ELSEVIER

Evolution and Human Behavior 25 (2004) 209–220

Evolution and Human Behavior

Trustworthiness and competitive altruism can also solve the "tragedy of the commons"☆

Pat Barclay*

Department of Psychology, McMaster University, 1280 Main Street West, Hamilton, Ontario, Canada L8S 4K1

Received 4 November 2003; accepted 19 April 2004

Abstract

The benefits of a good reputation can help explain why some individuals are willing to be altruistic in situations where they will not receive direct benefits. Recent experiments on indirect reciprocity have shown that when people stand to benefit from having a good reputation, they are more altruistic towards groups and charities. However, it is unknown whether indirect reciprocity is the only thing that can cause such an effect. Individuals may be altruistic because it will make them more trustworthy. In this study, I show that participants in a cooperative group game contribute more to their group when they expect to play a dyadic trust game afterwards, and that participants do tend to trust altruistic individuals more than nonaltruistic individuals. I also included a condition where participants had to choose only one person to trust (instead of being able to trust all players) in the dyadic trust game that followed the cooperative group game, and contributions towards the group were maintained best in this condition. This provides some evidence that competition for scarce reputational benefits can help maintain cooperative behaviour because of competitive altruism.

Keywords: Evolution; Competitive altruism; Game theory; Trust; Public goods

1. Introduction

Altruism towards unrelated individuals has puzzled evolutionary biologists for decades, and several theories provide possible explanations for its existence. Theories of direct

☆ This paper is based on the Young Investigator Award winning presentation at HBES2003.

* Tel.: +1-905-525-9140x24867.

E-mail address: barclapj@mcmaster.ca (P. Barclay).

1090-5138/04/$ – see front matter © 2004 Elsevier Inc. All rights reserved.
doi:10.1016/j.evolhumbehav.2004.04.002

reciprocity (Trivers, 1971) and indirect reciprocity (Alexander, 1987) suggest that organisms can succeed by reciprocating altruistic acts towards other altruists. Direct reciprocity occurs when individuals reciprocate generous acts towards others who have been generous to them in the past. Indirect reciprocity occurs when individuals provide benefits for others who have been generous to anyone, and in turn are rewarded for this benevolence by individuals other than the recipients. Many computer simulations and experimental games have shown that some forms of direct and indirect reciprocity can allow for the evolution of altruism, and people actually do engage in direct and indirect reciprocity (see especially Axelrod, 1984; Nowak & Sigmund, 1998; Wedekind & Milinski, 2000; but see also Leimar & Hammerstein, 2001).

However, these theories by themselves cannot account for altruistic acts that cannot be directed towards particular individuals, such as the provision of public goods. A public good is something that people have to incur costs to provide and yet all members of the group benefit from it whether or not they helped provide it (Davis & Holt, 1993), so the public good is open to exploitation by free riders. Examples of public goods include group protection, irrigation, and any collective action project. Individuals have an incentive to not provide public goods because the benefits of providing them are spread among many people, whereas only the altruists bear the cost. Thus, the provision of public goods is very much like the classic "tragedy of the commons" situation introduced by Hardin (1968). One would expect that altruism in such situations would be selected against, and yet many studies demonstrate that humans are willing to contribute to public goods (e.g., Fehr & Gächter, 2000).

People may be altruistic in these situations if there is a chance that they will earn a good reputation that will later be repaid in direct or indirect reciprocity (Alexander, 1987). Supporting this, Milinski, Semmann, and Krambeck (2002) had participants play an experimental game where they alternated between the opportunity to donate money to other players (an indirect reciprocity game from Wedekind & Milinski, 2000) or the opportunity to donate to a public good. They found that people were more likely to contribute to public goods when they expected future indirect reciprocity games, and that participants donated more often in the indirect reciprocity game towards people who contributed to the public good. However, the rewarding of altruists (one component of indirect reciprocity) is not the only way in which an individual might benefit from a reputation for altruism. People often engage in dyadic relations in which they have to trust another person, and competition to form these cooperative partnerships could also account for the importance of reputation. The present study examines whether humans are more willing to trust altruistic individuals than nonaltruistic individuals in a situation where they might be cheated. Alternately, people might not do so, because it would then become possible for individuals to send a dishonest signal by being generous in order to deceive others into trusting them. This would reduce the effectiveness of altruism as a signal of trustworthiness, such that people do not trust altruists any more than nonaltruists.

Given that coalitions and reciprocal altruism are integral parts of human interaction and carry great potential benefits and costs, we can expect careful choice of cooperative

partners (Cosmides & Tooby, 1992). If altruism can signal a willingness to cooperate in partnerships, then altruistic individuals will be desirable partners (Alexander, 1987; Brown & Moore, 2000). However, no one can interact with all people all of the time, and people tend to form friendships or interact more frequently with some individuals than others. This should create a subtle competition to be more altruistic than others in order to be preferred as an exchange partner (Roberts, 1998). Whether this occurs in humans is an open question, but it may be occurring in nonhuman grooming partnerships when good reciprocators prefer to interact with each other (Barrett, Henzi, Weingrill, Lycett, & Hill, 2000). It is similar to Seyfarth's (1977) model of primate grooming where there is competition to associate with the highest-ranking individuals (see Schino, 2001, for a review). More generally, competitive altruism could occur whenever the most altruistic individual in a group can stand to receive more benefits than other altruistic individuals, whether those benefits be better partnerships or not.

No studies have explicitly tested for the existence of competitive altruism in humans, so the present study investigated whether humans will compete to be the most altruistic member of a group. I used a public goods game (PGG) where each participant had an incentive to be selfish, but all participants could have done well if everyone was altruistic. Past studies using PGGs have shown that contributions tend to fall by the last round (Davis & Holt, 1993), and especially in the last round if participants know when the last round is. Therefore, the present study tested whether having an incentive to compete for the most altruistic reputation would maintain contributions better than if there were no extra incentive for being the most altruistic individual.

2. Methods

2.1. Participants

One hundred twenty participants (43 males, 77 females) were recruited through posters around McMaster University campus. The average age of participants was 23.9 (±S.D. 5.6) years.

2.2. General procedure

In groups of four, participants played a PGG followed by a trust game, and each player was given a pseudonym so that one could acquire a reputation in the game but still be anonymous. Participants were seated at a table with dividers that prevented them from seeing each other while they made their decisions, and prevented them from seeing each other's decisions. During the experiment, they earned "lab dollars" that would be exchanged at the end of the experiment to Canadian dollars at a rate of 15:1, with a 1 in 36 chance to change them on par for Canadian dollars (to provide further incentive to treat the "lab dollars" like real dollars). Participants were paid individually at the end of the experiment to reduce the chances of their interacting after the experiment.

2.3. Public goods game

In each of the five PGG rounds, each player was given 10 lab dollars and had the option of contributing any number of these to the public good. Each round, players' contributions were collected using envelopes that had the pseudonyms inside, and these contributions were written on a blackboard beside each participant's pseudonym. The total contributions in each round were multiplied by 1.6, and this new total was divided evenly amongst the participants. After they received their shares, the next round began. At the start of every round, players were told how many rounds of the PGG remained.

2.4. Trust game

After five rounds of the PGG, participants played a version of Berg, Dickhaut, and McCabe's (1995) trust game. They used the same pseudonym for the trust game as they had for the PGG. They were given a total of 30 lab dollars, which they could send to other participants. Any money that a player sent was tripled (by the experimenter) before the other player received it, and the recipient could then return any amount to the sender. Players put the amount they wanted to send into an envelope and wrote down the pseudonym of the player that they wanted to send it to. To maximize information about recipients' decisions, participants indicated how much money they would return for each possible amount they could receive regardless of who sent it: Thus, returns were not contingent on the identity of the sender. This decision was binding in that the experimenter used this information to calculate the amount of money returned to the senders.

2.5. Experimental conditions

There were three experimental conditions with 10 groups in each condition. (1) In the no reputation condition, participants were told they would play another monetary game after the PGG, but they did not know the details of the trust game before playing the PGG, and thus did not have any strategic reason to acquire a good reputation. (2) In the regular reputation condition, participants were informed (before playing the PGG) about the trust game they would play with each of the other three players. (3) In the competitive reputation condition, participants were informed (before playing the PGG) about the trust game they would play with only one partner of their own choosing. Thus, the no reputation condition acts as a control, where participants did not know they could gain a good reputation, whereas the regular reputation and competitive reputation conditions examine the effects of reputation and competition to have the best reputation, respectively.

In the no reputation and regular reputation conditions, participants played the trust game with each of the other three players, to whom they could send up to 10 lab dollars each. In the competitive reputation condition, each participant played one trust game with one other player *of his/her choice*, and could send up to 30 lab dollars to that one player. Thus, participants had the same number of lab dollars to trust to other players in all conditions, but a single "trustworthy" individual could receive more in the competitive reputation

condition. This creates an incentive to be the most altruistic individual in the PGG in order to be the one trusted with money in the trust game.

Although I predicted that groups with high PGG contributions would tend to have higher trust levels in the trust game, I did not predict any additional effect of experimental condition on trust levels. PGG contributions were visible on the blackboard throughout the trust game in all three conditions.

2.6. Practice rounds

Practice rounds were conducted to familiarize the players with the nature and the procedures of the games as well as to make the presence of the trust game salient in the two reputation conditions. In the practice rounds, participants were instructed on how much to contribute in the PGG and entrust to others in the trust game. The amounts were chosen by the experimenter so as not to bias the participants' decisions for or against the experimental hypotheses.

2.7. Statistical analysis

Each group of four players was treated as one unit of analysis because each participant's behaviour affects the behaviour of others in later rounds. Total group contributions in the PGG were analyzed with a repeated measures general linear model. Any violations of sphericity assumptions were corrected for using the conservative Greenhouse–Geisser correction. In PGGs and reciprocity games, contributions often drop in the last round if participants know which round is the last, so I also examined the change in contributions between the fourth and fifth rounds.

In the trust game, the three conditions were analyzed separately. In the no reputation condition, the instructions and practice rounds for the trust game were given in between the PGG and the trust game, causing a long separation between the two and likely reducing any effects of the former on the latter. Both games were played back to back in the regular reputation and competitive reputation conditions, but the number of people to whom money could be entrusted was different in those conditions. In addition to correlating players' contributions in the PGG with amount they received in the trust game, I ranked the players within each four-person session on their total contributions in the PGG, and ran within-group general linear model on the amount sent and received by each player in the trust game.

3. Results

3.1. Effects of reputation in the public goods game

Fig. 1 presents the results of the PGG. There was no overall significant difference between the regular and competitive reputation conditions ($F < 1$), so these conditions were

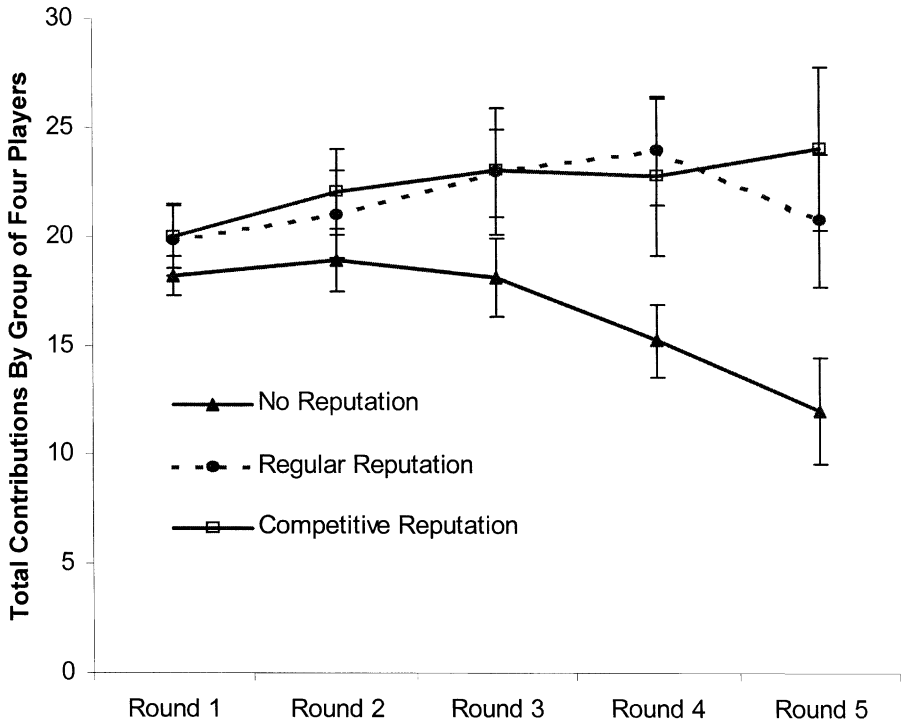

Fig. 1. Group contributions to the public good dropped in the no reputation condition (–▲–), but rose in the regular reputation (–•–) and competitive reputation (–□–) conditions, showing that having an opportunity for reputation makes people more likely to contribute to public goods. Contributions were less likely to drop in the final round of the competitive reputation condition than in the regular reputation condition, suggesting that when individuals have to compete for the most altruistic reputation, they are more likely to continue being altruistic to the end. The error bars represent standard errors of the means.

pooled to compare with the no reputation condition to test for the effects of reputation on PGG contributions. The two pooled conditions with reputation had significantly higher contributions than the no reputation condition [$F(1,28) = 5.91$, $p = .022$]. There was a significant interaction of reputation with round number [$F(1.61,45.06) = 3.69$, $p = .042$]. This interaction was caused by different linear trends in conditions with and without reputation [linear contrast analysis, $F(1,28) = 4.69$, $p = .039$], such that contributions tended to drop without reputation and slightly increase with reputation.

3.2. Competition for reputation in the public goods game

The regular reputation and competitive reputation conditions were very similar until the last round. In the final round, contributions dropped an average of \$3.20 (±S.E.M. 1.40) per group in the regular reputation condition, but rose an average of \$1.28 (±S.E.M. 1.40) per group in the competitive reputation condition (Fig. 1). This difference was significant [$F(1,18) = 5.16$, $p = .036$], indicating that contributions were less likely to drop in the

competitive reputation condition. Furthermore, this difference was significant even if Round 4 is included as a covariate to control for differences in Round 4 contributions [$F(1,17) = 4.85$, $p = .042$].

3.3. Trust game

Total amounts sent in the trust game were not significantly different across the no reputation ($M = \$11.6$), regular reputation ($M = \$11.9$), and competitive reputation ($M = \$10.9$) conditions ($F < 1$). Participants tended to entrust more money to players who contributed more in the PGG in both the regular reputation condition [$r(38) = .72$, $p < .001$] and the competitive reputation condition [$r(38) = .53$, $p < .001$]. In these two conditions, there was a strong correlation between total contributions in the PGG by all four players in each group and the total amount sent in the trust game by all four players [$r(18) = .75$, $p < .001$]. Even after factoring out this effect, players' total contributions still predicted the amount of money they were entrusted with in the trust game in the regular reputation condition [partial $r(37) = .59$, $p < .001$] and in the competitive reputation condition [partial $r(37) = .40$, $p = .012$]. All of the above correlations were positive but not significant in the no reputation condition (all $ps > .10$).

There were significant differences in how much differently ranked players received in both the regular reputation conditions [$F(3,27 = 7.60$, $p = .001$] and the competitive reputation condition [$F(3,27) = 5.79$, $p < .01$; Fig. 2]. In the regular reputation condition, the top-ranking PGG contributor in each group of four received significantly more than the

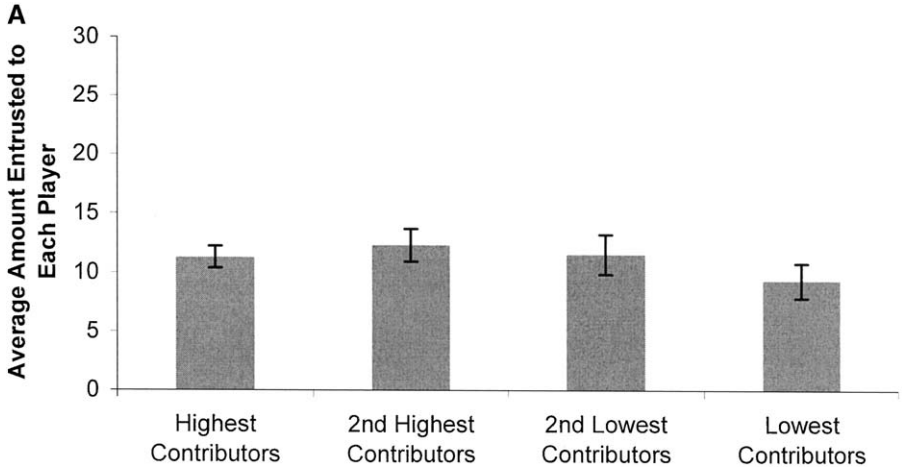

Fig. 2. Average amount (and standard errors of the means) sent to each player in the trust game as a function of their rank as contributors in the PGG. (A) There were no significant differences in the no reputation condition. (B) In the regular reputation condition, the highest-ranking PGG contributors were entrusted with more money than the second lowest or lowest-ranking contributors, and the second highest contributors were entrusted with more than the lowest-ranking contributors. (C) In the competitive reputation condition, the lowest-ranking PGG contributor received less than the other three players did.

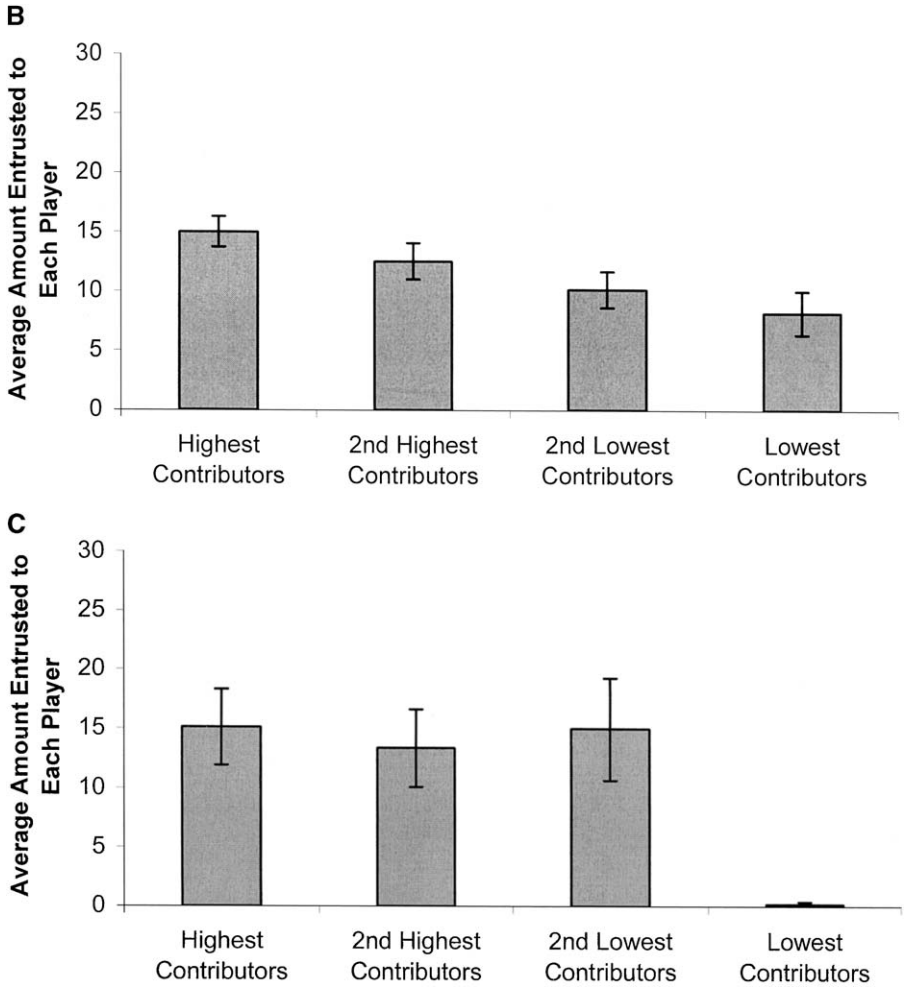

Fig. 2. (*continued*)

bottom-ranking and second lowest contributors [Fs(1,9) = 47.38 and 9.89, $p < .001$ and $p = .012$, respectively], and the second highest contributor received significantly more than the bottom-ranking contributor [$F(1,9) = 5.37$, $p = .046$], but there were no other significant differences. In the competitive reputation condition, the bottom-ranking PGG contributor in each group of four received significantly less than the other three players [all Fs(1,9) > 10, all ps < .01], but there was no effect of the other rank positions on amount entrusted (all Fs < 1). There were no significant differences in the no reputation condition, although the effect approached significance [$F(3,27) = 2.62$, $p = .07$] because the lowest contributor received less on average than any of the others (ps < .10).

These effects were not due to high contributors trusting less (i.e., sending less money) than low contributors in the reputation conditions. After factoring out the total PGG

contributions by all four players in each group (as above), there was no significant correlation between individuals' contributions in the PGG and how much they sent to others in the trust game [partial $r(77) = -.01$, $p > .92$]. Similarly, there were no significant differences among players in a group in how much they sent in the trust game, based on the rank of their total contributions within their group ($F < 1$). The only significant predictor of the proportion of money individuals were willing to return in the trust game was the amount that the individuals themselves sent to other players [$r(78) = .52$, $p < .001$].

4. Discussion

This study shows that people are more willing to contribute to a public good when they can benefit from having a reputation for being altruistic, because contributions were lower in the no reputation condition than in the two reputation conditions (regular reputation and competitive reputation). This study also shows that people may be most altruistic when they could benefit from being the most altruistic individual in a group: PGG contributions were less likely to drop in the last round if there were potential benefits to being the most altruistic. Thirdly, the study found that people were more likely to trust individuals who had been altruistic in the public goods game.

This is the first study to show that people will be most altruistic when they might have to compete to be the most altruistic member of the group, and is thus the first experimental evidence for the existence of competitive altruism in humans. A study by Clark (2002) gave participants the opportunity to give money to the most altruistic group member, but this opportunity did little to increase PGG contributions. Clark's participants had no vested interest in giving money, whereas participants in the present study could benefit by trusting if that trust was repaid. This incentive to trust (and expectation of trust) makes competitive altruism more likely to arise. This could occur in situations where only a subset of all people will benefit from signaling altruism, or when some people will benefit more than others will. Such situations may include (but are not limited to) times when individuals need to cooperate with others, yet cannot or will not form partnerships with all other group members. The most altruistic individuals may attract the best (or the most) cooperative partners (or mating partners if pair bonding is a type of cooperative relationship where free riding on a partner's efforts is possible).

The present study replicates and extends Milinski et al.'s (2002) findings that more people contribute in PGGs when they expect future indirect reciprocity games, and people donate money more often to persons who contributed in the PGG. The present experiment differs from the Milinski et al. study because it shows that reputation effects extend to trust. It does so by pairing a PGG with an experimental game that measures trust rather than indirect reciprocity, and shows another reason why good reputations may be valuable. The present results in the trust game are not likely to be the result of rewarding high PGG contributors for a few reasons. The amounts entrusted to other players were higher than one might expect anyone to send as a reward for contributing in the PGG because the average amount entrusted was over one third of the maximum possible, which was more than the

endowment in a round of the PGG. In other experiments, people have given about one tenth (Clark, 2002) or one fifth (Sefton, Shupp, & Walker, 2002) of the total amount possible to the highest PGG contributors, whereas in the present study the highest contributors were entrusted with about one half of the amount possible in the two reputation conditions (see Fig. 2B and C). Previous public goods experiments have shown that people do not reward low contributors but instead punish them by lowering their payoff (e.g., Fehr & Gächter, 2000, 2002; Sefton et al. 2002). The fact that people gave money to bottom-ranking contributors suggests that the money was not sent as a reward, but was entrusted in hopes of receiving some back. Finally, one would expect that high contributors in the PGG would be more likely to be the people who give to others in indirect reciprocity, as found by Clark (2002). This was not the case in this experiment because there was no relationship between the amounts that individuals contributed and the amount they entrusted to others (after group contributions were factored out).

There was, however, a positive relationship between the PGG contributions of each group and how much its members sent in the trust game, indicating that cooperation may have facilitated partnership formation. Despite this, amounts sent in the trust game were just as high in the no reputation condition as in the two reputation conditions, despite the lower PGG contributions. This may be because there was a reason to discount the honesty of an altruistic signal in the reputation conditions: Some participants could have contributed in the PGG to deceive others into trusting them. This is especially true in the competitive reputation condition, where the incentive to make such deceptive contributions was highest because the potential payoffs of being trusted were highest. Also, the potential cost to the truster was greatest in this condition because one could not spread out the risk by trusting more than one person. However, this discounting of altruism should not eliminate the incentive to try to gain the best reputation, because people might not want to risk being passed over as a partner in favour of more altruistic individuals.

If people are sensitive to the possibility of dishonest signals of altruism, then they should vary their trust according to the costs of being cheated and the potential benefits to a deceptive signaler. Since both of these were greatest in the competitive reputation condition, a signal of altruism would have been least effective in this condition. This might explain the surprising result that participants did not send more money to the highest contributors than to the second and third highest contributors in the competitive reputation condition (Fig. 2C). It also may explain why there was a slightly stronger correlation between a player's contributions and what the player was entrusted with in the regular reputation condition than in the competitive reputation condition. One might predict less wariness outside of the laboratory because pairing with more than one person reduces the cost of being cheated by any one individual. Furthermore, deceptive signaling may be less common outside the laboratory because the cost of altruism could be sufficiently high, making it worthwhile only for individuals who will not defect immediately in a cooperative relationship. This was not the case in the present study because the trust game was played only once. Despite the possibility of dishonest signals in this experiment, participants may have felt that predicting others' trustworthiness from their PGG contributions was better than using no information about other players' behaviour.

There are some limitations of the present design that may have hindered the search for competitive altruism. Anything that reduces the likelihood of money being entrusted to other players also reduces the incentive to compete for the best reputation because the benefits of a reputation depend on being trusted with money. There are a few things that might have reduced the likelihood of money being sent in the trust game. First, there was no guarantee that players would entrust any money to anyone, and each player should have recognized that and realized that one might not receive any money from other players. Secondly, participants had to indicate how much money they would return if they were entrusted with each possible amount, and they could not make these returns contingent on the identity of the sender. This was done to ensure that the effects of having a good reputation did not cause a ceiling effect in contributions by having high contributors benefit from having money preferentially sent and returned to them. However, this design meant that if players wanted to avoid returning money to those who had contributed little in the PGG, they could only do so by returning relatively little to anyone. Players may have feared that the others would return relatively little, and thus not entrusted as much money as they would have if they could respond differently to each person. Thirdly, and most importantly, there may have been little incentive in the competitive reputation condition to be the second or third most desirable trust partner, because participants could only trust one person. Thus, low contributors may have felt that there was no benefit to having a reputation and reduced their contributions in the PGG, affecting the entire group. In real life, the second and third most trustworthy people in a group might also benefit from good reputations because people can form multiple partnerships. Despite these limitations, the present study provides evidence for competitive altruism, which suggests that this is a potentially fruitful area for future research.

Acknowledgments

I thank M. Wilson, M. Daly, A. Muller, L. DeBruine, A. Clark, E. Bressler, S. Myles, D. Krupp, and A. Skye for comments and support; G. Charness, S. Gaulin, and two anonymous reviewers for comments on manuscripts; and A. Chang for logistical support. Financial support for this research was provided by a grant from the Social Science and Humanities Research Council of Canada to M. Wilson.

References

Alexander, R. D. (1987). *The biology of moral systems*. New York: Aldine de Gruyter.

Axelrod, R. (1984). *The evolution of cooperation*. New York: Basic Books.

Barrett, L., Henzi, S. P., Weingrill, T., Lycett, J. E., & Hill, R. A. (2000). Female baboons do not raise the stakes but they give as good as they get. *Animal Behaviour*, *59*, 763–770.

Berg, J., Dickhaut, J., & McCabe, K. (1995). Trust, reciprocity and social history. *Games and Economic Behavior*, *10*, 122–142.

Brown, W. M., & Moore, C. (2000). Is prospective altruist-detection an evolved solution to the adaptive problem

of subtle cheating in cooperative ventures? Supportive evidence using the Wason selection task. *Evolution and Human Behavior*, *21*, 25–37.

Clark, J. (2002). Recognizing large donations to public goods: an experimental test. *Managerial and Decision Economics*, *23*, 33–44.

Cosmides, L., & Tooby, J. (1992). Cognitive adaptations for social exchange. In: J. Barkow, L. Cosmides, & J. Tooby (Eds.), *The adapted mind: evolutionary psychology and the generation of culture* (pp. 163–228). New York: Oxford University Press.

Davis, D. D., & Holt, C. A. (1993). *Experimental economics*. Princeton: Princeton University Press.

Fehr, E., & Gächter, S. (2000). Cooperation and punishment in public goods experiments. *American Economic Review*, *90*, 980–994.

Fehr, E., & Gächter, S. (2002). Altruistic punishment in humans. *Nature*, *415*, 137–140.

Hardin, G. (1968). The tragedy of the commons. *Science*, *162*, 1243–1248.

Leimar, O., & Hammerstein, P. (2001). Evolution of cooperation through indirect reciprocity. *Proceedings of the Royal Society of London B*, *268*, 745–763.

Milinski, M., Semmann, D., & Krambeck, H.-J. (2002). Reputation helps solve the "tragedy of the commons". *Nature*, *415*, 424–426.

Nowak, M. A., & Sigmund, K. (1998). Evolution of indirect reciprocity by image scoring. *Nature*, *393*, 573–577.

Roberts, G. (1998). Competitive altruism: from reciprocity to the handicap principle. *Proceedings of the Royal Society of London B*, *265*, 427–431.

Schino, G. (2001). Grooming, competition and social rank among female primates: a meta-analysis. *Animal Behaviour*, *62*, 265–271.

Sefton, M., Shupp, R., & Walker, J. (2002). The effects of rewards and sanctions in provision of public goods. *Working Paper W00-16*, Centre for Decision Research and Experimental Economics, Nottingham, UK.

Seyfarth, R. M. (1977). A model of social grooming among adult female monkeys. *Journal of Theoretical Biology*, *65*, 671–698.

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

Wedekind, C., & Milinski, M. (2000). Cooperation through image scoring in humans. *Science*, *288*, 850–852.