual may take time, resources, or be dangerous to oneself—it therefore precludes other worthwhile activities, including assisting others. From this perspective, exchange relationships are analogous to economic investments. Individuals need to decide who they will invest in, and how much they will invest. Just as some economic investments are more attractive than others, some people should be more attractive as objects of investment than others.

Computational adaptations designed to regulate such decisions should certainly take into account whether an individual will be willing to repay in the future (i.e., are they a cheater?). But they should also assess whether the person will be in a position to repay (i.e., are they a good credit risk?), and whether the terms of exchange will be favourable (will this exchange partnership ultimately prove more profitable than the alternatives it will preclude?). If the object of investment dies, becomes permanently disabled, leaves the social group, or experiences a permanent and debilitating social reversal, then the investment will be lost. If the trouble an individual is in increases the probability of such outcomes when compared to the prospective fortunes of other potential exchange partners, then selection might be expected to lead to the hardhearted abandonment of those in certain types of need. In contrast, if a person's trouble is temporary or they can easily be returned to a position of full benefit-dispensing competence by feasible amounts of assistance (e.g., extending a branch to a drowning person), then personal troubles should not make someone a less attractive object of assistance. Indeed, a person who is in this kind of trouble might be a more attractive object of investment than one who is currently safe, because the same delivered investment will be valued more by the person in dire need. The attractiveness of extending the branch can be compared to nursing someone with a life-threatening disease for months: the cost is high, and the outcome is uncertain.

For hunter-gatherers, illness, injury, bad luck in foraging, or the inability to resist an attack by social antagonists would all have been frequent reversals of fortune with a major selective impact. The ability to attract assistance during such threatening reversals in welfare, where the absence of help might be deadly, may well have had far more significant selective consequences than the ability to cultivate social exchange relationships that promote marginal increases in returns during times when one is healthy, safe, and well-fed. Yet selection would seem to favour decision rules that caused others to desert you exactly when your need for help was greatest. This recurrent predicament constituted a grave adaptive problem for our ancestors—a problem whose solution would be strongly favoured if one could be found. What design features might contribute to the solution of this problem?

Becoming irreplaceable: The appetite for individuality

One key factor is replaceability or substitutability. Consider X's choice between two potential objects of investment, Y and Z. Each helps X in different ways; the magnitude of the benefits Z delivers are higher than the magnitude of the benefits that Y delivers, but the types of benefits that Y supplies can be supplied by no one else locally. Consider the alternative payoffs when one or the other enters a crisis and requires help. Extending 'credit' to a person in crisis may easily have a negative payoff if the *kind* of benefits that she customarily delivers could be easily supplied by others. To the extent an individual is in social relationships in which the assistance she delivers to her partners could easily be supplied in her absence by others, then there would be no necessary selection for her partners to help her out of difficulty. A 'replaceable' person would have been extremely vulnerable to desertion. In contrast, extending credit has a higher payoff if the person who is currently in trouble customarily delivers types of benefits (or has some other value) that would be difficult to obtain in her absence. Selection should favour decision rules that cause X to exhibit loyalty to Y to the extent that Y is irreplaceably valuable to X. In other words, Y's associates will invest far more in rescuing her than they would if she lacked these unique distinguishing properties (Tooby & Cosmides 1984, 1989a). Y may be helped, and Z abandoned even though the benefits Z delivers are greater.

If Banker's Paradox dilemmas had been a selection pressure, then one would expect to see adaptations that caused humans to:

1 have an appetite to be recognized and valued for their individuality or exceptional attributes;

2 be motivated to notice what attributes they have that others value but cannot obtain as easily elsewhere;

3 be motivated to cultivate specialized skills, attributes, and habitual activities that increase their relative irreplaceability;

4 be motivated to lead others to believe that they have such attributes;

5 preferentially seek, cultivate, or maintain social associations and participate in social groups where their package of valued attributes is most indispensable, because what they can differentially offer is what others differentially lack;

6 preferentially avoid social circles in which what they can offer is not valued or is easily supplied by others; and,

7 be jealous or rivalrous when someone within their social circle develops abilities to confer similar types of benefits, or when someone with similarly valued attributes enters their social circle. Such jealousy would motivate and organize actions that drive off attribute-rivals and that inhibit

individuals who value the actor from developing potential relationships with others who could supply the same type of assistance.

Although we are unaware of any experimental studies specifically of these traits, we think many aspects of human social and mental life show clear evidence of them. Much of social life seems to consist of a continual movement to find and occupy individualized niches that are unusually other-benefitting but hard to imitate, accompanied by a shuffling of social associations in search of configurations where the parties are most highly mutually valued. Indeed, the cross-culturally general motivation for status (as opposed to dominance) is arguably a product, in some measure, of this kind of selection pressure. Calling someone irreplaceable, or stressing how they will be (or have been) missed is a ubiquitous form of praise. Many other phenomena seem to be obvious expressions of a psychology organized to deal with the threat of social replaceability. These include everything from complaints about feeling anonymous in modern mass societies to the incessant fissioning off of smaller social groups whose members cultivate a mutual sense of belonging and discourage transactions with outgroup members. More significantly, the growth of irreplaceability as a feature of hominid life would have had powerful secondary impacts on hominid evolution. For example, individuals could pursue more productive, but more injury producing subsistence practices, such as large game hunting.

The motivation to discover and occupy unique niches of valued individuality is facilitated by the many forces that act to spontaneously locate individuals in unique 'starting positions' (Tooby & Cosmides 1988). These include, obviously, the fact that each individual's talents and shortcomings will be somewhat different due to random genetic variation, the accidents of ontogeny, and the different kinship, demographic and social circumstances they are born into. One might expect selection for adaptations that guide an individual not only to hone those skills that she can do well in an absolute sense, but to put special effort into those skills that she does relatively well, so that she 'product-differentiates' herself. Indeed, the most common and basic meaning humans apply to the issue of ability-acquisition is a social meaning—ability relative to others—rather than an absolute standard. Competences that everyone shares are not even noticed. In any case, Plomin & Daniels' work (1987) on the effects of nonshared environment provides strong evidence that individuals do product-differentiate themselves, even among their siblings, as does Sulloway's pioneering work on birth order (forthcoming).

Fair weather friends and deep engagement

The archetypal concept of the fair weather friend implies that there is also another kind of friend, a close or true friend—someone who is deeply

engaged in your continued survival and in your physical and social welfare (but not necessarily in promoting the propagation of the genes you carry). It is this kind of friend that the fair weather friend is the counterfeit of. If you are a hunter-gatherer with few or no individuals who are deeply engaged in your welfare, then you are extremely vulnerable to the volatility of events—a hostage to fortune. Indeed, the higher the variance or volatility of the environment inhabited, the more individuals ought to care about friendships.

But if you wait until you are in trouble to determine whether anyone cares, it may be too late, if the answer is 'no'. When times are good, close friends who are deeply committed to you and casual exchange partners for whom you are replaceable may behave very similarly to each other. Moreover, since it is advantageous for anyone to be categorized as a close friend by someone who is not in difficulty, humans face the adaptive problem of friendship mimicry. The adaptive problem of discriminating true friends from fair weather friends would have been a formidable signal detection problem for our ancestors. One would expect the human psychological architecture to contain subsystems designed to sift social events for cues that would reduce uncertainty about the relative engagement different individuals have in one's welfare, i.e., assess the genuineness of friendship. Of course, the most ecologically valid evidence is what people actually do when you are genuinely in trouble. One would expect that assistance received in such times would be far more computationally meaningful, and cause a far greater change in attitude toward the giver than assistance rendered at other times. Phenomenologically, individuals seem to be deeply moved at such times, find such acts deeply memorable, and often subsequently feel compelled to communicate that they will never forget who helped them.

Given these facts and hypotheses, modern life creates a paradox. For the purposes of friendship assessment, different events and time periods will vary substantially in their informativeness, and certain types of events such as a period of personal trouble will be particularly clarifying. Yet, the human psychological architecture will obviously have been selected to avoid genuine and unnecessary personal difficulties. Safer, more stable modern environments may, therefore, be leaving people in genuine and uncharacteristically protracted doubt as to the nature of their relationships, and whether anyone is deeply engaged in their welfare. Because of the lack of clarifying events, an individual may have many apparently warm social contacts, and yet feel lonely, uneasy, and hungry for the confident sensation of deep social connectedness that people who live in environments that force deep mutual dependence routinely enjoy.

Although there are other kinds of cues, the basic structure of the clarifying event our minds are designed to monitor is one in which a particular individual has the opportunity to help, and that help would be of great value to the recipient. If they fail to help you when such help would be a deliverance, and the cost to them would not have been prohibitive, then it is a mistake to waste one of your scarce friendship niches on them (see below). Their level of commitment is revealed by the magnitude of the cost they are willing to incur per unit of benefit they are willing to deliver. Although there are many other variables that are important—such as how alert they are for opportunities to help, and how effective they can be at helping—the presence of deep engagement is a key variable.

NICHE LIMITATION MODELS OF FRIENDSHIP

Human hunter-gatherers, along with all other prisoners of space and time, have finite time and energy budgets, and cannot be in more than one place at a time. The decision to spend time with some individuals is, therefore, the decision not to spend time with others. Close spatial association is the prime factor that produces opportunities to help and be helped. For a hunter-gatherer, who one chooses to associate with will facilitate or preclude, over time, the development of computational states in others that are beneficial over the long run. From this perspective, each individual can be thought of as having a restricted number of *friendship* or *association niches*, and faces the computational problem of filling these slots with individuals from whom they will reap the best long-term outcomes. If an individual has a limited number of association niches, then the logic of the adaptations underlying friendship may be considerably different than that suggested by the standard model of reciprocation.

What factors would a well-designed computational device take into account in deciding how these niches should be filled?

1 *Number of slots already filled.* Adaptations should be designed to compute how many individuals in one's social world are deeply engaged in one's welfare, and how much uncertainty there is in this computation. If the number is high, then other factors, such as efficiency in exchange relationships or short run return to investment, might be weighted more heavily. If the number is low, or the individual is uncertain about the commitment of her friends, then adaptations should motivate counter-measures: activities that increase the likelihood of friend recruitment or consolidation should become more appealing.

2 *Who emits positive externalities?* The ongoing rewards of interacting with a person can take many forms other than specific acts of altruism.

Behaviours that are not undertaken as intentional acts of altruism often have side-effects that are beneficial to others—what economists call positive externalities. Some potential associates exude more positive externalities than others. For a knowledge-generating and knowledge intensive species such as ours, such situations abound. Someone who is a better wayfinder, game locator, tool-maker, or who speaks neighbouring dialects is a better associate, independent of the intentional altruistic acts she might direct toward you. Similarly, there are an entire array of joint returns that come about through co-ordinated action, such as group hunting or joint problem-solving. Individuals may vary in their value as friends and associates because they contribute to the general success, or because their attributes mesh especially well with yours or with other members of your cooperative unit.

3 *Who is good at reading your mind?* Dyads who are able to communicate well with each other, and who intuitively can understand each other's thoughts and intentions will derive considerably more from co-operative relationships than those who lack such rapport.

4 *Who considers you irreplaceable?* All else equal, it is better to fill a friendship niche with a person who considers you difficult to replace. This person has a bigger stake in your continued health and well-being than an individual who can acquire the kind of benefits you provide elsewhere.

5 *Who wants the same things you want?* A person who values the same things you do will continually be acting to transform the local world into a form that benefits you, as a by-product of their acting to make the world suitable for themselves. Trivial modern cases are easy to see: e.g., a roommate who likes the same music or who doesn't keep setting the thermostat to a temperature you dislike. Ancestrally, associates who shared affinities would have manifested many important mutual positive externalities, such as those who share enemies; those who have the same stake in the status of a coalition; spouses or affines who share a joint stake in the welfare of a set of children, and so on. There are likely to have been recurrent disputes and stable social divisions, and an individual is automatically benefitted by the existence of others who shared the same interest in the outcome. A person who your enemies fear, or a person who attracts more suitors than she can handle, may be a more valuable associate than a reliable reciprocator whose tastes differ widely from your own.

These and many other factors should be processed by the computational machinery that generates what we phenomenally experience as spontaneous liking. Many of them are attributes rather than act-histories, which offers an explanation for why we often experience a spontaneous and deep liking for someone on first exposure.

In other words, not only do individual humans have different reproductive values that can be estimated based on various cues they

manifest, but they also have different association values. One dimension of this value is the partner-independent component, while the other component will vary specifically with respect to the individual attributes of each other potential partner. Adaptations that evolved to regulate association should be designed to fill niches with partners whose association delivers the most net rewards, and who value the individual highly and specifically. The tendency to dispense benefits contingent upon specific reciprocation is not the logic that defines association-value. Although the disposition to make alternating exchanges may not be completely irrelevant to an individual's value as an associate, it is neither a necessary nor a sufficient attribute. It can be trumped by other factors.

Of course, who you can associate with depends not only on who you like, but on who likes you, as well as larger scale structures of friend and family clustering. The computational architecture should be designed to deploy one's choices, acts, and attributes so as to make one's own association value high, and to attract the best distribution of friends into one's limited set of association niches. When this deployment is not effective enough to recruit a worthwhile set of friends, then the architecture should initiate other measures. Increasing the delivery of beneficial acts to others is one possibility, but the analysis above suggests other operations that might be effective: moving into new social worlds, initiating mateships (which have the potential to be a specialized kind of deep engagement association), conceiving children, increasing one's aggressive skills, searching for new positive externalities to exploit, moderating one's negative externalities, ending unfavourable relationships, chasing off association rivals, cultivating irreplaceability, resorting to extortion, and so on—each of which could lead to favourable reconfigurations of one's social world.

The dynamics of this kind of world are considerably different from what the co-operator-defector models, in isolation, suggest. In a world of limited friendship niches, the issue is not necessarily cheating *per se*, but the relative returns of different, mutually exclusive associations. Losing a valued friend, being able to spend less time with the friend, becoming less valued by that friend, or at the extreme, social isolation, may be more costly than being cheated. (This is not to say, however, that one cannot be cheated by a friend.) One way of modelling such a situation is as a Hobbesian bidding war of all against all, waged with the benefits of association, gated by the effectively limited number of friendship niches an individual has. The possibility that a friend will switch between friendships (or rather between mutually exclusive time-association budgets) on the basis of the relative rewards generated by each is the force that keeps the stream of benefits flowing and calibrated. In such a world, the adaptations will be designed to monitor all returns from a relationship, not just those from concrete acts of

material assistance, reciprocally exchanged. It will be advantageous to be a high quality associate, and so individuals should feel a spontaneous pleasure in discovering effective ways of helping their friends, without looking for any contingent return. Instead of being cheated, the primary risk is experiencing a world increasingly devoid of deeply engaged social partners, or sufficiently beneficial social partners, or both. Adaptations should be designed to respond to signs of waning affection by increasing the desire to be liked, and mobilizing changes that will bring it about.

Friendship versus exchange

Accordingly, the phenomenology of friendship unsurprisingly reflects the pleasure you experience in someone's company, the pleasure you feel knowing they enjoy your company, the affection generated by an ease of mutual understanding, the desire to be thoughtful and considerate, the satisfaction in shared interests and tastes, how deeply you were moved by those who helped you when you were in deep trouble, how much pleasure it gave you to be able to help friends when they were in trouble, the trust you have in your friends, and so on. Explicit contingent exchange and turn-taking reciprocation are the forms of altruism that exist when trust is low and friendship is weak or absent, and treating others in such a fashion is commonly interpreted as a communication to that effect. The injection of explicit contingent exchange into existing friendships (e.g., buying a friend's car) is experienced as awkward. It seems to be a pervasive expression of human psychology that people in repeated contact feel the need to rapidly transform relationships that began in commercial transactions into something 'more'—with signs that indicate the relationship is no longer one simply of contingent exchange, but of friendship. Those of us who live in modern market economies engage in explicit contingent exchanges—often with strangers—at an evolutionarily unprecedented rate. We would argue that the widespread alienation many feel with modern commercial society is the result of an evolved psychological architecture that experiences this level of explicit contingent exchange in our lives as a message about how deeply (or rather, how shallowly) we are engaged with others.

Runaway friendship

The issues of irreplaceability and association value have a variety of implications about the functional organization of human social psychology. One of the most interesting implications of this model is how the detection of strong valuation should select for design features that construct a strong reflected valuation: a mirroring effect. By the argument of the Banker's

Paradox, if you are unusually or uniquely valuable to someone else—for whatever reason—then that person has an uncommonly strong interest in your survival during times of difficulty. The interest they have in your survival makes them, therefore, highly valuable to you. The fact that they have a stake in you means (to the extent their support is not redundant to you) that you have a stake in them. Moreover, to the extent they recognize this, the initial stake they have in you may be augmented. Our psychological adaptations should have evolved in response to these dynamics. For example, because you may be the only route through which your maiden aunt can propagate the genes she bears, her psychological architecture may recognize you as being uniquely valuable to her. Because she would sacrifice everything for you (let us assume), that makes her in turn an unusual or perhaps uniquely valuable person in your social universe. Because she values you, you have a corresponding stake in her survival and in the maintenance of her ability to act on your behalf. A risky action to save her life would not be a case of reciprocal altruism, but of altruism through cyclic valuation.

In the same way that the initial impetus in Fisherian runaway sexual selection may have been minor, the initial stake that one person has in the welfare of another might be minor. But the fact that this gives you a stake in them, which gives them a greater stake in you, and so on, can under the right conditions set up a runaway process that produces deep engagements. The recursive nature of these cyclic valuations can reinforce and magnify each person's association value to the other, far beyond the initial valuations. Friendships may become extremely powerful, despite weak initial conditions. Of course, this requires mutual communication and the ability to detect when someone truly values you (in which deception is certainly possible). But against a background of impoverished social options, it might not take much of an initial asymmetric valuation to get such a mirror relationship running and mutually reinforcing. Indeed, under the right conditions, a simple arbitrary decision may be enough (as in oaths of friendship that are found in many cultures), provided it is in the form of an emotional 'commitment' in the sense meant by Hirschleifer (1987) or Frank (1988). When applied to mate choice, these and many of the other arguments made above may help to illuminate the functional design of the adaptations that regulate romantic love (see also Nozick 1989: Ch. 8).

Finally, we want to emphasize that the benefits that certain of the adaptations for altruism described above are designed to deliver are not necessarily benefits at all in the classical sense of increases in direct reproduction or inclusive fitness. The benefits delivered may sometimes have such effects on the recipient's fitness, but this will be as an incidental by-product of the design of the adaptation. It is not the functional product of the adaptation—that is, what the adaptation was designed to do. For

example, in some of these cases, the function of the altruistic act was to extend the recipient's lifespan or otherwise preserve whatever properties make the recipient willing and able to continue supplying benefits to you. If the recipient's fitness increases as a result, this is a side-effect of the computational design and, therefore, irrelevant to the selection pressure that shaped it. Meaningful alternative models of the evolution of altruism might be developed by looking at the delivery of energy, or survival through high-risk episodes, or what might be called agency altruism—increasing the ability of other agents to take effective action. By moving beyond the classical definition of altruism, which requires a fitness cost to the deliverer and a fitness benefit to the recipient, evolutionarily oriented researchers can construct a much richer family of models of altruism which may better account for the diverse array of altruistic adaptations in humans and other species.

Note. We would like to warmly acknowledge Irven DeVore, at whose Simian Seminar the first version of this paper was presented, for his input. He is particularly thanked for the intellectually stimulating effect that his frequent remark 'Tooby, you can be replaced!' had on the central ideas of this paper. We would also like to thank Daphne Bugental, David Buss, Steve Pinker and Don Symons who, as always, provided enlightening insights, and Betsy Jackson, Rob Kurzban, and Melissa Rutherford for their help on manuscript preparation. We are also deeply indebted to the many members of the Center for Evolutionary Psychology's regular seminar, and the seminar on the evolutionary psychology of coalitions, with whom many of these ideas were discussed. The preparation of this paper was supported, in part, by generous grants from the Harry Frank Guggenheim Foundation, the James S. McDonnell Foundation, and the National Science Foundation (#BNS9157–449 to John Tooby).

REFERENCES

Axelrod, R. & Hamilton, W.D. 1981: The evolution of cooperation. *Science* 211, 1390–1396.

Baron-Cohen, S. 1995: *Mindblindness: An essay on autism and theory of mind.* Cambridge, MA: MIT Press.

Baron-Cohen, S., Leslie, A. & Frith, U. 1985: Does the autistic child have a 'theory of mind'? *Cognition* 21, 37–46.

Boorman, S. & Levitt, P. 1980: *The Genetics of Altruism.* NY: Academic Press.

Boyd, R. 1988: Is the repeated prisoner's dilemma a good model of reciprocal altruism? *Ethology and Sociobiology* 9, 211–222.

Boyd, R. & Lorberbaum, J. 1987: No pure strategy is evolutionarily stable in the repeated Prisoner's Dilemma game. *Nature* 327, 58–59.

Cheney, D. & Seyfarth, R. 1990: *How Monkeys See the World.* Chicago: University of Chicago Press.

Cosmides, L. 1985: *Deduction or Darwinian algorithms? An explanation of the 'elusive' content effect on the Wason selection task.* Doctoral dissertation, Department of Psychology, Harvard University: University Microfilms, #86-02206.

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.

Cosmides, L. & Tooby, J. 1981: Cytoplasmic inheritance and intragenomic conflict. *Journal of Theoretical Biology*, 89, 83–129.

Cosmides, L. & Tooby, J. 1989: Evolutionary psychology and the generation of culture, Part II. A computational theory of social exchange. *Ethology and Sociobiology* 10, 51–97.

Cosmides, L. & Tooby, J. 1992: Cognitive adaptations for social exchange. In *The Adapted Mind: Evolutionary psychology and the generation of culture* (ed. J. Barkow, L. Cosmides & J. Tooby), pp. 163–228. NY: Oxford University Press.

Dawkins, R. 1982: *The Extended Phenotype*. NY: Oxford.

Dawkins, R. 1986: *The Blind Watchmaker*. NY: Norton.

Frank, R. 1988: *Passions within Reason: The strategic role of the emotions*. NY: Norton.

Gallistel, C.R. 1990: *The Organizations of Learning*. Cambridge, MA: MIT Press.

Garcia, J. & Koelling, R. 1966: Relations of cue to consequence in avoidance learning. *Psychonomic Science* 4, 123–124.

Gelman, R. 1990: First principles organize attention to and learning about relevant data: Number and the animate-inanimate distinction as examples. *Cognitive Science* 14, 79–106.

Gigerenzer, G., & Hug, K. 1992: Domain-specific reasoning: Social contracts, cheating and perspective change. *Cognition* 43, 127–171.

Haig, D. 1993: Genetic conflicts in human pregnancy. *Quarterly Review of Biology* 68, 495–532.

Hamilton, W. D. 1963: The evolution of altruistic behavior. *American Naturalist* 97, 31–33.

Hamilton, W.D. 1964: The genetical theory of social behavior. *Journal of Theoretical Biology* 7, 1–52.

Hirschleiffer, J. 1987: On emotions as guarantors of threats and promises. In *The Latest on the Best: Essays on evolution and optimality* (ed. J. Dupre). Cambridge, MA: MIT Press.

Humphrey, N. 1984: *Consciousness regained*. Oxford: Oxford University Press.

Leslie. A. 1988: Some implications of pretense for the development of theories of mind. In *Developing Theories of Mind* (ed. J.W. Astington, P.L. Harris, & D.R. Olson), pp. 19–46. NY: Cambridge University Press.

Leslie, A 1994: ToMM, ToBY, and Agency: Core architecture and domain specificity. In *Mapping the Mind: Domain specificity in cognition and culture* (ed. L.A. Hirschfeld & S.A. Gelman), pp. 119–148. NY: Cambridge University Press.

Leslie, A. & Thaiss, L. 1992 Domain specificity in conceptual development: Neuropsychological evidence from autism. *Cognition* 43, 225–251.

Maljkovic, V. 1987: *Reasoning in evolutionarily important domains and schizophrenia: Dissociation between content-dependent and content-independent reasoning*. Undergraduate honors thesis, Dept. of Psychology, Harvard University.

Maynard Smith, J. 1964: Group selection and kin selection. *Nature*, 201, 1145–1147.

Nozick, R. 1989: *The Examined Life*. NY: Simon & Schuster.

Pinker, S. & Bloom, P. 1990: Natural language and natural selection. *Behavioral and Brain Sciences*. 13, 707–727.

Plomin, R. & Daniels, D. 1987: Why are children in the same family so different from one another? *Behavioral and Brain Sciences* 10, 1–16.

Premack, D. & Premack, A. 1994: Origins of human social competence. In *The Cognitive Neurosciences* (ed. M Gazzaniga). Cambridge, MA: MIT Press.

Sugiyama, L., Tooby, J. & Cosmides, L. 1995: Testing for universality: Reasoning adaptations among the Achuar of Amazonia. *Meetings of the Human Behavior and Evolution Society*, Santa Barbara, CA.

Sulloway, F. (Forthcoming) *Born to Rebel*.

Symons, D. 1990: A critique of Darwinian anthropology. *Ethology and Sociobiology* 10, 131–144.

Symons, D. 1992: On the use and misuse of Darwinism in the study of human behavior. In *The Adapted Mind: Evolutionary psychology and the generation of culture* (ed. J. Barkow, L. Cosmides & J . Tooby), pp. 37–159.

Thornhill, R. 1991: The study of adaptation. In *Interpretation and Explanation in the Study of Behavior* (ed. M. Bekoff & D. Jamieson). Boulder, CO: Westview Press.

Tooby, J. & Cosmides, L. 1984: Friendship, reciprocity, and the Banker's Paradox. *Institute for Evolutionary Studies Technical Report 84-1.*

Tooby, J. & Cosmides, L. 1988: Can non-universal mental organs evolve? Constraints from genetics, adaptation, and the evolution of sex. *Institute for Evolutionary Studies Technical Report 88-2.*

Tooby, J. & Cosmides, L. 1989a: Are there different kinds of cooperation and separate Darwinian algorithms for each? *Meetings of the Human Behavior and Evolution Society*, Evanston, IL.

Tooby, J. & Cosmides, L. 1989b: The logic of threat. *Meetings of the Human Behavior and Evolution Society*, Evanston, IL.

Tooby, J. & Cosmides, L. 1990a. The past explains the present: Emotional adaptations and the structure of ancestral environments. *Ethology and Sociobiology* 11, 375–424.

Tooby, J. & Cosmides, L. 1 990b. On the universality of human nature and the uniqueness of the individual: The role of genetics and adaptation. *Journal of Personality* 58, 17–67.

Tooby, J. & Cosmides, L. 1992: The psychological foundations of culture. In *The Adapted Mind: Evolutionary psychology and the generation of culture* (ed. J. Barkow, L. Cosmides & J. Tooby), pp. 19–136. NY: Oxford University Press.

Tooby, J. & Cosmides, L. (forthcoming) Cognitive adaptations for threat.

Trivers, R.L. 1971: The evolution of reciprocal altruism. *Quarterly Review of Biology* 46, 35–57.

Whiten, A. 1991: *Natural Theories of Mind.* Oxford: Blackwell.

Williams, G.C. 1966: *Adaptation and Natural Selection.* Princeton: Princeton University Press.

Williams, G. C. 1985: A defense of reductionism in evolutionary biology. *Oxford Surveys in Evolutionary Biology* 2, 1–27.

Williams, G.C. & Williams, D.C. 1957: Natural selection of individually harmful social adaptations among sibs with special reference to social insects. *Evolution* 17, 249–253.