

Behaviorism and altruistic acts

J. McKenzie Alexander

Department of Philosophy, Logic and Scientific Method, London School of Economics and Political Science, London WC2A 2AE, United Kingdom.
jalex@lse.ac.uk

Abstract: Rachlin's idea that altruism, like self-control, is a valuable, temporally extended pattern of behavior, suggests one way of addressing common problems in developing a rational choice explanation of individual altruistic behavior. However, the form of Rachlin's explicitly behaviorist account of altruistic acts suffers from two faults, one of which questions the feasibility of his particular behaviorist analysis.

Rational choice explanations of altruistic behavior tend to flounder when they try to reconcile individual maximization of expected utility with the fact that altruistic behavior confers a lower utility than other available choices. Incorporating other-regarding interest into the conception of self-interest makes this reconciliation easier to achieve, but at the expense of a notion of self-interest that can reasonably be attributed to most individuals. In this light, Rachlin's explanatory strategy of accounting for altruistic acts by seeing them as particular instances of a highly valued pattern of behavior – the value of which overrides the value of non-altruistic acts – seems to be the right way to proceed. However, as promising as Rachlin's strategy is, I find the extent to which he adopts a behaviorist position troubling. Indeed, his overarching idea that altruistic acts belong to a general pattern of conduct does not commit him to behaviorism, and his final suggestions regarding how patterns of behavior are maintained in fact undermine an explicitly behaviorist account.

To begin, it seems that Rachlin's definition allows for inconsistent classifications of acts. Consider the following example: A father of two children enjoys spending a short time with them at home, but because of their young age prefers being at work all day to being at home all day. Conditions 1 and 2 are thereby satisfied. Then, one day the father contemplates staying at work an extra hour before going home and, reluctantly, decides to stay at work. Is this act altruistic? Because Condition 3 leaves the identity of the group benefited by the father's choice a free variable, whether the act is altruistic or not depends on which group we select. Because the company benefits from the father's choice, by Rachlin's definition the act is altruistic. On the other hand, if we consider the family and see that the father's choice does not benefit it, the act can simultaneously be seen as not altruistic. Note that this is a different point from saying that altruistic acts are context-dependent, a point Rachlin accepts, believing that one can always find contexts that render acts altruistic.¹ While I readily admit that different act-tokens of the same act-type may be differentially classified as altruistic or not depending on the context in which the particular act-token occurs, it does not make sense for the same act-token to be identified as either altruistic or not on the basis of how *we* carve up the world into groups that are or are not affected beneficially by the act.

More importantly, though, I find that in Rachlin's definition of an altruistic act a tension obtains between the behaviorist account of altruism as choices violating individual preference and the underlying behaviorist account of preference presupposed by Conditions 1 and 2. Consider whether a coherent behaviorist gloss may be given to Conditions 1 to 3. Condition 3 poses no difficulty because choice can be defined operationally, but what about the references to individual *preferences* in Conditions 1 and 2? Can these be given a suitable behaviorist interpretation? One behaviorist response might adopt the traditional economist view, which says individual preferences are revealed through choice. However, this explanation of what it means to talk of individual preferences in Conditions 1 and 2 proves difficult to reconcile with Condition 3: if an individual chooses a *t*-length fraction of the longer activity over the brief activity (i.e., Condition 3 obtains), in what sense can one say that the individual *prefers* the brief activity to a *t*-length fraction of the longer activity (i.e., Condition 2 obtains)? The preference is not revealed through the choice of the

individual, because the choice runs directly counter to the supposed preference. Moreover, one may not even be able to say that in *previous* instances, the individual has revealed a preference for the briefer activity, for this instance may be the first time that the individual is presented with the choice opportunity. It seems that behavioral evidence supporting Condition 3 provides evidence *against* Condition 2. How, then, can a behaviorist determine when Conditions 1 to 3 obtain?

It is important to note that these criticisms only target Rachlin's particular definition of altruistic acts and their relation to acts of self-control. They have little impact on his primary observation that *because* altruism "for most of us . . . is not profitable and would not be chosen considering only its case-by-case, extrinsic reinforcement," altruistic behaviour is best explained by appealing to benefits conferred by our choosing to adopt abstract patterns of behaviour. In choosing to follow such patterns, we "forego making decisions on a case-by-case basis" even to the point of being altruistic "at the risk of death" (target article, sect. 10). This seems right, yet need not commit one to a behaviorist position. Moreover, this approach to understanding altruistic behavior raises important questions for future research. How do people acquire preferences regarding these valued abstract patterns of behaviour, and why do they choose to maintain them? Rachlin acknowledges that "extrinsic social reinforcement . . . at home or school or church" may explain the initial acquisition of such pattern. Yet when Rachlin says that "such acts must be maintained not by extrinsic reinforcement but by intrinsic reinforcement" (sect. 9), one wishes for more. The transition from extrinsic to intrinsic reinforcement asks for further explanation, while simultaneously underlining the need to move away from an explicitly behaviorist understanding of altruistic acts.

NOTE

1. A "context" for Rachlin seems to involve only the specification of the longer activity *T*, another free variable in his account: "[c]ondition 1 does not specify the appropriate context (the longer activity, *T*) for a particular act. Is there *any* context (any relatively long-duration activity, *T*) in which a given altruistic act would also be a self-controlled act? I believe that it will always be possible to find such a context." This suggests that the context of an act is solely determined by specifying the long-duration activity. This passage is somewhat confusing because it is not clear how one should understand the expression "altruistic act" appearing within it. I assume that should be read as referring to Rachlin's account. Yet in the sentence immediately preceding the quote, Rachlin asks "Are there altruistic acts under Conditions 2 and 3 above where Condition 1 does not obtain?" According to his definition, this is impossible because "an altruistic act is defined as a choice of the *t*-length fraction of the longer activity over the brief activity under Conditions 1, 2, and 3." Condition 1 must obtain for an altruistic act *by definition*.

Rationality and illusion

Jonathan Baron

Department of Psychology, University of Pennsylvania, Philadelphia, PA 19104-6196. baron@psych.upenn.edu
<http://www.psych.upenn.edu/~baron>

Abstract: Commitment to a pattern of altruism or self-control may indeed be learnable and sometimes rational. Commitment may also result from illusions. In one illusion, people think that their present behavior causes their future behavior, or causes the behavior of others, when really only correlation is present. Another happy illusion is that morality and self-interest coincide, so that altruism appears self-interested.

Consider two patterns for yourself, behaving selfishly and behaving altruistically. Behaving altruistically can be seen as a commitment, what Irwin (1971) called an "extended act." Rachlin may well be right in arguing that the altruistic pattern is better in terms of your long-run self-interest and that you can learn this in a variety of ways.