learn rule-governed behaviour, does reinforcement provide a fully adequate explanation of all that is going on in the process?

If such language is used, why not use a cognitive model that makes such lapses unnecessary? The concept of “self-system” can cover the phenomena with which Rachlin is concerned and much else besides, and recent studies of the “self” (Baumeister 1999) are generating an approach which is as hard-headed as teleological behaviourism. Rachlin uses the “self” only as “existing contingently in a series of temporal intervals during which behavior occurs in patterns” (sect. 4), a model that he suggests would imply that people’s selves evolve and change. But of course they do. It is well established that the “self” (in the sense of a self-description) changes with age and context (Harter 1998; McGuire & McGuire 1988). Rachlin seems unaware of recent developments in this field.

Self-descriptions include references to moral precepts (“I try to be honest”). It is a not unreasonable suggestion that what we describe as “conscience” involves comparison between past, present, or intended action and the moral code incorporated in the self-system (Hinde 2002). This is not incompatible with Rachlin’s description of altruism – “What is highly valued is a temporarily extended pattern of acts into which the particular act fits” (sect. 1.2). Even Rachlin postulates a “coherent sense of self” for the maintenance of altruism. This is entirely in keeping with the results of studies of extraordinary individuals who served as exemplars to many. Such individuals had great certainty about the decisions they made, as though morality were completely integrated into the self-system and altruistic actions involved no conflict with other personal motivations (Colby & Damon 1995; Youniss & Yates 1999). Indeed, studies of personal relationships in the context of exchange theory suggest that being overbenefited as well as being underbenefited can provoke compensatory behaviour (e.g., Prins et al. 1993). If this is confirmed, it indicates that not only a moral code but a social contract is incorporated in the self-system.

How are moral issues incorporated? Young infants show a great deal of proto-prosocial behaviour – sharing, caregiving, showing sympathy, and so on (Rheingold & Hay 1980) – as well as selfish or egocentric behaviour. Rachlin rejects the idea of an inherited mechanism for altruism, but the evidence points to a predisposition to learn to please caregivers. Developmental psychologists have shown how prosocial tendencies are moulded through relationships, especially those within the family, and thus come to form part of self-descriptions (Turie1 1998). Yes, of course reinforcement plays a part, but broad moral precepts, like “you should protect others from danger,” may be incorporated even if the opportunity to act on it has never occurred. There is no denying that reinforcement may play some role in the genesis and maintenance of (apparent) altruism. Minor altruistic acts often receive nods of approval and may contribute to the actor’s status (see discussion of meat-sharing by hunter-gatherers, Hawkes et al. 2001). And reinforcement, if used in a strict Skinnerian sense, is a more hard-headed concept than the “self-system,” which still has fuzzy borders. But reinforcement loses its edge if it is pushed beyond its limits. The use of the “self-system” as something more than an intervening variable but perhaps not quite a hypothetical construct (MacCorquodale & Meehl 1954) can embrace not only the behaviour but also the experience of individuals.

Toward a better understanding of prosocial behavior: The role of evolution and directed attention

Stephen Kaplan & Raymond De Young

Departments of Psychology and of Electrical Engineering and Computer Science, University of Michigan, Ann Arbor, MI 48109; School of Natural Resources and Environment, University of Michigan, Ann Arbor, MI 48109.

skap@umich.edu rdeyoung@umich.edu

http://www.snre.umich.edu/~rdeyoung

Abstract: Rachlin’s thought-provoking analysis could be strengthened by greater openness to evolutionary interpretation and the use of the directed attention concept as a component of self-control. His contribution to the understanding of prosocial behavior would also benefit from abandoning the traditional (and excessively restrictive) definition of altruism.

Discussions of altruism routinely exclude from consideration any behavior from which the actor receives pleasure or other benefit. That Rachlin adopts this traditional approach is understandable but unfortunate. In a perceptive and inadequately appreciated analysis, Wallach and Wallach (1983) point out that there are two distinct meanings of self-interest (or, in their terms, “egoism”). In their example, you can be motivated by helping someone because you expect something in return or because “the other person’s relief from distress or the other person’s happiness is itself what you want to achieve and what would make you happy” (p. 201). As they point out, the two situations are equally self-interested only in the most trivial sense. Yet it is this trivial sense that studies of altruism call upon when they use the traditional definition, that is, that one is acting against one’s self-interest. The result is that the enormously important topic of what motivates prosocial behavior tends to be neglected in favor of a focus on special and atypical cases.

Even accepting this limitation, however, Rachlin’s argument is flawed by his determination to eliminate the potential role of evolution as a component of altruism. This commitment harks back, unfortunately, to an earlier era, when a then-dominant behaviorism argued that the existence of a behaviorist explanation demonstrated that all other explanations were irrelevant. This notion that an explanation at one level usurps the possibility of a useful explanation at another level has been sufficiently pervasive to have received several colorful appellations, such as “nothingbutism” and MacKay’s (1965) more elegant “fallacy of ‘nothing buttery.’” This way of thinking is no more acceptable now than it was then; if indeed there is a demonstrable role of habit in altruism, this in no way eliminates the possibility that there is a role for evolution as well.

A particularly interesting component of Rachlin’s discussion is his use of intrinsic motivation. It is also, however, a topic where a bias against an evolutionary perspective is a serious handicap. His interpretation of intrinsic motivation as arising from a string of habits is less than convincing. The fascination with crossword and jigsaw puzzles seems far more likely to be an expression of the human inclination to solve problems, a tendency humans share with nonhuman primates (Harlow 1953). The very widespread character of this motive strongly suggests its evolutionary origins.

Closely related is Rachlin’s argument that “most of us would indeed choose to be heroes rather than cowards” (sect. 9, last para.). His explanation for the origin of this motivation is not clear. A fairly straightforward explanation arises from Campbell’s (1975) suggestion that humans innately have both self-interested and social motivations, and Goldschmidt’s (1990) impressively documented argument that the inclination to work for the respect of one’s fellows is a central component of human nature. In fact, much prosocial behavior may well be traceable to the way in which cultures use respect as a reward for such behavior. This also provides a nice example of the way in which an innate inclination could provide the leverage for a great deal of learning. Far from being in conflict with an explanation based on learning, the evolutionarily based motivation would be what makes the learning possible.
Rachlin is undoubtedly correct in his assertion that self-control is learned. However, his analysis would be strengthened by including the role of inhibition in managing our behavior. Inhibition is essential to self-control. Without the ability to inhibit the effect of the immediate environment, long-term goals cannot possibly affect present behavior. If people were unable to inhibit any stimuli, essentially forced to attend and respond to every next thing that the environment presented, then contemplation, recollection, and behavioral continuity, necessary for all of Rachlin’s examples, would be unattainable.

Rachlin supposes that self-control is accomplished by an innate learning mechanism. Yet such a mechanism would be unable to inhibit immediate stimuli so as to allow a longer-term pattern to come into play. The mechanism of self-control involves more than just learning habits; there is also the need to direct one’s attention.

Directed attention (Kaplan 1995; Mesulam 1985) is useful in dealing with just the sorts of short-term versus long-term behavioral choices that Rachlin sets up: inhibiting the power of the immediate environment so as to allow consideration of less salient but nonetheless valued patterns. Directed attention allows for a variety of prosocial behaviors (e.g., pursuit of an important social goal despite interesting competition in the immediate setting, helping others despite unmet personal needs, and resisting temptation to maintain devotion to a larger pattern).

A re-analysis of Rachlin’s examples offers some insight on the role of inhibition. The example of a woman entering a burning building is ambiguous because many of the stimuli present (e.g., onlookers screaming that someone is trapped, a child’s scream for help) are both involuntarily fascinating (James 1892) and conceivably capable of prompting a short-term pattern (e.g., entering the building) that is closely linked with the longer-term pattern of prosocial behavior. The behavior of a recovering alcoholic better demonstrates the enormous adaptive advantage offered by inhibition. Here the environmental stimuli conspire mightily against sobriety. Yet, the recovering alcoholic’s self-control is only possible because of the ability to hold the immediate environment at bay and the insertion of cognition between stimulus and response.

The desire not to have to use self-control is a most interesting and useful contribution which fits well with the recent work showing that directed attention is a scarce and labile resource. When under continual demand, our ability to direct our inhibitory process tires, resulting in a condition called directed attention fatigue (DAF). This condition reduces mental effectiveness and makes consideration of abstract long-term goals difficult. A number of symptoms are commonly attributed to this fatigue: irritability and impulsivity that results in regrettable choices, impatience that has us making ill-formed decisions, and distractibility that allows the immediate environment to have a magnified affect on our behavioral choices. The symptoms of DAF can be summarized as a reduced ability to make and follow plans, and the inability to mentally restrain impulsive thought or action. In short, DAF makes prosocial behavior at any temporal scale less likely.

We would thus like to commend Rachlin for his fascinating treatment of the problem of long-term versus short-term interests, for his focus on self-control (and its limitations), and his linking all this to prosocial behavior. At the same time, we would encourage him to consider evolutionary perspectives less extreme than those he has apparently been reading, and to explore the possible role of directed attention as a useful tool in his further exploration of the self-control concept.
learn it is genetic; if you show that a certain behavioral trait is genetically determined I can always show that the expression of that trait depends on the environment. Whether the properties of a ripe red apple in my left hand are genetically or environmentally determined depends on what I am holding in my right hand. If it’s an orange, then the apple’s properties will seem genetic; if it’s another apple in my right hand, say, a green, unripe one from the shady side of the tree, then the apple’s properties will seem environmentally determined. To say that altruism may be learned is not to say that “there is no genetic variance in propensity toward altruism” (Hartung, Zizzo), or “to eliminate the potential role of evolution” (Kaplan & De Young), or to be “averse to considering the genetic basis of behavior” (Zentall), any more than the fact that calculus may be learned means that there is no genetic variance in the ability to learn calculus or that the role of evolution in our ability to learn calculus has been eliminated. The fact that, in free choice tests, a hungry rat spends more time in eating than in wheel-running, and more time in wheel-running than in lever-pressing, has two kinds of implications: one, the behavioral implication that certain contingent relations among these three activities will be found (e.g., wheel-running will reinforce lever-pressing but punish eating); another, that some internal mechanism or mechanisms that are partly innate, partly themselves learned, underlie the behavior. To say that one implication is important does not deny that the other is also important.

Krebs seems to believe that I was trying to say something about how altruism evolved. But I did not even mean to say that self-control came before altruism in the course of human evolution; the reverse might well be true. A behavioristic theory would be silent on how altruism was selected. I did call the self-control mechanism an “innate learning mechanism” but Premack’s (1969) reinforcement theory would not distinguish such a mechanism from an evolved strategy. All I intended to claim with respect to a common mechanism is that in both cases, self-control and altruism, evolution must select behavioral patterns rather than individual acts. Individual acts would be reinforced to the extent that they formed part of a valuable pattern. Whatever mechanism did this would be a “learning mechanism.”

Wilson & Miller accuse me of setting up a straw man in claiming that biologists posit a specialized altruism mechanism. Then they proceed to make virtually the same claim, only now it’s a “specialized form of learning.” I had presumed that when Tooby and Cosmides (1992) compared the supposed specialized mechanism to the eye they were minimizing the contribution of learning. If I was wrong, I apologize. If there were a specialized form of learning, as Wilson & Miller claim, the question becomes, “What does that specialized mechanism do?” If, as I claim, it organizes low-valued particular acts into high-valued patterns, and if we had one such mechanism for self-control and one for altruism, then we would have two mechanisms doing the very same thing. It seems to me that in the absence of physiological evidence for two such redundant mechanisms we ought to assume that only one exists.

R6. Morality

Being from the Bronx, I certainly do not belong among Hartung’s five known pure altruists. That leaves only four remaining. Still, we should try to explain their behavior because, on a lower level, altruism is a pattern in all of our lives. The players of the prisoner’s dilemma game illustrated in Figure 1 who anonymously chose Y are good examples. When I was a Boy Scout I occasionally helped old ladies across the street (I still do, although now I’m less interested). I did not do so in fear of hell or hope of heaven, as Hartung and Levine would seem to require. Despite the rhetoric, fear of hell and hope of heaven are not by themselves good explanations of altruistic behavior. We still would need to explain, in behavioral terms, how such fears and hopes work. As I said in the target article, there is a distinction to be made between altruism and morality. The behavior of the firemen and the hijackers on September 11th may have been equally altruistic but not, from our viewpoint, equally moral. Wagstaff cites cases where altruistic acts turn out to be socially harmful: “The man who impulsively sacrifices himself to help anyone, including gangsters and tyrants, may be acting selflessly, but he is also a social liability.” Correct. This was the point of my example in the target article of the Nazi soldier sacrificing himself for his unit. Wagstaff makes the point, implicit in Lacey’s commentary, that no account of altruism is complete without an understanding of justice. Such an understanding would enable us to distinguish more clearly between altruism and morality. I agree.

References

[Note: The letters “a” and “r” before author’s initials stand for target article and response references, respectively.]


Bandura, A. (1972) Modeling theory: Some traditions, trends, and disputes. In:
References/Rachlin: Altruism and selfishness


(1953) *Science and human behavior.* Macmillan. [RS]


