RESEARCH REPORT

GIVE AND YOU SHALL RECEIVE. GIVE MORE AND
YOU SHALL BE HONORED. EXPERIMENTAL EVIDENCE
FOR ALTRUISM AS COSTLY SIGNALING

Siegfried Dewitte • David De Cremer

OR 0415

Give and you shall receive.

Give more and you shall be honored.

Experimental evidence for altruism as costly signaling

Siegfried Dewitte

Catholic University Leuven

David De Cremer

Tilburg University

Abstract

In two studies, we measured the degree and manner of reciprocation of a public good in subsequent two-by-two interactions. Both studies consisted of two phases: a public good phase and a subsequent give some game (i.e. a gradual Prisoner's dilemma game). In the first study, the public good was a financial game in the lab. In the second study, the public good games were real life student projects. The observed behavior in the subsequent interactions was consistent with the reciprocity rule, but only up to the fairness norm of equality. Students who had invested more than their fair share, did not receive more financial returns than those who had invested a fair share. However, despite the lack of financial benefits, these high investors were preferred more as future team mates (Study 1), or received more social rewards (Study 2). We interpret these findings in terms of altruism as costly signaling.

Keywords: Altruism, costly signaling, cooperation, reputation

1. Introduction

Altruistic behavior has been puzzling evolutionary and social scientists for decades. The main challenge is the reconciliation of the selfish processes that govern evolution with the phenotypic unselfish, costly behavior (Gurven, in press). Several potential mechanisms that bridge that gap have been suggested. One such mechanism is costly signaling. In this view, the cost of altruism allows it to be an honest sign for some unobservable underlying quality or good intent. Altruism as costly signaling is characterized by the unconditional donation of resources to the community (Zahavi & Zahavi, 1997). Gintis, Smith, and Bowles (2001) showed that such a type of altruism leads to an equilibrium under plausible circumstances. However, empirical evidence is scarce. Bliege-Bird & Smith (2001) and Gurven, Allen-Arave, Hill and Hurtado (2000) provided anthropological evidence that costly signaling might be behind some altruistic behavior that is observed in non-market societies. In this paper, we report the first experimental evidence supporting costly signaling as an explanation for altruistic behavior.

Altruism is generally defined as providing a benefit to another individual at a cost to the agent. The mechanisms proposed to explain altruistic behavior to date fall into two major categories. The first category refers to genuine altruism at the individual level. The altruist bears the cost, but does not gain any benefit. This type of altruism makes evolutionary sense if one acknowledges that selection also affects other levels than the individual. We distinguish selection forces at a smaller and a broader scale than the individual. Kin selection (Hamilton, 1964) reflects selection at the gene level. At the individual level, altruism towards strongly related individuals is genuine, because the donor incurs real costs. However, at the gene level, genes facilitating altruistic behavior towards kin actually benefit copies of themselves. Specifically, kin are likely to carry the same gene facilitating altruism. This way,

phenotypical altruism based on kin selection relies on gene competition. There is ample evidence throughout the animal kingdom that kin selection shaped a portion of altruistic behavior (e.g. Chapuisat & Keller, 1999) and also among humans, it accounts for some portion in the variance in sharing (Gurven, in press).

Selection also affects levels higher than the individual. Clade or group selection (e.g. Bergström, 2002, Williams, 1966) reflects selection on the level of the group or species. Unconditional altruism towards unrelated people can survive when it is temporarily protected or isolated from selfish invaders (e.g. the haystack game, Bergström, 2002). Groups consisting of altruists thrive (as long as they are not invaded) and therefore outnumber egoists in the long run. In this case, phenotypical altruism relies on group competition. Empirical evidence that this mechanism accounts for any variance in human sharing is hard to find. However, all else equal and irrespective of the mechanisms explaining that selfishness did not invade the altruism, it is difficult to deny that in the long run, altruistic groups fare better than selfish groups (Gintis et al., 2001).

The second type of mechanisms invoked to explain altruistic behaviors relies on the benefits for the altruistic individual. Although in the short run, the altruist incurs a cost, he is better off in the long run because he gets something in return. One type of such altruism is cooperation or reciprocal altruism (Trivers, 1971; Axelrod, 1984). The long term material benefit (repayment of group benefits) is the core mechanism sustaining this type of altruism. There is ample evidence for cooperation based on reciprocity, not only in humans (Axelrod, 1984, Gurven, in press, Wedekind & Milinski, 2000) but also in other animals (e.g. Wilkinson, 1984) and between species (e.g. Maynard-Smith & Axelrod, 1981). In humans, reciprocity can even be vicarious, in the sense that trust that an agent will reciprocate can emerge from observation rather than interaction (Wedekind & Milinski, 2000). Another type of benefit that altruists can receive is social status in the community, with all its entailing

fitness benefits (Richards, 1974, Wilson, 1975). Individuals with a high internal quality can signal this quality by displaying unconditional altruism. Through this behavior, the community members learn that this individual is worth having in a coalition (e.g. for mating) or better is avoided in a competition. As a result, this person becomes more wanted, and hence increases in status. Status increases mating success and access to resources. So, the altruistic individual is reimbursed with increased fitness during his or her lifetime.

At first sight, cooperation and altruism as costly signaling might appear similar because in both cases altruism is under the selection control of benefits for the individual. The major aim of the present paper is to distinguish both mechanisms conceptually and empirically. Two studies are presented to show that reciprocity cannot explain all aspects of altruism in our experimental session.

Altruism as costly signaling is not cooperation

Cooperation has several essential characteristics that follow from the selection process governing it. We will focus on these mechanisms that set it apart from altruism as costly signaling (see also Gurven, in press). The first essential characteristic of cooperation is (social) *conditionality*. Cooperation involves an interaction between two or more individuals. The level of cooperation strongly depends on the opponent's previous behavior (Axelrod, 1984, Nowak & Sigmund, 1998; Sethi & Somanathan, 2003; Wedekind & Milinski, 2000). To give just one example, van Lange, Ouwerkerk, and Tazelaar (2002) showed that subtle communication distortions in a give some game (= PDG with a more fine-grained option scale) can lead to serious drops in cooperativeness. However, when the opponent always returns the player's investment plus one (i.e. Tit for Tat + 1), the negative effects of noise are avoided. Further, Fehr and Gächter (2002) showed that punitive feelings and behavior toward a player heavily depended on the player's cooperative behavior in a public good game. These

and many other findings in the literature strongly attest to the importance of the opponent's behavior in the decision to act cooperatively and hence to conditionality.

Conditionality does not play a role in altruism as costly signaling. The benefits that are provided by the altruist are not directed at a specific person, so they cannot be conditional. If they are, they do not depend on the receivers' history (Gintis, Smith, & Bowles, 2001) but possibly on receiver's need for food (Gurven, in press). In fact, the situations in which this type of altruism is displayed seem to be designed in such a way that conditionality is not easily applied. For instance, when a team of hunters bring home sea turtles, they are donated to the host of a large feast and the hunter does not have any control over who enjoys the meat (Bliege-Bird & Smith, 2001). Similarly, the chiefs throwing community parties or potlaches cannot easily exclude people from joining the feast (Boone, 1998).Significantly, unconditionality is not identical to what Gurven (in press) called producer control. Producer control refers to the control a producer has on the sharing of the product. However, unconditionality implies control combined with the deliberative decision not to differentiate among receivers. The first major difference between cooperation and altruism as costly signaling, hence, is conditionality.

A second essential element in cooperation is trust. Trust is essential for cooperation to occur. In fact, people with dispositionally low levels of trust, do not cooperate as much, unless they are pushed, for instance by salient reputation concerns (De Cremer, Snyder, & Dewitte, 2001). In fact, the communication and commitment effect on cooperation (e.g. Lindskold, Han, & Bettis, 1986; Orbell, Vandekragt, & Dawes, 1988) all rely on the increase of trust in the opponent, with the result that the players reduce the risk of being exploited. Van Lange's (2002) procedure involving an opponent playing TFT+1 can be interpreted as a trust building mechanism. Further, the shadow in the future (reputation concern) increases the

expectation that the opponent will cooperate (Axelrod, 1984; Murnighan & Roth, 1983) and hence reflects the same essence of cooperation.

In sharp contrast, for altruism as costly signaling, trust is irrelevant. The first reason is that no return is expected from the receivers. The second reason is that running the cost (which implies no reimbursement) is an *essential* part of the signal. In other words, the expectation that the receivers would pay back the altruist would ruin the function of altruism as costly signaling.

To summarize, the essential difference between cooperation and altruism as costly signaling seems to be that cooperation requires a reasonable expectation that the other will return the favor, whereas such an expectation is at odds with the nature of costly signaling. In fact, this expectation implies that the cooperator's behavior is regulated by a social contract or a norm. In sharp contrast, the altruist's behavior is designed to signal more than normal quality, and therefore, he or she has to surpass the norm in some way or another. To conclude, irrespective of the sharing norm amongst a certain group of people (Gurven, in press; Gintis, Bowles, Boyd, & Fehr, 2003), cooperation reflects norm conforming behavior whereas altruism reflects norm surpassing behavior.

A delicate issue is whether or not altruists and cooperators are conscious of the function of their altruism, and more importantly, whether or not the awareness is different for cooperation and altruism. That is, altruists might be unaware of the ultimate rewards that motivate their efforts, or they might be very skilled in avoiding the impression that they consciously aim at harvesting material or even social rewards in the future. To avoid that this complicated issue renders our operationalization equivocal, for the time being we focused on social consequences of naturally occurring cooperation and altruism. We leave the awareness question for future research. Specifically, we expected that behavior that conforms to the norm (=cooperation) would be reciprocated materially but that behavior that surpasses the

norm (=altruism) would not be reciprocated with material means. Rather, norm surpassing behavior should be rewarded with social rewards or 'honor'.

In the present paper, we present two studies that test this prediction. In the first study, we measured the impact of altruism, cooperation, and defection that naturally occurred in the lab on anonymous monetary and non-monetary rewards. In the second study, we assessed the impact that naturally occurring altruism, cooperation, or free-riding in student teams had on anonymous monetary and non-monetary rewards measured. Twice it was found that doing more than the norm is not returned monetarily: opponents return the normative amount of money. However, altruists receive more social rewards.

Study 1. Give and you shall be honored: the laboratory version

We created an opportunity for participants to display altruism in a social dilemma in the lab. They played a public good game (without provision point) and had to announce how much of their endowment they would invest in the public good. Afterwards, they played anonymous give some games (e.g. Van Lange et al., 2002) against each other and did so two-by-two. We further measured with whom they would like to team up in the future (e.g. to make group assignments). As a control variable, we also measured the level of friendship between all participants. Because the game was public, knowing each other might have a substantial effect on sharing decisions. Finally, in many game experiments, the endowments are obtained without effort. Possibly, this situation is not appropriate to display altruism because the cost of sharing is too low (the windfall effect, e.g. Kameda, Takezawa, Tindale, & Smith, 2002). On the other hand, it might also enhance altruistic behavior exactly because of the low cost. To avoid that this artificial feature would hide or boost altruistic behavior, we manipulated how much effort would be required to obtain the individual endowment. In one condition, participants had to work for their endowment. In the other (yoked) condition, they

received the endowment for free. If sharing depends on effort (e.g. Miller, 1977), the predicted effect should be obtained only in the work condition. If sharing depends on possession in itself, then the effect should generalize irrespective of effort invested.

Method

Participants

Students were invited to the lab in groups of 4-6. There were 4 groups of 4, 2 groups of 5, and 8 groups of 6, totaling 74 participants. Half of the groups were assigned to the "work" public good game condition and the other half to the "free" public good game condition (yoked to the first condition). Forty participants were women and 34 were men. The ages ranged between 18 and 25, with an average of 21 years. Both factors did not have effects and are not discussed further. Participants were recruited from a diversity of study orientations, with a majority from the social sciences. They participated in return for a fee that varied with their performance during the games (from €7.90 to €9.30, €1 ≈ $1).

Procedure

Once all the participants had arrived, they were seated around a group of 6 tables. The tables were arranged in a square. Before them there was a number from 1 to 4, 5, or 6 (depending on group size), clearly written on a cardboard. The experimenter introduced himself and invited them to start reading instructions.

In the '*work*' public good game, they were explained that they would have to generate as many examples of a concept as possible during two minutes. We further explained that they would be asked to contribute to the group afterwards (but did not mention that this would happen publicly). They were told that every example of the concept they would contribute would be doubled and be redistributed among the group members. A numeric example showing the behavior and outcomes for a group of three players was provided, one of which had contributed a lot, another a little, and another one had made an average contribution.

They were also told that the number of concepts they would end up with were worth one cent each.

In the actual generation phase, they had to generate car brands. When the two minutes had passed, the experimenter asked the players to count the car brands they had generated (ranging from 9 to 30). They were asked to decide how many they wished to contribute to the group (for simplicity and comparability, we ignored doubles). They then had to read in public the number of car brands they had generated, and the number of car brands they contributed to the group.

In the '*free*' public good game, the participants received a similar introduction, but the example of generating concepts was omitted. Instead of having to generate their possessions, they were assigned a certain number of eurocents (yoked to the other condition, see above). They then had to decide how much of their endowment they wished to contribute to the group. They had to read aloud the initial amount they had been assigned and the amount they contributed, in exactly the same way as in the 'work' public good game.

From hereon, the procedure was identical in both conditions. They had to play a give some game with each of the other group members separately. For each other group member, they received €0.20, and they had to decide how many cents they would keep to themselves, and how many cents they would give to the other. The amount given was doubled. Appendix A shows the pay-off matrix for this game.

Finally, they received a questionnaire checking whether they had understood the procedure. All participants understood that everyone would be better off if everybody contributed rather than kept all their money. We also asked with whom they would like to team up in the future (on a seven-point likert scale) and to what extent they knew the others beforehand. They had to check the highest category that applied for each opponent on a 6-point scale: 1: 'never seen before'; 2: 'know his or her face', 3: 'know his or her name', 4:

'spoke to him or her before', 5: 'I already worked with him or her', 6: 'is a good friend'. This measure was found to be nearly dichotomous: friends (6) vs. strangers (0) and was treated as such in the remainder. The scores below 4 were merged (never interacted) as well as those above 3 (interacted).

Results and discussion

The unit of analysis was the pair wise decision from the point of view of the receiver (n= 328). The two dependent variables were the amount of money (0-20) and the score on the popularity rating (1-7) each player received from each other player in their group. We call these respectively the monetary return and the social return. Both measures were standardized to allow comparisons between the measures. The independent variables were effort (two levels: the work condition vs. the free condition), friendship (two levels, high, n = 65 or low, n = 263, measured), and relative investment, which reflects the proportion of the endowment the player had invested in the public good in phase 1. The relative investment was also standardized to allow interpretation of the estimated parameter. All interaction terms were also included. To test the core hypothesis, we also included the squared relative investment, and the interaction with the two other factors (effort and friendship). In this type of experiment with public interaction, social behavior could also be affected by the group atmosphere. To control for this, we added the group as an independent variable (nested in condition, Kirk, 1995). Further, because each the players might also differ with respect to how much they give in general, we included giver as an additional independent factor (nested in group). These additional controls produce more reliable tests of the target effects (Kirk, 1995).

We will first report the results pertaining to the predicted effect, and then proceed with the moderating results. All omitted ps are smaller than .01 or beyond. Relative investment in phase 1 was positively related to return in general: $F(1,245) = 9.64$ and type of return did not

interact with relative investment, $F(1,245) < 1.0$, ns. This replicates the reciprocity effect that is typical for cooperative behavior. More interestingly, we found an interaction between the quadratic relative investment and type of return: $F(1,245) = 5.61$, $p < .02$. Specifically, for the monetary return, the quadratic relative investment was significant: $F(1,245) = 8.94$. For the social return, it was not, $F < 1.0$, ns. Figure 1 shows the best fitting curve for the monetary return. It shows that the relation between the relative investment in phase 1 and monetary return increases upwards to around a relative investment of 50%, and then levels off. This level is close to 10 (= 50% of the endowment in the give some game in phase 2). Assuming that the social norm is giving half of one's endowment in these types of games (see Gintis et al., 2003), this result suggests that reciprocity rules shape behavior up to the norm and that return ceases to follow relative investment beyond the norm. Interestingly, the work manipulation did not interact with neither the linear nor the quadratic effects mentioned above (both Fs < 1.0, ns).

Figure 1. Received amount of money in the give-some game of phase 2 as a function of the relative investment in the public good in phase 1
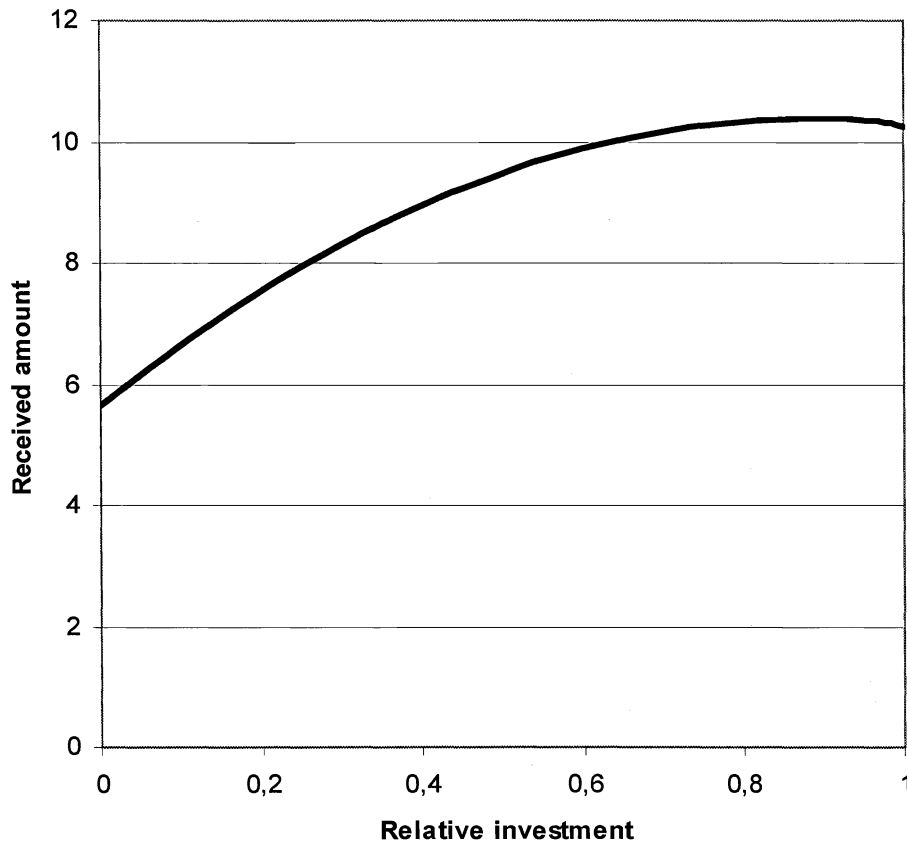
There are some additional effects worth mentioning. Obviously, friendship had an important

main effect, $F(1,245) = 75.19$, but it was stronger for the measure 'preference as a partner'

($F(1,245) = 98.9$) than for received money ($F(1,245) = 10.1$; interaction: $F(1,245) = 67.51$).

Furthermore, friends received only slightly more in the effort condition than in the free

condition ($F(1,245) = 3.80, p = .08$) but this was due to the effect of effort condition on

relative investment (see below). Finally, the effect of relative investment was somewhat

smaller for friends than for strangers ($F(1,245) = 2.93, p = .09$), but this two-way interaction

was qualified by a three-way interaction with type of measure: $F(1,245) = 3.39, p < .07$.

Specifically, the buffering effect of friendship was evident only for the social measure

($F(1,245) = 4.28, p < .04$) and not for the monetary measure ($F(1,245) < 1.0$). Obviously, one

public good decision, no matter how asocial, should not ruin the friend's mutual preference to collaborate. These plausible findings add to the validity of our procedure. Finally, the effects of both the group factor ($F(12,60) = 4.72$, and donor ($F(60,249) = 8.58$) were substantial.

We conducted two additional analyses. First we evaluated contribution in phase 1 as a function of effort condition, group (nested in condition), and possession. We found that having to work for the endowment reduces the relative investment from $M = 76\%$ to $M = 54\%$: $F(1,12) = 7.98$, $p < 0.02$. Combined with the lack of interaction between the quadratic effect of relative investment and effort condition in the previous analysis, this means that effort does not change the reaction of the observers, but rather the willingness of the giver to display altruism. Interestingly, high possession reduced the relative investment slightly ($F(1,327) = 4.91$, $p < .03$).

Finally, it is interesting to explore to what extent relative investment in phase 1 actually predicted the same person's cooperative behavior in phase 2. Across all participants, we found a strong correlation between average money contributed to the public good in phase 1 and average amount given to the others in the give some game (phase 2, averaged across opponents): $r = 0.56$, $p < .0001$, $n = 74$. This correlation was not affected by effort condition ($F(1,70) < 1$). This suggests that behavior in phase 1 is actually predictive for behavior in phase 2, irrespective of means, and hence that reputation information is indeed a good guide to predict whether or not a player will cooperate.

Four issues need to be considered. First, one could object to our use in phase 1 of relative investment rather than absolute investment as a marker of altruistic behavior. However, altruism is defined in terms of benefits to others *relative* to costs to the giver. Further, in this study, possession was not related to the trait of interest at all. Therefore, in this study, relative investment is a better operationalization of altruism than absolute investment.

The second issue that we need to consider is the fact that people may stop returning investments beyond a certain level because those who invest more than the norm are 'suckers'. The fact that the altruists are popular to team up with can be consistent with the sucker interpretation if one considers that it might be easier to free ride in a team with a *sucker*. To rule out this interpretation, we included a stronger social return measure in the next study that is not consistent with the *sucker* interpretation.

A third issue is the vagueness of the norm in phase 1 of this study. Because people did not agree on the norm, it is difficult to be sure that they surpassed the norm. To reduce these concerns, the second study made use of a naturalistically evolving reputation and a subjective evaluation of the norm in this situation.

A final issue is related to a possible methodological artifact for our main result. That is, the lack of quadratic relation between relative investment and popularity score could be related to the type of scale. Specifically, the higher scale points might be relatively closer to each other than the lower scale points, which might hide an actual quadratic trend. In this case the whole quadratic trend (both in the material and social measures) might be due to logarithmically decreasing perceived value of higher investments. To rule out this interpretation in Study 2, we measured the social and monetary return measures with a highly similar scale.

Study 2: Give and you shall be honored: the real life version

The major aim of Study 2 was to replicate the major result of Study 1, namely the quadratic relation between relative investment in a public good and monetary return and the lack of this quadratic relation for social return. We adapted the procedure to deal with the potential interpretational problems. First, to increase the ecological validity of the first phase,

we used student groups with a substantial history of productive interactions. The first phase was not played in the lab. Rather, it was a history of interactions in changing study teams during a semester (13 weeks). The relative investment was measured by a simple question with five options centered on the option 'student x did his or her fair share'. In this way, the norm became relative and hence comparable over individuals.

To solve the *sucker* interpretation, the social return was more valuable than a popularity rating as a team mate. Specifically, the recipient received something that had social value. We had two versions of social return. The first version included points in an emotional intelligence contest. If you think of a person as a *sucker*, you certainly would not give that person a high score in this contest. In the second version points earned the right to construct teams in the following semester. If you find someone to be a *sucker* and you are a *free rider* yourself, you certainly would not give the *sucker* the right to build a team.

To solve the methodological artifact issue, all return types (i.e. social or monetary) in phase two were the stakes in a similar give-some game. The only difference was that the points had different exchange values in the different games. Therefore, any difference we observed in the relation between altruism and type of returns can be related only to the nature of the exchange, not to the difference in measuring method.

Method

Participants

Participants were a group of 42 students enrolled in the "Master of Science in Marketing" program (Catholic University of Leuven) that had collaborated in several teams during one semester. They participated after having finished an exam. Twenty-seven of them volunteered to participate (13 men). They could earn money in return for their efforts. Their age ranged between 22 and 26.

Procedure

After the students had finished an exam that took less time than anticipated, they read an introductory text inviting them to participate in a research. It was stressed that they were completely free to withdraw (13 actually did). The students played give-some games against each student of the entire group. The stakes in give-some games were points (20, for the pay-off matrix, see appendix A) that had three different meanings. In one third of the games, each point was worth €0.01. In the second third of the games, the points earned the right to build student teams for the second semester. The student with the highest score in this ranking could choose with which students he or she would team up with. The second highest scorer's preferences (if not already chosen) were used to build the next team, until all teams were built. To that end, they also had to give a top five of students they would like to share their team with. In the last third of the games, the points were part of a contest. The student with the highest score in this contest would be publicly announced as the most emotionally intelligent student of the group.

A table containing the names of all students was provided on the first page following the instruction sheet. They had to start with barring their own name. Next to each name, there were two columns. One was headed with 'you give' and one with 'you keep'. The sum of the two numbers had to equal 20 for each row. The order of the games was systematically varied, whereas the order of the students was alphabetical. On the second sheet, the same list of students was shown. All participants had to rate each student on two dimensions. They had to rate fairness of contribution and friendship with each student. To measure fairness of contribution (i.e. defection, cooperation, or altruism), we asked "You all collaborated with each other in teams this year. To what extent did class mate X cooperate in these team tasks? (You can use both your personal experience if you have any, and the information from others.) They had to rate the investment on a five-point scale with the following anchors: 'very poorly', 'less than fair share', 'OK: fair share', 'more than fair share', and 'outstanding

effort and input'. Secondly, we adapted the measure of friendship to the current situation because all people had at least seen each other before: 'To what extent would you consider X as a friend? Give your impression on a seven-point scale. The anchors were 'I avoid X', 'No friend at all', 'no friend', 'we get along, that's it', 'get along rather well', 'is my friend', 'is my best friend'.

Finally, we also recorded the score on the exam preceding the experiment (n = 42), the grades on an individual assignment, and two grades on team assignments. We did so to gauge whether intelligence might increase the likelihood that people signal by means of altruism.

Results

Not all students knew all of the other students well enough to evaluate each other's contribution. Therefore, of the 1107 possible data points (27 by 41) we could use only 524.

Again, the unit of analysis was the receiver. We conducted an ANOVA with the received points as the dependent measure and with type of game (3 levels: money, E.Q, or team building), fairness of the contribution in the student teams (5 levels, from defection to surpassing the norm), and friendship (2 levels) and all interactions as independent variables. We again included the giver as a control variable. All omitted ps are smaller than .01 or beyond.

Again, the contribution in the first phase exerted a main effect ($F(4,467) = 5.77$) which was almost entirely due to the linear trend. Having done more in the study teams yielded more points in the games: $F(1,467) = 21.08$. The cubic trend was also significant, $F(1,467) = 6.50, p < .02$, but this is due to the interaction with type of game (see below). The other trends (quadratic and quartic) were not significant (all Fs < 1.0). More interestingly, we found an interaction between contribution in the first phase and type of game: $F(8,467) = 2.06, p < .04$. The linear, cubic, and quartic trends were comparable for the three games (Fs<1.38, ns). In

line with the predictions, the quadratic trend interacted (marginally) with the type of game: $F(2,467) = 2.64$, $p = .07$. Specifically, we found a significant quadratic trend in the game played for money: $F(1,467) = 5.51$, $p<.02$. For the E.Q game ($F= 0.0$) and the team construction game ($F=0.65$), there were no indications of a quadratic trend. Figure 2 shows the interaction between the trends and the type of game.

Figure 2. Points received in the give some game (0-20) as function of type of game and investments in the study teams.
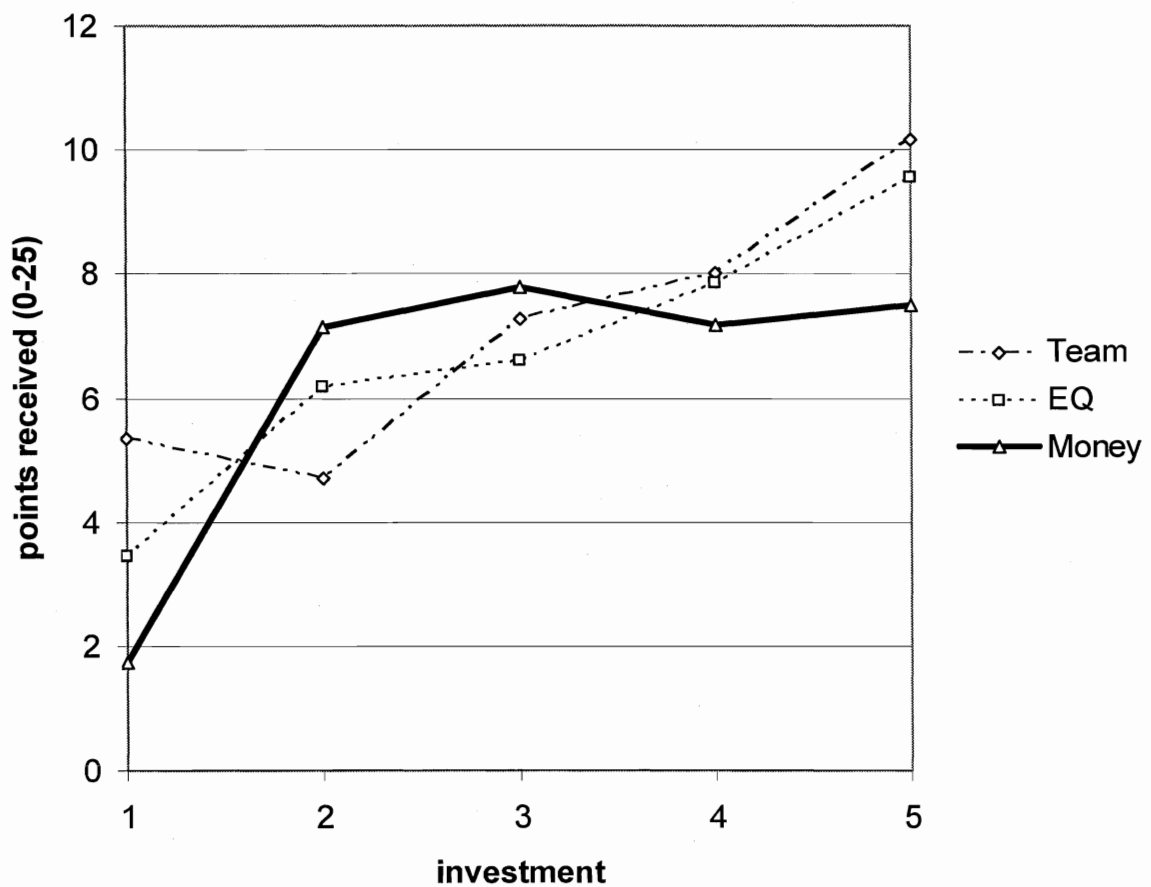
To summarize, we replicated the general reciprocity effect. The more a student invests in his or her team works, the more points he or she receives in the game. However, students who go beyond the norm (i.e. by doing more than their fair share) receive differential rewards

according to the stakes of the game. Their excess investment did not pay off in the money game, but it did pay off in the games in which social rewards were at stake.

In addition to the major results, it is also worth mentioning that friendship had an important positive effect on points received: $F(1,469) = 93.3$. This effect was not different for the games ($F<1.0$) but it moderated the investment effect: $F(4,469) = 3.49$. To explore the friendship by investment interaction, we followed Aiken and West's (1991) procedure. We calculated the b of reputation at low and high levels of friendship. We found that the relation between the opponent's investment and cooperation increased with decreasing levels of friendship (b at low levels of friendship: 0.24; C.I. 0.16-0.32 and b at high levels of friendship: 0.08; C.I. -0.01-0.17), suggesting that friendship buffered the reciprocity effect (like in study 1). The three-way interaction between friendship, investment, and the game was not significant, $F(8,469) = 1.42, p > .15$.

It is also interesting to note that the correlation between cooperative behavior in phase 1 and cooperative behavior in phase 2 (averaged over opponents) was very low: r = .04, ns. (n = 27). Finally, we calculated the correlation (n=39) between grades on four different tasks and the average perceived fairness of their investment in team works. The correlation between average perceived fairness and the score on two team assignments was close to zero: -0.10 and -0.5, both ns. The correlation between average perceived fairness and (1) the score on an individual assignment was 0.52, $p < 0.001$, and (2) the score on the exam was 0.39, $p <.02$.

Discussion

In this Study, we replicated the differential reaction to altruism and cooperation. Altruistic behavior in phase 1 was not returned in monetary value, but it was returned in social pay-off. So, in comparison with students who did their fair share, students who did more than their fair

share received more points in the EQ contest, were entitled to more rights to build teams, but did not receive more money. They gained status, not resources.

Participants deciding how to share their monetary endowment seemed to follow the basic rule "share equally", unless the other really exploited the team. Doing a lot yielded as much money as doing slightly less than the norm. In contrast, participants deciding how many points they should give to others in social contests, clearly follow a more fine-tuned strategy. Points received were clearly linearly related to investment in previous encounters, and this reward followed even beyond the norm.

The present method solves the ambiguities in Study 1. First, the interpretation in terms of altruists as *suckers* does not hold here, because the altruists receive more points in the E.Q. contest, and more points in the time construction game. Winning these contests has clear advantages, so it is not plausible that *suckers* would receive these benefits. Second, the behavior in phase 1 was a naturalistic behavior, and the perception was subjectively centered on the norm. This solves the problem that the norm was too vague in Study 1. Third, the explanation in terms of a methodological artifact is ruled out because the rewards were paid in the same currency for the three types of return.

Interestingly, academic performance was positively related to the average perceived fairness. Students who were perceived, on average, as having done more than the norm obtained higher scores on two individual tasks but not on team assignments. One interpretation is that intelligence might be the quality that the students signal. Another interpretation might be that students assessing each other's contribution have difficulties distinguishing effort from intelligence. Therefore, further data are required to support this claim.

General discussion

In the two studies reported, we found evidence for the distinction between cooperation and altruism. To avoid debate about the awareness of the underlying purposes, we focused on the receiver's reaction to altruism and cooperation. We found that receivers, deciding how much of their monetary endowments they would share, return investments to the public good (such as a study project) up to the norm, but do not follow the altruist beyond the norm. However, receivers, deciding on the social rewards they can provide, return investments in the public good not only up to the norm but also beyond. So, altruism does not buy resources, but yields honor. We first turn to the question why this differntiation might occur. We then turn to the question whether this is costly signaling, to the question about the nature of the quality signaled, and do suggestions for future research.

The most puzzling question is why receivers return social rewards but not material rewards. A first possibility is that altruism as costly signaling evolved from reciprocity. Reciprocity implies equality. However, if the extent of exchange increases, some people might fail to return what they received. If the receiver cannot return what he received, he is in debt. Norms might result from such escalating reciprocity. It could reflect a point beyond which a large majority of people cannot reciprocate anymore without risking immediate survival. Given the reciprocity rule, receivers people incur psychological debts. Returning social rewards might be the only way for them to ward off punishment from the giver. Another mechanism might be that receivers realize the importance of altruists in the group. As a way of making sure that these individuals stay in the group, they might reward altruists with status. These rewards could serve as 'social glue' (e.g. van Vugt and Hart, 2004).

Is the altruism displayed in our studies costly? In Study 1, altruists give more than half of their endowments, but they receive only 50% of the other's endowment in return in phase 2. That is, if they had given 50% of their endowments instead of more, they would also have received 50% of the other's endowment. Thus, altruists lose money due to their altruistic

behavior. In Study 2, altruists do not score higher grades than their lazy peers (there is no correlation between contribution and grade). That means that their effort is only poorly rewarded in terms of grades. Furthermore, receivers of the altruism do not return the investment in terms of money. In later phases, their investment does not pay off in terms of additional resources. We can conclude that the behavior identified as altruism in the present two studies is indeed costly.

Is the behavior an example of signaling? Signaling implies a message and a receiver. The data indicates that the message is received, because the receiver's behavior is affected by altruism. That is, altruists are wanted as a partner (Study 1) and are entitled to more social rewards (Study 2). What is less clear is the content of the message. The findings are consistent with any quality (e.g. altruism, intelligence) or with good intent. In Study 2, the signal correlates with academic performance. This relation can be due to the fact that intelligent people signal quality by behaving altruistically, but it could also be the case that intelligent behavior during team meetings is interpreted as altruistic, because these students contribute so much to the public good. The question which quality altruistic behavior actually signals is an important one and certainly deserves empirical attention.

The fact that few people follow altruists in phase 2 could shed some light on this issue. Because all participants actually *have* 20 cents, sharing is not merely related to current resource availability. If this had been the case, altruism would have been more common in phase 2. It seems safe to suggest that altruism is related to some individual difference variable that determines whether the present resources are shared or not, or at least interacts with the situational cues to determine whether resources are shared.

References

Aiken, L.S., & West, S.G. (1991). *Multiple regression: testing and interpreting interactions*. California: Sage Newbury Park

Axelrod, R. (1984). *The evolution of cooperation*. New York: basic books.

Axelrod, R. & Hamilton, W.D. (1981). The evolution of Cooperation. *Science 211*:1390-96.

Bergström, T.C. (2002). Evolution of Social Behavior: Individual and Group Selection. *Journal of Economic Perspectives 16*(2): 67-88.

Bliege-Bird, R., Smith, E.A., & Bird, D.W. (2001). The hunting handicap: costly signaling in human foraging strategies. *Behavioral Ecology & Sociobiology 50*: 9-19.

Boone, J.L. (1998). The Evolution of Magnanimity: When Is It Better To Give Than To Receive? *Human Nature 9*(1): 1-21.

Chapuisat, M., & Keller, L.(1999). Testing kin selection with sex allocation data in eusocial Hymenoptera. *Heredity 82,* 473-478.

De Cremer, D., Snyder, M., & Dewitte S.(2001). 'The less I trust, the less I contribute (or not)?' The effects of trust, accountability and self-monitoring in social dilemmas. *European Journal of Social Psychology 31(1)*: 93-107.

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

Gintis H., Bowles S., Boyd R., Fehr, E. (2003). Explaining altruistic behavior in humans *Evolution and Human Behavior 24,* 153-172.

Gintis H., Smith, E.A., & Bowles, S. (2001). Costly Signaling and Cooperation. *Journal of Theoretical Biology 213,* 103-119.

Gurven, M. (in press) To give or not to give: an evolutionary ecology of human food transfers. *Behavioral and Brain Sciences*.

Gurven, M.D, W. Allen-Arave, K. Hill, A.M. Hurtado. (2000). It's a Wonderful Life: signaling generosity among the Ache of Paraguay. *Evolution and Human Behavior 21*:263-282.

Hamilton, W.D. (1964). The genetic evolution of social behaviour. *Journal of Theoretical Biology, 7*, 1-52.

Kameda T., Takezawa M., Tindale R.S., Smith C.M. (2002). Social sharing and risk reduction - Exploring a computational algorithm for the psychology of windfall gains. *Evolution and Human behavior, 23*, 11-33.

Kirk, R.E. (1995). *Experimental design: Procedures for the behavioral sciences (3rd Ed.)*. California: Brooks and Cole Belmont.

Lindskold S., Han G., Betz B. (1986). Repeated persuasion in interpersonal conflict. *Journal of Personality and Social Psychology 51 (6)*: 1183-1188

Miller, D.T. (1977). Personal Deserving versus Justice for Others: An Exploration of the Justice Motive. *Journal of Experimental Social Psychology 13 (1)*: 1-13.

Murnighan, J.K., & Roth, A.E. (1983). Expecting continued play in prisoner's dilemma games. A test of several models. *Journal of Conflict Resolution 27(2)*: 279-300.

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

Orbell, J.M, Vandekragt, A.J.C, & Dawes, R.M. (1988). Explaining Discussion-Induced Cooperation. *Journal of Personality and Social Psychology 54(5)*: 811-819.

Richards P.W. (1974) The concept of dominance and methods of assessment. *Animal Behaviour, 22*: 914-930

Sethi, R. & Somanathan E. (2003). Understanding reciprocity. *Journal of Economic Behavior and Organization 50 (1)*: 1-27

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

Van Lange PAM, Ouwerkerk JW, Tazelaar MJA. (2002). How to overcome the detrimental effects of noise in social interaction: The benefits of generosity. *Journal of Personality and Social Psychology 82*: (5) 768-780.

Van Vugt, M., & Hart, C.M. (2004). Social Identity as Social Glue: The Origins of Group Loyalty. *Journal of Personality and Social Psychology 86*: (4) 585-598.

Wedekind, C; & M. Milinski (2000). Cooperation through image scoring in humans. *Science 288 (May 5, 5467)* 850-852.

Wilkinson G.S. (1984). Reciprocal food sharing in the vampire bat. *Nature 308 (5955)* 181-184

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

Wilson E.O., 1975. *Sociobiology.* The Belknap Press of Harvard University Press, Cambridge.

Zahavi, A. & Zahavi, A. (1997). *The Handicap Principle: A Missing Piece of Darwin's Puzzle.*

New York: Oxford University Press.

Appendix A

A's Earnings in €cents		B gives				
		20	...	10	...	0
A gives	20	40 *(R)*		20		0 *(S)*
	...					
	10	50		30		10
	...					
	0	60 *(T)*		40		20 *(D)*