Evolution and Human Behavior 31 (2010) 131–140

Evolution and Human Behavior

Giving it all away: altruism and answers to the Wason selection task

Laurence Fiddick[a,*], Nicole Erlich[b]

[a]*James Cook University, Townsville, QLD 4811, Australia*
[b]*University of Queensland, Brisbane, QLD, Australia*

Initial receipt 9 April 2008; final revision received 21 August 2009

Abstract

The Wason selection task, a standard test of conditional reasoning, has featured prominently in experimental studies of cognitive adaptations for cooperation. The most prominent of these is Cosmides' investigations of cheater detection on social contract versions of the Wason selection task [*Cognition* 31 (1989) 187–276]. Subsequent to Cosmides' initial investigations, several researchers [*Evol Hum Behav* 21 (200) 25–37; *Manage Decis Econ* 19 (1998) 467–480; *J Genet Psychol* 163 (2002) 425–444; *Evol Hum Behav* 27 (2006) 366–380] have argued that people also are competent at detecting altruism on the Wason selection task, suggesting that there is nothing privileged about the detection of cheaters. However, an analysis of the selection tasks on which these claims are based suggests that participants may have solved these altruism-detection tasks correctly because the scenarios explicitly or implicitly provide the answer to the task in the scenario [*Evol Hum Behav* 21 (200) 25–37; *Manage Decis Econ* 19 (1998) 467–480; *J Genet Psychol* 163 (2002) 425–444], or due to confounds in the cheater-detection tasks leading to the (misleading) appearance of enhanced altruist-detection performance [*Evol Hum Behav* 27 (2006) 366–380]. We tested our conjecture by giving participants selection tasks with and without the answer embedded in the scenario. Performance dropped significantly on the altruism-detection tasks when the embedded answers were removed, whereas performance on cheater-detection versions was unaffected by the manipulation. A reanalysis of the findings of Oda et al. suggested that participants performed significantly worse on their altruism-detection problems than their cheater-detection problems — a finding that we replicate after removing confounds from the cheater-detection tasks of Oda et al. The results reaffirm the specificity of cheater-detection.

Keywords: Altruism; Reciprocity; Cheater-Detection; Social contract theory; Wason selection task

One of the most widely discussed and controversial findings in evolutionary psychology is Cosmides' (1989) demonstration of elevated levels of performance on cheater-detection versions of the Wason selection task in support of the proposal that humans possess a dedicated "look for cheater" mechanism. Several evolutionarily minded theorists have argued that this focus is too narrow and that people ought likewise to be able to detect altruists in reciprocal exchanges and have purportedly demonstrated such an ability (Brown & Moore, 2000; Evans & Chang, 1998; Lawson, 2002; Oda, Hiraishi & Matsumoto-Oda, 2006). We provide arguments and experimental evidence against an altruist-detection mechanism.

* Corresponding author. Department of Psychology, James Cook University, Townsville, QLD 4811, Australia. Tel.: +61 7 4781 4972; fax: +61 7 4781 5117.

E-mail address: larry.fiddick@jcu.edu.au (L. Fiddick).

1090-5138/$ – see front matter © 2010 Elsevier Inc. All rights reserved.
doi:10.1016/j.evolhumbehav.2009.08.003

1. Altruism vs. cooperation

Before proceeding further, it is important to clarify that what previous theorists have proposed is not that one can detect cooperators, but that one can detect altruists. Where a *cooperator* is someone who fulfills his or her obligations in a reciprocal exchange, an *altruist* is someone who acts in a supererogatory manner, exceeding what is required. This can be illustrated by considering behavior with respect to social contract rules of the form, "If you take the benefit, then you must pay the cost." Typically, the *cost* is a benefit granted to some other party. Hence, a cheater is someone who *accepts the benefit* and *does not pay the cost*, a cooperator is someone who *accepts the benefit* and *pays the cost*, and an altruist is someone who *does not accept the benefit* but *pays the cost*. There may, of course, be other ways of defining altruism, but the preceding analysis is the one operationally adopted by all studies of altruist detection on the Wason selection task. We stress that our analysis only applies to altruist detection and should not be taken to minimize the importance of tracking cooperation. Even if there are reasons for doubting the existence of altruism detection, there might all the same be sound reasons for believing that humans track cooperation and assess cooperative intent.

2. Punishments and rewards in the maintenance of cooperation

Studies in experimental economics that contrast how people deal with free-riders and cooperators can potentially shed some light on whether one would expect people to be equally competent at detecting both cheaters and altruists. Two recent studies have specifically contrasted punishments with rewards as a means of establishing and maintaining cooperative interactions (Andreoni, Harbaugh & Vesterlund, 2003; Sefton, Shupp & Walker, 2006). In both studies, the provision of rewards was less effective in maintaining cooperation/fairness than the provision of punishments, although in both studies the combination of rewards and punishments was the most effective in promoting cooperation/fairnesss. For example, in the Sefton et al. (2006) study, contributions to a public good account in their *rewards-only* condition eventually deteriorated to 20% of allocated funds, below the 30% allocation level observed in their baseline condition, whereas contributions remained above 50% in their *punishment-only* condition. Moreover, both studies found that participants rewarded fellow players in a nonintuitive manner, often rewarding players who were less than fair in their behavior and failing to reward more generous behavior with larger rewards. Whereas when punishing, both studies observed that participants rarely punished cooperative/fair behavior and increased their punishments proportional to the deviation from norms of fairness, similar to previous studies (e.g., Fehr & Gächter, 2002). Similarly, in an earlier study by Andreoni, Brown and Vesterlund (2002) in which participants played a sequential two-person public goods game, 30% to 40% of second movers punished free-riding first movers, while nearly no participants rewarded generous behavior, and when they did it was usually only by a negligible amount, whereas punishers typically forfeited a much more substantial amount of money to punish free-riders.

In short, the available evidence from studies that directly contrast the provision of rewards and punishments as a means of establishing and maintaining cooperation suggests that, while people will use rewards, their application of rewards is typically inconsistent and ineffective in promoting cooperation. By contrast, punishments are allocated in a more intuitive fashion — primarily punishing free-riders and in proportion to their deviation from fairness norms — and appear to be more effective in maintaining cooperation. If, as these studies would seem to suggest, people do not have an intuitive grasp of how to effectively allocate rewards to promote cooperation, then it is not clear why they would have evolved with an ability to detect altruists.

The argument is not simply empirical. There are also theoretical grounds for proposing that cheater detection and punishments will be more important in the maintenance of cooperation than altruism detection and rewards. Effective punishment is a negative feedback system (Boyd, Gintis, Bowles & Richerson, 2003). As punishment decreases the levels of free-riding, the cost imposed by the need to punish likewise decreases. Eventually, the threat of punishment alone is enough to maintain cooperation. Effective rewards, however, constitute a positive feedback system (Sutter, Linder & Platsch, 2006). As levels of cooperation increase, the cost paid rewarding cooperation increases. Moreover, the mere promise of rewards, unlike the threat of punishment, is not likely to promote cooperation — cooperation is costly for rewarders, but cheap for punishers.

There is, of course, also the standard evolutionary assumption that true altruists (as opposed to cooperators) should be rare or nonexistent. Even if they are not rare, it is not clear that the failure to detect true altruists would impose a serious fitness cost. Unlike cooperators, true altruists provide aid nondiscriminatively; hence, they inflict no costs of retaliation. Furthermore, it is not clear what advantage there might be in distinguishing between nondiscriminative altruists and discriminative cooperators. If the assumption is that altruists can be exploited with impunity, then the potential costs to be paid for mistaking a cooperator for an altruist can be substantial (Delton, Krasnow, Tooby & Cosmides, 2009). If, however, the argument is that people should be competent at detecting either altruists or cooperators without any distinction made between them, then this is a different proposal that is not supported by the selection task results reviewed below. Hence, both empirical observation and theoretical analysis suggest that cheater detection will be more important to the maintenance of cooperation than will altruism detection.

3. Altruism detection and the Wason selection task

Contrary to the above analysis, there are several studies that purportedly demonstrate that people are also competent at detecting altruists (Brown & Moore, 2000; Chang & Wilson, 2004; Evans & Chang, 1998; Lawson, 2002; Oda et al., 2006). Perhaps the most persuasive demonstration of enhanced altruism detection is that reported by Brown and Moore (2000). Brown and Moore gave participants two altruist-detection tasks to solve. One task involved participants being required to detect a genuinely altruistic babysitter. They were informed that babysitters who abided by the rule, “If they volunteer, then they seek credit,” were considered unacceptable. Thus, if looking for an altruist, they would seek a babysitter who volunteered but did not seek credit. Similarly, the second altruist-detection task involved participants searching for altruistic friends, by observing