ELSEVIER

Evolution and Human Behavior xx (2008) xxx – xxx

Evolution and Human Behavior

Original Article

Spotting altruistic dictator game players and mingling with them: the elective assortation of classmates

Julia Pradel[a,*], Harald A. Euler[b], Detlef Fetchenhauer[a,c]

[a]Department of Economic and Social Psychology, University of Cologne, Germany
[b]Department of Economics, Institute of Psychology, University of Kassel, Germany
[c]Interuniversity Center for Social Science Theory and Methodology, University of Groningen, Netherlands

Initial receipt 5 September 2008; final revision received 11 September 2008

Abstract

Altruism can evolve through assortation if the selfish advantage of egoistic individuals is outcompeted by the benefits of mutual cooperation between altruists. This selection process is possible if (a) individuals can distinguish altruists from egoists and (b) altruists cooperate electively with other altruists, leaving egoists no chance but to mingle with each other. This study investigates whether these two conditions are fulfilled in a natural setting. One hundred twenty-two students of six secondary school classes (age 10 to 19 years) played an anonymous dictator game, which functioned as a measure of altruism. Afterwards and unannounced, the students had to estimate their classmates' decisions and did so better than chance. Sociometry revealed that the accuracy of predictions depended on social closeness. Friends and disliked classmates were judged more accurately than liked classmates or those met with indifference. Moreover, altruists were friends with more altruistic persons than were egoists. The results confirm the existence of the two prerequisites for the evolution of altruism through assortation: the predictability of altruistic behavior and the association of altruists.

Keywords: Altruism; Assortation; Dictator game; School students; Social perception
PsycInfo classification: 2360 Motivation & Emotion; 3040 Social Perception & Cognition

1. Introduction

From birth to their grown-up years, humans daily interact with others and feel a need for social attachment (Baumeister & Leary, 1995). The dependence of humans on their social world can be risky in terms of resources, reproduction, and survival, if individuals choose the wrong interaction partners. The evolutionary perspective therefore suggests that humans should have developed adequate skills to judge other persons accurately in order to find the right lovers, friends, and allies, but to avoid the villains (Haselton & Funder, 2006).

One criterion for a good social interaction partner is altruism. Altruism denotes an individual's willingness to give up resources in order to benefit others. The evolution of altruism has puzzled scientists for generations, and although there has been considerable progress in its theoretical explanation, the question how such self-detrimental behavior could survive the pressures of natural selection has not been answered unanimously. Most game-theoretical models of the evolution of altruism assume random encounters between interaction partners (Aktipis, 2004). However, for many species, encounters with others are nonrandom. It has thus been supposed that the evolution of altruism may have been driven by assortation processes (Wilson & Dugatkin, 1997), that is, the gathering of like-minded individuals (for a general overview of social selection via the cooperation between correlated characters, see Frank, 1998, 2006).

Consider a population with egoists and altruists. If individuals were really equipped with a "personality judgment instinct" (Haselton & Funder, 2006), they should be able to distinguish altruists from egoists. As altruists should have an interest in assorting themselves, they would

* Corresponding author. Department of Economic and Social Psychology, University of Cologne, Herbert-Lewin Str. 2, 50931 Cologne, Germany.
E-mail address: julia.pradel@uni-koeln.de (J. Pradel).

1090-5138/$ – see front matter © 2008 Published by Elsevier Inc.
doi:10.1016/j.evolhumbehav.2008.09.003

leave self-interested individuals no chance but to stay among each other. As altruists could consequently reap extra benefits through mutual cooperation, the selfish advantage of egoistic individuals would be outcompeted. Altruism would become adaptive as it would provide a long-term benefit (West, Griffin, & Gardner, 2007, have argued that such behavior should be termed *mutualism*, rather than *altruism*).

Several scholars (Boorman & Levitt, 1973; Eshel & Cavalli-Sforza, 1982; Frank, 1988; Peck, 1992; Wilson & Dugatkin, 1997) construed scenarios similar to the one outlined above, in which altruists prefer like-minded interaction partners against the selfish rest of the population. To make assortation, as specified in the scenario, work as a driver of the evolution of altruism, two conditions have to be fulfilled: (a) individuals have to be able to distinguish altruists from egoists and (b) altruists have to elect like-minded individuals for mutual cooperation in order to reap synergetic extra benefits. The present study was designed to test whether these two assumptions hold true in a natural context, as only then assortation may reliably be invoked as an additional explanation for the evolution of altruism.

As to the first condition, there is indeed evidence that humans are equipped with a psychological mechanism to detect cheaters in social interactions (Cosmides & Tooby, 1992). A few studies examined whether humans are also capable of predicting altruistic behavior reliably. Frank, Gilovich, and Regan (1993) offered groups of participants the opportunity to get acquainted with each other before playing a prisoner's dilemma game. Subjects were able to discuss the paradigm for 30 min and make nonbinding pronouncements to their interaction partners about their game decision (defection or cooperation). Subsequently, the participants indicated their actual response covertly and were asked to predict the responses of their counterparts. Predictions were better than chance. However, Ockenfels and Selten (2000) criticized these results, suggesting that the accuracy of the predictions might have been due to the explicit pronouncement of defectors. As defectors would have had little interest in deviating from their announced decision, predicting their behavior was easy. Brosig (2002), therefore, replicated the experiment under restricted conditions. Pre-communication sessions were filmed so that subjects who explicitly announced their intention to defect were excluded from the analysis. Accuracy of predictions exceeded chance level under these conditions as well.

Going a step further, Verplaetse, Vanneste, and Braeckman (2007) investigated whether cooperation is predictable by only minimal visual information about the person. Subjects were shown photos of unknown target persons that had been taken when the target persons made a decision in a one-shot prisoner's dilemma game, and they had to rate how cooperative these target persons had behaved. The subjects were able to differentiate between cooperative and noncooperative target persons. Fetchenhauer, Groothuis, and Pradel (unpublished data) could even prove that subjects may predict the level of altruistic behavior on the basis of 20 s of silent video clips that were recorded in a neutral setting unrelated to altruistic behavior. Thus, first impressions seem to give humans a clue of what kind of person they are dealing with.

As to the second condition for the evolution of altruism by assortation, namely, that altruists choose like-minded persons for mutual cooperation, assortation with respect to mating ("assortative mating") is a well-established empirical phenomenon (Mascie-Taylor, 1995; Spuhler, 1968) that has been shown for a variety of somatic and psychological characteristics. The phenomenon, furthermore, exceeds mate choice as it has been evidenced as well for friendship (Berscheid, 1985). Sheldon, Sheldon, and Osbaldiston (2000) investigated assortative partner choice with particular respect to prosociality. The authors asked university freshmen to recruit three peers to participate in an *N*-person prisoner's dilemma game. Subjects with a prosocial value orientation tended to assort with one another. They thus achieved a group-level advantage in the game returns, which counteracted the individual-level advantage of antisocial participants.

In sum, data show that persons seem to be capable to predict the level of altruism in other persons and, moreover, tend to associate with those persons who show a level of prosociality similar to their own. Yet, the empirical evidence so far is based on laboratory studies with unknown ecological validity. The present study investigates whether the prediction of altruistic behavior and the grouping of altruistically like-minded individuals hold true also in genuine social groups in a natural setting. School classes are such genuine social groups, and a study of students of different grades offers the opportunity to investigate the possible age dependence of prosociality assessments and assortation processes.

In school classes, relationships between students vary in intensity. A typical student has a few friends among classmates, several she simply likes, and others she might dislike. Additionally, the relationship to some classmates could be one of indifference. This setting, therefore, offers the opportunity to examine the relationship between social closeness and the accuracy of judgments about other students' altruistic behavior. In the present study, altruistic behavior was operationalized as the decision in a dictator game, in which each student was asked to secretly divide a sum of money between himself or herself and another anonymous classmate. Subsequently, each student was asked to predict how each classmate had divided the money in the dictator game. We assumed that individuals might be better in predicting the behavior of their friends, with whom they have shared many experiences, than the behavior of more distant persons. As to the study of assortation, the school setting furthermore allows one to observe as to how far students choose friends who make dictator game decisions like they themselves do.

It might be noted that when making predictions about the accuracy of character assessments, evolutionary theory takes a position very different to the one represented by mainstream social psychology. The latter would not expect adolescents to be capable of accurately assessing the behavior of their peers in such a situation as social cognition in general is regarded to be biased (Kunda, 1999). For example, Ross, Greene, and House (1977) described the so-called false consensus effect, which is due to an individual's unfounded assumption of similarity between himself and a target person. Fetchenhauer and Dunning (2006) showed the false consensus effect to be present in a variety of game-theoretical paradigms. However, the assumption of similarity does not always provoke a false consensus effect but may induce accurate predictions, if the target person is indeed similar to the judge (Dawes & Mulford, 1966; Hoch, 1987). We shall, therefore, refer to this effect neutrally as "consensus effect". Because the consensus effect has been shown to be frequent within close relationships (Schul & Vinokur, 2000), it needs to be controlled for when studying social perception.

Another factor to be taken into account is the better-than-average effect, the pervasive tendency of individuals to assume that they are superior to others. This effect has been shown for a large range of socially desirable attributes, altruism included (Dunning, 2005). Therefore, persons can be expected to underestimate the trustworthiness and altruism of others (Fetchenhauer & Dunning, in press). Both effects, the consensus effect and the better-than-average effect, are independent and may coexist as individuals may base their predictions on their own behavior (consensus effect) while at the same time using a somewhat lower anchor (better-than-average effect) to estimate the behavior of others (Fetchenhauer & Buunk, unpublished data).

The present study tests the following three hypotheses: (a) Students are able to predict the level of altruistic behavior of their classmates in a dictator game. (b) Classmates who play altruistically in the dictator game are more often labeled as likable or as friend than egoists are. (c) Classmates positively assort themselves in their friendships along the dimension of altruism; that is, altruists have friends who play altruistically, too, and egoists have friends who play egoistically, too.

2. Methods

2.1. Participants

Participants were 127 students (60 female, 67 male) from six secondary school classes of a different grade each. Age varied from 10 to 19 years: 5th grade (aged 10 to 11), 8th grade (aged 13 to 14), and 10th, 11th, 12th, and 13th grade (aged 15 to 19). Size of classes varied from 14 to 29 students. Parents had given written consent for their child's participation and were debriefed after the experiment.

2.2. Procedure

The experiment was conducted in two secondary schools in Cologne, Germany. Sessions, one for each class, were carried out in social science lessons. No detailed information about the experiment was given to the participants in advance. In order to guarantee experimenter–subject anonymity, a subject number was handed out to every participant. To enable participants to link subject numbers to classmates, which was necessary for a later task, the participants were seated in a circle so that they could see each other with their subject number. This procedure assured that no real names entered the data sets.

The students were first familiarized with the logic of the dictator game without naming the game as such. The participants then received the first part of the questionnaire, which was coded with their subject number, and were asked to secretly mark down their own dictator decision. This means that they had to imagine a situation with two persons, a dictator and a recipient, in which the dictator receives a certain amount of money. Taking the perspective of the dictator, students had to consider the distribution of the money between themselves and the recipient, who was an anonymous student of the class. Neutral vocabulary was used all throughout the instructions to avoid influencing the participants in any possible way. For example, the dictator game was referred to as a "distribution task" and the dictator was referred to as "Person A" while the anonymous recipient was called "Person B".

The participants were informed that their decisions had a tangible consequence in that a payment would be made a week later. To maintain comparable incentives, the sum to be divided increased with the students' age, ranging from €6 (about US$9) in 5th grade to €10 (about US$15) in 13th grade. These payments corresponded approximately to the age-specific recommendations for allowances of the German youth welfare offices (Sport-, Schul-, & Jugendamt Springe, n.d.). The participants could pass money in whatever whole-numbered proportion they wanted but could also keep the entire amount for themselves. The recipients were unknown to the participants and the participants were told that—although for the payment session, recipients were secretly allotted to them later—they would never be informed about the identity of their respective classmate.

Communication was not allowed during the whole experiment. It was always made clear that information on the actual behavior of any person could never be traced back to a face or name, neither by any other subject nor by the experimenter, as data management and analysis for each class were made by different persons other than the experimenter.

After all participants had made their dictator decision, they received the second part of the questionnaire in which they were told that they now had to act as judges and assess how each of their classmates had decided in the preceding distribution task. The participants could not have expected

this task when they had made their own dictator game decisions. For motivational reasons, it was announced that the participant with the most accurate predictions would win a ticket for the local cinema.

For a sociometric analysis, the participants were finally asked to write down which of their classmates they characterized as friend, as likable, or as dislikable. The participants were free to nominate as many classmates as they wanted for each of these three categories.

For the purpose of payment management, half of the participants were randomly assigned the role of the dictator after the experiment, while the other half were assigned the role of the recipient. Each dictator was randomly matched with a recipient. Resulting pairs were paid 1 week later according to the distribution chosen by the respective dictator. The participants received the money in a closed envelope labeled with the subject number. No participant obtained any information about the identity of the other classmate with whom they had been paired.

2.3. Data preparation

As the amount of money that had to be distributed in the dictator game increased with the participants' age, payment values were transformed into percentages for reasons of comparability. Five out of 127 students (3.9%) passed more money to the recipient than they kept for themselves. According to Camerer (2003), dictator game contributions above 50% are rare in Western cultures. The fairness principle of equality, which demands parity of resources, would suggest contributions of 50% but nothing more. So would the equity rule, which demands a division of resources according to effort. Since both students, the dictator and the recipient, expended the same—namely, none—effort in the acquisition of this money, an equal split would have been suggested. Thus, it could not be ruled out that outliers who passed more than 50% might have misunderstood the experiment, especially since all these five participants were from lower grades. For this reason, these five participants were excluded from all further analyses. In order to maintain consistency, those 3.7% of the predictions in which students had expected their classmates to keep less than 50% for themselves were excluded as well. However, results were only minimally influenced by these exclusions.

For the sociometric analysis, the participants had been asked to classify their relations to every other classmate according to three levels of social closeness: friendship, liking, or dislike. Because students were free to make as many nominations as they wanted, some interindividual relations were not classified at all. These relations were consequently labeled as "indifferent". As class size varied, relationship nominations between students of different classes were incomparable (i.e., the number of nominations was likely to depend on the number of students in the relevant class). Thus, relationship nominations were transformed into relative values taking into account the possible number of nominations that could be made in the respective class.

Two sets of data were created. Data Set 1 comprised 122 cases, one for each participant, with information about the participants' own decision in the dictator game, their average prediction concerning others, how the behavior of the participants was predicted on average by their classmates, and the relative number of nominations of friendship, liking, and dislike each individual received. In Data Set 2, each case represented a single prediction from one student (judge) concerning another (target). As students were asked to predict all their classmates' behavior, (n_i-1) predictions were received from each judge, where n refers to the number of students in the respective class. A total of 2437 cases were obtained from all six classes. Moreover, Data Set 2 included information on the nature of the relationship between each judge and the respective target (i.e., the level of social closeness).

3. Results

3.1. Actual and predicted dictator game contributions

Participants, on average, contributed 37.3% towards the other person. The largest fraction (49% of the participants) handed half of the money to Person B. A total of 8% of the participants kept the entire amount for themselves, while 43% made a contribution in-between. Participants slightly underestimated the altruism of their classmates as the average predicted contribution amounted to 34.0% [$t(122)=3.09$, $p<.01$, $d=0.24$].

Girls acted more altruistically (contribution of 42.1±11.6%) than boys [32.8±18.1%, $t(122)=-3.39$, $p<.001$, $d=0.61$]. This sex difference was reflected by the participants' predictions about their classmates, as girls were predicted to make higher contributions than boys [37.9±6.4% vs. 30.0±9.4%, $t(122)=-5.36$, $p<.001$, $d=0.98$].

3.2. The ability to estimate the altruistic behavior of others

Our first hypothesis stated that students are able to estimate the level of altruistic behavior of their classmates in the dictator game. Because every participant rated every other participant, predictions were nonindependent. Moreover, the variance of individual predictions could be assumed to be partly attributable to class membership. To control for these facts, we used a method developed by Warner, Kenny, and Stoto (1979). The average prediction of several individuals concerning a single student was adjusted so that the effects of repeatedly occurring judges and class membership were sorted out (see Appendix A). These adjusted average predictions were z-transformed for reasons of clarity and comprehensibility.

We explored the relationship between actual contribution and adjusted average predictions by a partial correlation analysis controlling for class membership. This was necessary as the variance of actual contributions was affected

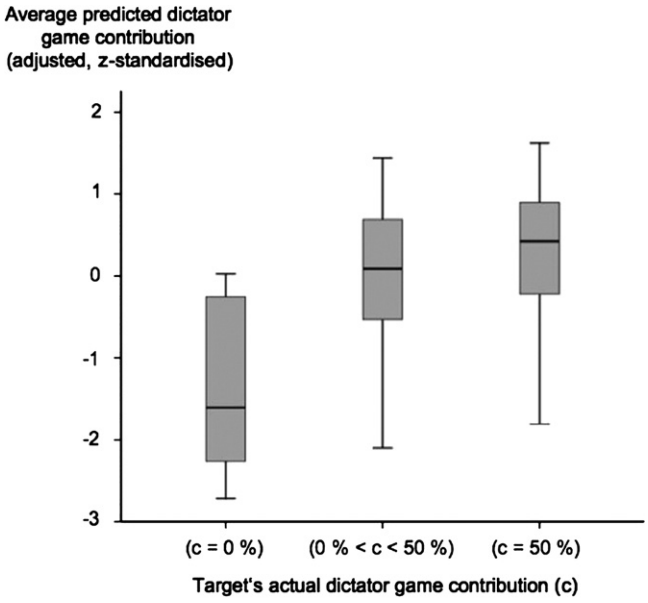


Fig. 1. The average prediction of classmates concerning the dictator game behavior of a target (N=122) depending on the target's actual contribution in the dictator game. Prediction values are adjusted for nonindependence due to class membership and repeatedly occurring judges and targets.

by class membership too. Five dummies were created, one for each class, with Class 5 as the reference category. The average prediction of a target's behavior indeed correlated with her actual behavior (r>=.39, p<.001, one tailed), showing that students were able to predict the dictator game behavior of their classmates better than chance. This result persisted when the target's sex was controlled for (partial r>=.31, p<.001, one tailed), indicating that classmates were able to make out those variations in the degree of altruistic behavior that went beyond differences due to the sex of the target. To highlight this result, we trichotomized the actual dictator game behavior into three categories: egoistic behavior (no contribution), moderate altruistic behavior (0%<contribution<50%), and altruistic behavior (contribution=50%). As can be seen in Fig. 1, the effect of classified actual dictator contributions was highly significant [F(2, 119)=9.27, p<.001]. The average predicted contribution pertaining a person who kept all the money was more than 1.5 standard deviations lower than the average predicted contribution pertaining a person who transferred 50%. Post hoc tests revealed that the average predictions for those students who had contributed nothing differed significantly from the average predictions for those who belonged to the altruistic categories. However, average predictions for moderate altruistic contributions (0%<contribution<50%) and altruistic contributions (50%) did not differ significantly from each other. Fig. 1 thus accentuates that students performed particularly well in spotting those classmates who had kept all the money for themselves.

Having focused on the analysis of average reputation so far, we now wanted to find out whether individual predictions concerning the behavior of a target were a good indicator for actual behavior, too. We thus repeated our analysis by using Data Set 2, in which a single prediction (again adjusted for nonindependence of cases, see Appendix A) of one student concerning the behavior of another was treated as the unit. Students' individual predictions matched the behavior of their classmates significantly better than chance (r=.21, p<.001, one tailed). The match was maintained when the target's sex was partialled out (partial r=.16, p<.001, one tailed). Thus, the hypothesis that persons are able to identify altruistic disposition in classmates was confirmed.

3.3. The social appraisal of altruism

Our second hypothesis was that classmates who distribute the money evenly in the dictator game are more often labeled as likable or as friend than egoists are. To test this, we initially explored the relation between the number of received liking nominations and the individual's level of altruism. The correlation was slightly positive with class membership and sex partialled out (partial r=.14, p=.07, one tailed). Likewise, the effect of altruism on received dislike nominations was slightly negative (partial r=−.12, p=.10, one tailed). As to the question of whether altruists had more friends than egoists, the correlation between the number of received friendship nominations and altruism was not significant (partial r=−.08, p=.20, one tailed). Summarizing these results, the hypothesis that individuals who share the money are more popular than individuals who keep everything for themselves could not be confirmed.

3.4. The assortation of altruists

The third hypothesis claimed that classmates positively assort themselves along the dimension of altruism. To test this, we investigated the relationship between the dictator contributions of friendship nominators and the contributions of their friendship nominees that correlated significantly with each other (partial r=.21, p=.01, one tailed), with class membership partialled out. The correlation persisted when controlling for sex (partial r=.19, p=.02, one tailed), which excludes the explanation of an artifact due to sex specificity of friendship dyads, namely, that altruistic girls are friends with altruistic girls and self-interested boys are friends with self-interested boys. Thus, our hypothesis that altruists choose like-minded persons as friends was confirmed.

3.5. Accuracy of predictions and grade

In order to investigate the relationship between the accuracy of predictions and grade, a variable was needed, which indicated the quality of predictions of every individual. For this purpose, the individual accuracy rate of each participant was calculated by correlating the predictions of judges with the actual behavior of their respective targets. A total of 122 correlation coefficients resulted, which can be considered to be individual accuracy rates. However, a relationship between the individual accuracy rates and grade

Table 1
Accuracy of prediction for different types of relationships (Pearson correlation coefficients)

	Complete subsamples		Random parts of subsamples	
Relationship	*n*	Accuracy of prediction	*n*	Accuracy of prediction
Mutual friendship	418	.34**$_a$	52	.35**$_a$
Mutual liking	246	.11*$_b$	51	.16$_a$
Mutual dislike	186	.18**$_{a,b}$	41	.30*$_a$
Mutual indifference	274	−.07$_c$	35	.09$_a$

Correlation coefficients that do not share subscripts differ significantly at $p<.05$. *n*, number of predictions.
* $p<.05$.
** $p<.01$.

could not be proven ($r=.10$, $p=.30$, one tailed). Younger students performed as well as older students in predicting their classmates' decisions.

3.6. Accuracy of predictions and social closeness

We further aimed at finding out whether the level of social closeness between two persons (i.e., friendship, liking, dislike, indifference) influenced the accuracy of their predictions. Of all 2437 dyadic relationships, 17% were mutual friends (both participants nominated each other as friends). Mutual liking occurred in 10% of the dyads, mutual dislike in 8%, and mutual indifference in 10%. The remaining dyads (55%) were mixed relationship declarations (e.g., Person A declared friendship and Person B declared liking), which were excluded for further analyses.

Accuracy rates (correlations between actual behavior and predictions) for the four relationship types are shown in Table 1 (complete subsamples). The accuracy of predictions varied with social closeness and was highest for mutual friendships, followed by dislike relationships and, finally, mutual liking relationships. The accuracy of predictions of unrelated individuals was lowest and reached chance level only. Thus, students were most accurate when they judged friends and individuals they disliked.

Dyads had been treated as independent cases until now. However, some dyads were connected such that one of the dyad members occurred repeatedly (e.g., because he or she was a mutual friend with more than one person). Hence, the present subsamples could be construed as asymmetric Social Relations Model designs whose estimation is complex (Kenny, Kashy, & Cook, 2006). We refrained from controlling for the interconnections between cases as any corrections should have had no influence on effect sizes but on standard errors only. Yet, the stability of social closeness effects on accuracy of predictions was tested under tightened conditions so that every dyad member only appeared once. Cases were randomly selected, which means that if a person was a friend with two other individuals, only one of these dyadic relationships was included in the sample of mutual friends via random sampling. As it can be seen in Table 1 (random parts of subsamples), accuracy rates were still significant for friendships and dislike relations. Accuracy rates for liking relations and indifferent peers were not significant. These results were comparable to those of the complete samples: individuals judged those persons best to whom they had a decided opinion—being either positive or negative. The following analyses were thus based on the complete samples, treating predictions of individuals as independent ones.

Table 2
Actual similarity between dyad members (Pearson correlation coefficients)

Relationship	*n*	Objective similarity
Mutual friendship	209	.16**$_a$
Mutual liking	123	.09$_a$
Mutual dislike	93	−.22*$_b$
Mutual indifference	137	.08$_a$

Correlation coefficients that do not share subscripts differ significantly at $p<.05$. *n*, number of dyads.
* $p<.05$.
** $p<.01$.

In the analysis so far, the influence of a possible consensus effect had not been controlled for. The potential influence of a consensus effect can be seen in Table 2, which shows that individuals who disliked each other behaved differently in the dictator game, whereas mutual friends behaved alike. The high accuracy rates for friends could thus have been driven by a consensus effect.

To find out how much of the reported accuracy rates in dyads was really due to the judge's explicit ability to identify the target's level of altruism ("true accuracy") and how much of the prediction was due to bias in form of a consensus effect, we applied the model of Kenny and Acitelli (2001) for the simultaneous measure of accuracy effect and consensus effect (see Appendix B). Note that, subsequently, the term *accuracy effect* denotes the proportion of variance within predictions due to true accuracy and the term *overall accuracy* denotes the correlation between predictions and actual behavior of targets, which had been given in Table 1.

For all types of relationships, Table 3 shows which proportion of the predictions was due to the observers' distinct ability to identify altruists (accuracy effects) and which proportion was due to bias (consensus effects). Predictions were indeed inflated considerably by the consensus effect for all relationships. However, under control of the consensus effect, true accuracy could be proved for friends and dislike

Table 3
Accuracy and consensus effects (unstandardized regression coefficients)

Relationship	*n*	Accuracy effects	Consensus effects
Mutual friendship	209	.28*$_a$	.58*$_a$
Mutual liking	123	.10$_b$	.40*$_b$
Mutual dislike	93	.27*$_{a,b}$	.34*$_b$
Mutual indifference	137	.05$_{a,b}$	.53*$_{a,b}$

Regression coefficients in the same column that do not share subscripts differ significantly at $p<.05$. *n*, number of dyads.
* $p<.001$.

Table 4
Partitioning of the overall accuracy correlations (r)

Relationship	n	r	True accuracy	Increment due to consensus effect
Mutual friendship	209	.34	.25	.09
Mutual liking	123	.11	.08	.03
Mutual dislike	93	.18	.25	−.07
Mutual indifference	137	−.07	n.a.	n.a.

As the accuracy correlation for mutual indifferences was insignificant, the validity of the decomposition could not be assured. n, number of dyads.

relations. Thus, the hypothesis that altruistic behavior can truly be predicted was at least confirmed for these two groups.

To analyze how the concurrence of accuracy and consensus effects contributed to overall accuracy rates, we decomposed the correlations between actual and predicted contributions into true accuracy and the increment due to bias. As can be seen in Table 4, mutual friends benefited from the consensus effect as they had made identical dictator contributions. In contrast, the overall accuracy of individuals who disliked one another was reduced as they, too, assumed similarity, which did not actually exist. Predictions of individuals who liked one another were mainly driven by true accuracy, though a positive impact of the consensus effect could be shown nonetheless. For unrelated individuals, a decomposition of the overall accuracy rate was impossible as overall accuracy was insignificant.

To summarize, beyond the consensus effect, a true ability to identify altruists could be demonstrated. Biased overall accuracy as well as true accuracy was highest for friends and disliked persons, indicating that predictions were more valid when the judge had a firm attitude towards the target, whether this attitude was positive or negative, than when the attitude was one of indifference or just acceptance.

4. Discussion

The purpose of the present study was to test the hypotheses that (a) individuals are able to predict the level of altruistic behavior in a dictator game of individuals familiar to them, (b) individuals who play altruistically in the dictator game are judged more positively than egoists and have more friends than egoists, and (c) individuals positively assort themselves in their friendships along the dimension of altruism. Hypotheses 1 and 3 were supported. Individuals are able to identify altruism, and altruists preferably choose other altruists for mutual cooperation. Hypothesis 2 could not be answered unambiguously as judgments of affection concerning target persons only marginally correlated with their degree of altruism. Furthermore, we wanted to explore the influence of age and social closeness on the validity of social perception. While the accuracy of predictions was linked to social closeness, it was independent of age.

Accuracy of predictions was measured by asking subjects to predict the dictator contributions of their classmates. This task can be considered demanding for the participants as Dunning, Griifin, Milojkovic, and Ross (1990) showed that accurate predictions of someone's behavior strongly depend on the availability of two types of information: A judge needs to have accurate information about the exact nature of the situation confronting the target and about the target's construal or understanding of the situation at the point the target must respond. Even though participants had precise information on the objective features of the situation, as they all took part in the same anonymous dictator game, they still needed to accurately anticipate the meaning of this situation from the individual perspective of each classmate. This was likely to be a difficult task as it has to be presumed that students might have interpreted the situation in different ways.

Following Hagen and Hammerstein (2006), cues of the degree of reputational consequences in economic games are processed in a manner that may be intuitive or affective and, thus, cognitively impenetrable. Admittedly, our experimental setting, in which we seated participants in a circle to enable them to link classmates with subject numbers, might have created a context that students perceived as partly public, despite the explicit propositional information in the instructions of the dictator game, which announced that anonymity would always be guaranteed.

There is evidence that removing anonymity from economic games increases individuals' contributions to a public good (Andreoni & Petrie, 2004). Moreover, it has been shown that already subtle cues of observability, namely, images of eyespots, affect generosity (Burnham & Hare, 2007; Haley & Fessler, 2005). It can be assumed that participants who behaved strictly egoistically when sitting face to face with their classmates almost surely would have behaved equally egoistically if anonymity had been even more certain. Per contra, it must be presumed that the behavior of those who transferred money could have been partly evoked by cues of publicity that activated evolutionarily relevant psychological mechanisms that manage reputation independent of explicit reasoning.

Observed altruistic behavior in the noniterated anonymous dictator game, therefore, cannot unconditionally be interpreted as "pure" altruism in the sense that it was independent of reputation management efforts. It rather seems plausible that there were some individuals who were strongly inclined to manage their reputations and cooperated because they perceived the experimental situation as public and, thus, followed rules that are rational in their everyday life (McKenzie, 2003) while there were others who were intrinsically motivated to behave altruistically regardless of the specification of the situation.

With reference to the accuracy of predictions, we do not know whether participants' predictions grounded on the exact assessment of the extent to which a participant perceived the situation as private or public and acted accordingly; nevertheless, students were able to predict with some accuracy how their classmates would behave. Hence, whatever the level of anonymity that was perceived by an actor, it seems reasonable to assume that either the

same level of anonymity factored into his or her predictions about others' behavior or—even more elaborate—he or she made predictions about how the respective target person would interpret and, thus, act in this ambiguous situation. In any event, individuals were able to estimate their classmates' decisions to some extent and they performed especially well in predicting egoistic decisions. This can be considered adaptive because altruistic persons should always—irrespective of their intrinsic or extrinsic motivation—be favored over egoists as cooperation partners.

The average dictator contribution of students in our study amounted to 37%. As Camerer (2003) reported, the mean dictator contribution of adults generally ranges from 20% to 40%. In a study by Takezawa, Gummerum, and Keller (2006), mean dictator contributions of 11- to 14-year-olds amounted to 44%. Thus, the level of contributions in our study is similar to that of earlier studies. Moreover, our results replicate the findings of Harbaugh, Krause, and Libay (2003) that the dictator contributions of children and adults are comparable. Benenson, Pascoe, and Radmore (2007) suggested that even young children have biologically based altruistic dispositions that are additionally fostered by socialization practices. In line with earlier studies (Eckel & Grossman, 1998), females in our study acted slightly more altruistically than males did.

Although the participants succumbed to the better-than-average effect and underestimated the absolute level of dictator contributions of their classmates, they correctly differentiated between the altruistic behavior of boys and girls in that they predicted girls to pass a higher amount than boys. As to the relative accuracy of predictions, the students performed well as the variation of average predictions matched the variation of actual behavior. Individual predictions, too, were significantly better than chance. Moreover, the students did not base their predictions on sex alone, as it could be shown that over and above the ability to differentiate between altruistic tendencies of boys and girls, students took into account the variation of altruism within both sexes. Thus, adolescents are indeed able to distinguish self-interested individuals from altruists.

The accuracy of predicting dictator contributions did not increase with grade; 11-year-olds performed as well as 19-year-olds. This result seems especially noteworthy when one considers that the 19-year-old participants had known each other for up to 9 years, while the youngest students had only been acquainted for several months. Moreover, in a follow-up-study, we could show that even 9-year-old children achieve comparable outcomes (Pradel & Fetchenhauer, unpublished data). As recent results indicate, infants as young as 6 months can take into account an individual's helping or hindering actions towards others by evaluating this individual as appealing or aversive (Hamlin, Wynn, & Bloom, 2007). The early development of capabilities to evaluate the prosocial behavior of others and their developmental stability support the proposition that the ability to evaluate altruistic dispositions in others is not primarily dependent on general learning or general socialization experiences but is the result of biological adaptation.

Although altruistic students favored like-minded persons as friends, altruists were not generally favored as interaction partners. Liking and antipathy nominations only marginally correlated with the degree of altruism. Moreover, altruists and egoists had a comparable number of friends. This finding contradicts the hypothesis about the popularity of altruistic players. However, it should be mentioned that these German students nominated, on average, only three classmates as friends. The term *friend* is used quite selectively in Germany and more equivalent to the term *best friend(s)* in English-speaking countries. With the benefit of hindsight, one might wonder whether the hypothesis that egoists are spurned in general was justified as it is not actually generated from the models of assortment that we investigated. For example, Frank (1988) did not suggest that egoists do not have friends at all. He supposed instead that altruists preferentially cooperate with other altruists, and this supposition is indeed supported by our data. Nevertheless, it seems understandable that egoists have friends too. They just choose or get friends who are less altruistic than friends of altruists.

Predictions about the dictator contribution of a friend were more valid than predictions about the contributions of liked and unrelated classmates, as can be expected from the intensity of communication and interaction in friendships. Interestingly, the overall accuracy for predicting the dictator contribution of disliked classmates was high as well and did not differ significantly from the accuracy of predicting the behavior of friends. When controlling for the consensus effect, a similar pattern emerged. True accuracy was highest when students evaluated friends, but similar levels of true accuracy were reached for the evaluation of explicitly disliked persons. In ancestral environments, a valid evaluation of exactly these two groups must have conferred fitness benefits. Persons ought to know about the true intentions of their best friends as they are interacting with them regularly and are thus highly dependent on correct assessments of their characters. People should further be aware of the intentions of disliked individuals, that is, those with different interests, because they are potential adversaries. However, we have to admit that the direction of causation could be the opposite. It could be that individuals befriend those who they believe to have good intentions (and who truly have good intentions), whereas they dislike those who they believe to have evil intentions (and who indeed have evil intentions).

In sum, our study shows that humans—although succumbing to some well-known biases of social perception—are able to predict altruistic or self-interested behavior of others and that they moreover tend to assort themselves in their friendships along the dimension of altruism. Speculating about the consequences of these findings, it seems reasonable to expect that altruists, who mingle electively with each other, reach extra benefits through mutual cooperation. As these benefits may counteract the individual advantage of egoists, it

can be assumed that altruism may have evolved due to assortation processes between altruists who acknowledged each other's cooperative tendencies.

Our empirical results thus support the theories of assortation and complement findings on various mechanisms that have probably contributed to the evolution of altruism, like reciprocity (Trivers, 1971), social norms and punishment (Fehr, Fischbacher, & Gächter, 2002; Fehr & Gächter, 2002, 2003), to name only a few. We therefore think that this study adds a small but important piece to solve the puzzle of human altruism.

Although the ability to spot altruists was proven, the exact nature of cues to altruism remains unknown. Frank (1988) argued that altruism is motivated by moral emotions. Drawing on the fact that emotions are linked to nonverbal behavior and involuntary facial expressions (Darwin, 1872), which are difficult to produce on command (Ekman, 1985), altruism might have been signaled through emotional displays. However, as this explanation is rather vague, future research is needed to open the black box of processes underlying the intuitive assessment of altruism. Thus, the conclusion that persons can predict the altruistic behavior of acquaintances and that they draw consequences from their insights stands. Birds of a feather flock together—so do altruists.

Acknowledgments

We thank Lisa Maria Podleyski and Jens Schäfer for their assistance with data collection; Bernd Weiss for methodical support; and Werner Greve, Alexandra Haferkamp, and Sascha Schwarz for valuable feedback on an earlier draft of the manuscript. Further thanks go to several participants of the HBES 2007 conference (Williamsburg, VA) and the MVE 2008 conference (Hamburg, Germany), especially to Ulf Toelch, for helpful comments and suggestions on this study.

Appendix A. Controlling for nonindependence of prediction scores within school classes

Warner et al. (1979) showed that in a so-called Social Relations Model in which every individual of a certain group rates every other individual of this group, a prediction X of an actor i concerning partner j can be expressed by the following equation:

$$X_{ij} = m + a_i + b_j + g_{ij}$$

where

- m denotes the mean prediction within the group;
- a_i is the actor effect for person i, which is his tendency to exhibit a consistent level of predictions across interaction partners;
- b_j is the partner effect of person j, which again measures the tendency of group members to judge a person in a consistent way; and
- g_{ij} is the relationship effect, which is set at the dyad level. It measures the prediction of actor i concerning partner j after removing their individual-level tendencies (i.e., the actor and the partner.

The estimation of the actor effect equals:

$$a_i = \frac{(n-1)^2}{n(n-2)} M_{i.} + \frac{n-1}{n(n-2)} M_{.i} - \frac{n-1}{n-2} M_{..}$$

where

- n is the group size,
- $M_{i.}$ is the mean of predictions exhibited by person i,
- $M_{.i}$ is the mean of predictions received by person i, and
- $M_{..}$ is the mean of all the predictions.

The estimation of the partner effect equals:

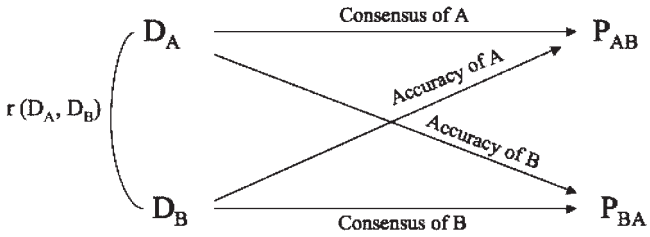
$$b_i = \frac{(n-1)^2}{n(n-2)} M_{.i} + \frac{n-1}{n(n-2)} M_{i.} - \frac{n-1}{n-2} M_{..}$$

As the partner effects are precisely those values, which signify the tendencies of targets to be predicted as behaving altruistically by all their classmates, we used partner effects as adjusted prediction scores when testing Hypothesis 1 on an aggregated level (e.g., the question whether the average reputation of a student matched his or her actual behavior). This means that actual dictator contributions of the students were correlated with their partner effects (N=122).

When testing Hypothesis 1 on the individual level (e.g., whether individual predictions, too, matched the behavior of their classmates), we additionally integrated the unique perception of an actor concerning a partner into our analysis. That is, the sum of partner effects and relationship effects belonging to a single prediction (b_i+g_{ij}) was correlated with the actual dictator contributions of the student (n=2437).

Appendix B. Measuring true accuracy and the consensus effect in mutual predictions of two persons

Kenny and Acitelli (2001) depicted a paradigm for the simultaneous measure of accuracy and consensus effect in two-person relationships: the Actor–Partner Interdependence Model (APIM). Our replica is presented in the following figure:



It consisted of four variables: (a) the dictator contribution of Person A "D_A", (b) the dictator contribution of Person B "D_B", (c) the prediction of Person A concerning the dictator

contribution of Person B "P_{AB}", and (d) the prediction of Person B concerning the dictator contribution of Person A "P_{BA}". While the actual dictator contributions of the two persons D_A and D_B were treated as determinants, the predictions of the dictator contributions P_{AB} and P_{BA} (in which the first subscript refers to the judge and the second subscript refers to the target) were treated as outcome variables. The four paths represent accuracy and consensus effects in the predictions of the two persons: diagonal paths from D_A to P_{BA} and from D_B to P_{AB} denote accuracy effects, whereas horizontal paths from D_A to P_{AB} and from D_B to P_{BA} denote consensus effects as they refer to the extent to which the judge implicitly assumes that the target will make a dictator contribution like he himself did. Whether the consensus effect lowers or betters predictions depends on whether Person A and Person B are actually similar (correlation between D_A and D_B).

We estimated the paths of the APIM using multilevel modeling with SPSS. The dyad was treated as the unit. As Person A and Person B were indistinguishable (i.e., both were friends or both nominated each other as likable, etc.), accuracy and consensus effects were assumed to be similar for both persons.

Having calculated accuracy and consensus effects, regression coefficients were standardized. Subsequently, overall accuracy rates (i.e., the correlation of actual behavior and predictions) could be split into true accuracy and the increment due to bias, with the latter being the product of actual similarity and the consensus effect. The equation reads as follows:

$$\text{overall accuracy} = \text{true accuracy} + \text{increment due to bias}$$
$$r(\mathrm{D},\ \mathrm{P}) = \text{accuracy effect} + r(\mathrm{D_A},\ \mathrm{D_B})\text{*consensus effect.}$$

References

Aktipis, C. A. (2004). Know when to walk away: Contingent movement and the evolution of cooperation. *Journal of Theoretical Biology, 231*, 249−260.

Andreoni, J., & Petrie, R. (2004). Public goods experiment without confidentiality: A glimpse into fund-raising. *Journal of Public Economics, 88*, 1605−1623.

Baumeister, R. F., & Leary, M. R. (1995). The need to belong: Desire for interpersonal attachments as a fundamental human motivation. *Psychological Bulletin, 117*, 497−529.

Benenson, J. F., Pascoe, J., & Radmore, N. (2007). Children's altruistic behavior in the dictator game. *Evolution and Human Behavior, 28*, 168−175.

Berscheid, E. (1985). Interpersonal attraction. In G. Lindzey, & E. Aronson (Eds.), *The handbook of social psychology* (3rd ed., pp. 110−168). New York, NY: Random House.

Boorman, S. A., & Levitt, P. A. (1973). A frequency-dependent natural selection model for the evolution of social cooperation networks. *Proceedings of the Natural Academy of Sciences of the USA, 70*, 187−189.

Brosig, J. (2002). Identifying cooperative behavior: Some experimental results in a prisoner's dilemma game. *Journal of Economic Behavior & Organization, 38*, 43−57.

Burnham, T. C., & Hare, B. (2007). Engineering human cooperation. Does involuntary neural activation increase public goods contributions? *Human Nature, 18*, 88−108.

Camerer, C. (2003). *Behavioral game theory: Experiments on strategic interaction.* Princeton, NJ: Princeton University Press.

Cosmides, L., & Tooby, J. (1992). Cognitive adaptation for social exchange. In J. Barkow, L. Cosmides, & J. Tooby (Eds.), *The adapted mind* (pp. 163−228). New York: Oxford University Press.

Darwin, C. R. (1872). *The expression of the emotions in man and animals.* London: John Murray.

Dawes, R., & Mulford, M. (1966). The false consensus effect and overconfidence: Flaws in judgment or flaws in how we study judgment? *Organizational Behavior and Human Decision Processes, 65*, 201−211.

Dunning, D. (2005). *Self-insight: Roadblocks and detours on the path to knowing thyself.* New York: Psychology Press.

Dunning, D., Griffin, D., Milojkovic, J., & Ross, L. (1990). The overconfidence effect in social prediction. *Journal of Personality and Social Psychology, 58*, 568−581.

Eckel, C., & Grossman, P. (1998). Are women less selfish than men? Evidence from dictator games. *The Economic Journal, 108*, 726−735.

Ekman, P. (1985). *Telling lies.* New York: W. W. Norton.

Eshel, I., & Cavalli-Sforza, L. L. (1982). Assortment of encounters and evolution of cooperativeness. *Proceedings of the National Academy of Sciences of the USA, 79*, 1331−1335.

Fehr, E., Fischbacher, U., & Gächter, S. (2002). Strong reciprocity, human cooperation, and the enforcement of social norms. *Human Nature, 13*, 1−25.

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

Fehr, E., & Gächter, S. (2003). Egalitarian motive and altruistic punishment. *Nature, 425*, 392−393.

Fetchenhauer, D., & Dunning, D. Do people trust too much or too little? *Journal of Economic Psychology* (in press).

Fetchenhauer, D., & Dunning, D. (2006). Perceptions of prosociality and solidarity in self and others. In D. Fetchenhauer, A. Flache, A. P. Buunk, & S. Lindenberg (Eds.), *Solidarity and prosocial behavior* (pp. 225−242). New York: Kluwer Academic/Plenum Publishers.

Frank, R. H. (1988). *Passions within reason.* New York: Norton.

Frank, R. H., Gilovich, T., & Regan, D. T. (1993). The evolution of one-shot cooperation: An experiment. *Ethology and Sociobiology, 14*, 247−256.

Frank, S. A. (1998). *Foundations of social evolution.* Princeton, NJ: Princeton University Press.

Frank, S. A. (2006). Social selection. In C. W. Fox, & J. B. Wolf (Eds.), *Evolutionary genetics: Concepts and case studies* (pp. 350−363). New York: Oxford University Press.

Hagen, E. H., & Hammerstein, P. (2006). Game theory and human evolution: A critique of some recent interpretations of experimental games. *Theoretical Population Biology, 69*, 339−348.

Haley, K. J., & Fessler, D. M. T. (2005). Nobody's watching? Subtle cues affect generosity in an anonymous economic game. *Evolution and Human Behavior, 26*, 245−256.

Hamlin, J. K., Wynn, K., & Bloom, P. (2007). Social evaluation by preverbal infants. *Nature, 450*, 557−559.

Harbaugh, W. T., Krause, K. S., & Liday, S. J. (2003). Bargaining by children. University of Oregon Economics Working Paper No. 2002-4, doi:10.2139/ssrn.436504 (Available at SSRN).

Haselton, M. G., & Funder, D. C. (2006). The evolution of accuracy and bias in social judgment. In M. Schaller, J. A. Simpson, & D. T. Kenrick (Eds.), *Evolution and social psychology* (pp. 16−37). New York: Psychology Press.

Hoch, S. J. (1987). Perceived consensus and predictive accuracy: The pros and cons of projection. *Journal of Personality and Social Psychology, 53*, 221−234.

Kenny, D. A., & Acitelli, L. K. (2001). Accuracy and bias in the perception of the partner in a close relationship. *Journal of Personality and Social Psychology, 80*, 439−448.

Kenny, D. A., Kashy, D. A., & Cook, W. L. (2006). *Dyadic data analysis.* New York: The Guilford Press.

Kunda, Z. (1999). *Social cognition.* Making sense of people. Cambridge, MA.

Mascie-Taylor, C. G. N. (1995). Human assortative mating: Evidence and genetic implication. In A. J. Boyce, & V. Reynolds (Eds.), *Human populations. Diversity and adaptations* (pp. 86−105). Oxford, UK: Oxford University Press.

McKenzie, C. R. M. (2003). Rational models as theories—not standards—of behavior. *Trends in Cognitive Sciences, 7*, 403−406.

Ockenfels, A., & Selten, R. (2000). An experiment on the hypothesis of involuntary truth signaling in bargaining. *Games and Economic Behavior, 33*, 90−116.

Peck, J. R. (1992). Group selection, individual selection, and the evolution of genetic drift. *Journal of Theoretical Biology, 159*, 163−187.

Ross, L., Greene, D., & House, P. (1977). The "false consensus effect": An egocentric bias in social perception and attribution processes. *Journal of Experimental Social Psychology, 13*, 279−301.

Schul, Y., & Vinokur, A. (2000). Projection in person perception among spouses as a function of the similarity in their shared experiences. *Personality and Social Psychology Bulletin, 26*, 987−1001.

Sheldon, K. M., Sheldon, M. S., & Osbaldiston, R. (2000). Prosocial values and group-assortation within an *N*-person prisoner's dilemma. *Human Nature, 11*, 387−404.

Sport-, Schul- & Jugendamt Springe (n.d.). Infoblatt: Taschengeld für Kinder und Jugendliche. Retrieved August, 18th, 2008, from http://www.springe.de/Download/40-Info%20Taschengeld-neu.pdf.

Spuhler, J. N. (1968). Assortative mating with respect to physical characteristics. *Eugenics Quarterly, 15*, 128−140.

Takezawa, M., Gummerum, M., & Keller, M. (2006). A stage of the rational tail of the emotional dog: Roles of moral reasoning in group decision making. *Journal of Economic Psychology, 27*, 117−139.

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

Verplaetse, J., Vanneste, S., & Braeckman, J. (2007). You can judge a book by its cover: The sequel. A kernel of truth in predictive cheating detection. *Evolution and Human Behavior, 28*, 260−271.

Warner, R. M., Kenny, D. A., & Stoto, M. (1979). A new round robin analysis of variance for social interaction data. *Journal of Personality and Social Psychology, 37*, 1742−1757.

West, S. A., Griffin, A. S., & Gardner, A. (2007). Social semantics: Altruism, cooperation, mutualism, strong reciprocity and group selection. *Journal of Evolutionary Biology, 20*, 415−432.

Wilson, D. S., & Dugatkin, L. A. (1997). Group selection and assortative interactions. *American Naturalist, 149*, 336−351.