ELSEVIER

Evolution and Human Behavior 30 (2009) 201–211

Evolution and Human Behavior

Collective action in culturally similar and dissimilar groups: an experiment on parochialism, conditional cooperation, and their linkages☆,☆☆

Ruud Koopmans[a,b], Susanne Rebers[a,*]

[a]Faculty of Social Sciences, Department of Sociology, Vrije Universiteit, Amsterdam, De Boelelaan 1081, 1081 HV, Amsterdam, The Netherlands
[b]Department "Migration, Integration, Transnationalization," Wissenschaftszentrum Berlin für Sozialforschung (WZB), Reichpietschufer 50, 10785 Berlin, Germany

Initial receipt 8 July 2008; final revision received 13 January 2009

Abstract

This study examines the effects of ingroup favoritism and outgroup hostility ("parochialism"), as well as of conditionally cooperative strategies, in explaining contributions to experimental public goods games. The experimental conditions vary group composition along two culturally inheritable traits (political party preference and religious affiliation) and one trivial, "minimal" trait (birth season). We contrast ingroup, outgroup, and random group conditions and investigate the relation between the own contribution to the public good and the expectations about other group members' behavior in each one of them. We find evidence for ingroup favoritism but no support for a separate tendency towards outgroup hostility. Further, conditional cooperation and ingroup bias are, to some extent, linked. Subjects had higher expectations of the contributions of ingroup members, and their own behavior was more strongly conditioned on other group members' expected behavior in the ingroup conditions. In ingroup contexts, subjects displayed a form of "strong reciprocity" by giving more than they expected others to at high expectation levels but less at low expectation levels. Once these interactions are taken into account, we do not find a direct effect of ingroup bias anymore. We discuss these results in the light of theories of cultural group selection and conclude that too much emphasis may have been laid on direct intergroup conflict. Our results suggest that differential cooperativeness, rather than parochialism, may characterize the behavior of individuals in cultural ingroups and outgroups.

Keywords: Collective action; Public goods; Cultural group selection; Parochialism; Conditional cooperation

1. Introduction

Humans often cooperate on a large scale with nonrelatives in order to obtain public goods. Public goods are goods that cannot be obtained by individual action and that are nonexcludable: once a public good is obtained, it is not possible to withhold group members from benefiting from the good, even if they have not contributed to its production. It has long been recognized that, although it is in the interest of a group of individuals that many people work together to obtain public goods, this does not explain why an individual would contribute to such "collective action." The influence an individual has on obtaining the good decreases as a function of group size, or in other words, the cost/benefit ratio increases with group size. Therefore, it is harder to explain cooperation in collective action than in dyadic cooperation (Hardin, 1968; Olson, 1965).

Many scholars have tried to explain the evolution of this human capacity for collective action. After all, the fact that people participate in it while they would be better off free riding on the participation of others poses a problem for individual-level selection theories. Traditional mechanisms to explain the evolution of cooperative behavior fail to explain the evolution of collective action (for an overview, see Henrich, 2004). One of the mechanisms that is able to explain cooperation among unrelated individuals is reciprocal altruism. This mechanism can explain the evolution of cooperation in repeated dyadic interactions. Essential for reciprocal altruism to evolve is that the cooperation is conditional: cooperative behavior must be targeted towards cooperators, and defections must be targeted towards those

☆ We gratefully acknowledge the financial support for our research by the Evolution and Behaviour programme of the Netherlands Organisation for Scientific Research (Grant Number 051-14-046).
☆☆ Both authors contributed equally to this article.
* Corresponding author.
E-mail address: s.rebers@fsw.vu.nl (S. Rebers).

1090-5138/$ – see front matter © 2009 Elsevier Inc. All rights reserved.
doi:10.1016/j.evolhumbehav.2009.01.003

who defect (Axelrod, 1984; Trivers, 1971). Conditional cooperation is also a common finding in collective action. Social psychological research has shown that the more people expect others to participate in a political protest event, the more likely they are to join themselves (Klandermans, 1984; 1997). Further, experimental research on the public goods game, which is a good experimental representation of the collective action problem because it entails the same tension between individual and group interests, shows that people contribute more the more they expect others to contribute (e.g., Dawes, 1980; Falk & Fischbacher, 2002; Fischbacher, Gächter, & Fehr, 2001; Marwell & Ames, 1979; Messick & Brewer, 1983). However, even though the cooperation is conditional, individual-level selection cannot explain the evolution of this large-scale cooperation. In collective action, targeting of cooperation and defection is impossible: when an individual cooperates, both cooperators and free riders benefit equally from the obtained good (e.g., Gil-White & Richerson, 2003). The fact that public goods are, by definition, nonexcludable implies that it is not possible to prevent free riders from enjoying the public good. Mathematical models have indeed shown that the larger the group, the more unlikely it is that cooperation for public goods can evolve (Boyd & Richerson, 1988).

An explanation for the evolution of collective action that has recently been proposed is cultural group selection (Boyd & Richerson, 1982, 1985; Henrich, 2004). In short, this theory states that if groups differ with respect to culturally evolved traits, groups with the most group-beneficial traits, like norms promoting participation in collective action, will be selected when there is between-group competition. Mathematical models have shown that if cultural group selection is responsible for the evolution of the human capacity for collective action, people should be more likely to participate in collective action when the public good that can be obtained will benefit members of one's own "cultural ingroup" than when it will benefit one's "cultural outgroup" (e.g., Boyd & Richerson, 1982). However, there has not been much research specifically aimed at testing whether present-day humans behave in accordance with the predictions of the theory (inspiring exceptions include Bernhard, Fischbacher, & Fehr, 2006; Gil-White, 2003). Therefore, we want to investigate whether participation in collective action indeed increases when the benefits of cooperation are shared with cultural ingroup members.

Findings from experimental research reveal differential behavior towards ingroups and outgroups that fits a cultural group selection explanation. Numerous "minimal group experiments" have shown that even if groups are based on trivial traits, like preferences for different abstract painters or colors, people behave more cooperatively towards their ingroup than towards outgroups (Tajfel, 1982; Tajfel et al., 1971; Vaughan, Tajfel, & Williams, 1981). Although we also include "minimal" groups (based on a person's birth season) in this article, our main focus is on more meaningful cultural traits, namely, political party preference and religious affiliation. Importantly, from the point of the cultural group selection perspective, these are traits that are transmittable by way of social learning and thereby potentially subject to cultural evolution. Ingroups and outgroups constructed on the basis of cultural traits may be expected to have a stronger behavioral effect than minimal group traits, a difference that, to our knowledge, has never been investigated before.

To our best knowledge, the role of cultural traits has only been investigated in other experimental games, not in the public goods game (Bernhard et al., 2006; Gil-White, 2003). Differential behavior towards ingroups and outgroups has been found in many dyadic games like the two-person prisoner's dilemma (e.g., Yamagishi et al., 2005) or the design used in the minimal group paradigm, in which an individual allocates money to both an ingroup and outgroup member (e.g., Tajfel et al., 1971). However, this does not automatically imply that the same behavioral pattern will be found in larger groups. As explicated above, cooperation in the collective action dilemma and, thus, in the public goods game is an evolutionary problem that is structurally different from dyadic forms of cooperative behavior, and it can therefore not be subsumed under the theory of reciprocal altruism.

Wilson and Sober have been among the most prominent advocates of the cultural group selection solution to the collective action problem (Wilson, 2002; Wilson & Sober, 1994). They emphasize that group selection can explain why the normative systems of religious, political, and other cultural groups are characterized by a double morality, which prescribes prosocial conduct towards members of the own group but, at the same time, allows, or even prescribes, exploitative or even outright hostile behavior towards other groups, a behavioral pattern known as parochialism (Bowles & Choi, 2003; Bowles & Gintis, 2004a; Choi & Bowles, 2007). However, there are two reasons why ingroup favoritism and outgroup hostility do not necessarily need to be linked. First, mathematical models have shown that large-scale cooperation can evolve by cooperating with ingroup members, without being hostile to outgroup members (e.g., Boyd & Richerson, 1982). Second, there are reasons to assume they are two distinct psychological mechanisms (e.g., Brewer, 1999). Therefore, in our experiment, we include ingroup conditions, outgroup conditions, and a random condition. Comparing the random condition to the other conditions, we are able to distinguish between ingroup favoritism and outgroup hostility. Since all existing theories on ingroup–outgroup differentiations predict that there is ingroup favoritism, we add this as one of our hypotheses. However, since there is theoretical disagreement on whether or not there will be outgroup hostility, we will add both possibilities as alternative hypotheses.

Several experiments point into the direction that the ingroup–outgroup differentiation is caused by ingroup favoritism, not by outgroup hostility. For instance, Kiyonari and Yamagishi (2004) find that if the partner in a two-person prisoner's dilemma is unaware of the focal participant's

group membership, the focal participant does not give more to ingroup members than to outgroup members. However, again, this finding in a dyadic game cannot be automatically assumed to be found in larger groups. One of the few studies on this topic in which groups larger than two have been utilized finds that in a three-person game with intergroup competition, ingroup–outgroup differentiations are caused by ingroup favoritism and not by outgroup hostility (Halevy, Bornstein, & Sagiv, 2008). Their experimental design, however, included neither minimal group nor cultural traits.

The hypothesis that people should cooperate more with their ingroup than with their outgroup is shared by adherents of the "bounded generalized exchange" hypothesis (e.g., Yamagishi, Jin, & Kiyonari, 1999; Yamagishi & Kiyonari, 2000). According to this hypothesis, exchanges of favors take place within demarcated groups. As a result, making costs to help an ingroup member will pay back, because group members will reciprocate your help. Differences in cooperation with ingroup and outgroup members can thus, according to this hypothesis, be explained by the higher cooperation levels people expect from ingroup members than from outgroup members, and not by a direct effect of the division in groups. Experimental evidence supports the idea that people have higher expectations of ingroup members and, therefore, contribute more in ingroups. For example, an experiment by Kiyonari and Yamagishi (2004) shows that in a two-person prisoner's dilemma, participants had higher expectations of ingroup members that were aware of their shared group membership than of ingroup members that were not aware of their shared group membership. Further, cooperation levels were significantly higher in the former than in the latter condition. Their findings are in line with a more recent experiment using existing social categories by Goette, Huffman, and Meier (2006). Integrating the generalized exchange hypothesis in our experiments, we want to investigate whether (part of) the ingroup–outgroup differentiation is caused by differences in expected cooperation levels between ingroups and outgroups. Again, although the fact that higher expectations in ingroups and the subsequent higher cooperation levels have been found in dyadic games, their effects have not been investigated in larger groups. Since cooperation in larger groups is a meaningfully different evolutionary problem, our investigation fills an important hiatus in existing knowledge.

A further linkage between the ingroup–outgroup differentiation and expected cooperation levels that we want to investigate is whether conditionally cooperative behavior is more pronounced in ingroups than in outgroups, a linkage that, to our best knowledge, has never been investigated before in either prisoner's dilemmas or public goods games. Although we do not know of any cultural group selection models or even published thought experiments that combine conditionally cooperative behavior and the division of individuals in ingroups and outgroups, we suspect that competition between groups could also lead to a differential response to expected cooperation levels between ingroups and outgroups. Responding more strongly to expectations of group members is likely to increase or at least stabilize cooperation in a group. Therefore, we expect that conditionally cooperative behavior is stronger in ingroups than in other groups.

In this study, we conduct public goods games in which participants can choose to divide an allocated sum of money between a private good and a public good. Before making this allocation decision, participants were shown information on group composition along two culturally inheritable traits, namely, political party preference and religion, and one minimal group trait, namely, birth season. For each trait category, we created an ingroup condition, in which most other group members held the same trait, and an outgroup condition, in which most group members held a different trait. In a seventh condition, each trait was varied randomly.

2. Research design

2.1. Participants and procedure

We first conducted a pilot study to determine whether group size mattered and whether we could conduct the experiment online, which was more convenient both for us and for participants. We compared groups composed of 3, 9, and 24 persons and found, in line with earlier studies (e.g., Marwell & Ames, 1979), no significant differences regarding the amount contributed to the public good (for details, see Appendix A). We also found no significant differences in contribution levels between public goods games conducted in the laboratory and those conducted online (for details, see Appendix B). This is in line with Amichai-Hamburger (2005), who found no differences in ingroup favoritism between participants playing a decision task online or in the laboratory. Of course, experimenters have a reduced degree of control over the experiment when conducting it online. For instance, it is possible that participants discuss their decision with other people. However, the results of both our pilot study and the study by Amichai-Hamburger convince us that if there is an influence of the noncontrolled setting, this only causes randomly distributed deviations that do not affect the results of the experiment. We chose to conduct the experiments online using the web survey tool Examine and opted for a group size of six. Even though our pilot results indicated that a group size of three would have been sufficient, we preferred a somewhat greater group size in which the impact of an individual's contribution on the group account (and thus the marginal return of an individual investment in the group account) is smaller, in order to more closely simulate the collective action dilemma. Moreover, a group size of six offered more possibilities to vary group composition.

Participants were 233 (80 male and 153 female) students recruited through posters and flyers on the VU University campus in Amsterdam, the Netherlands. They registered on a website, after which they immediately received

a pre-experimental questionnaire that took about 10 min to fill out. They received €1.00 (about US$1.50) for filling out the questionnaire, provided that they completed the entire experiment. Next to demographics and some other questions, participants were asked the following questions: "Imagine there would be Dutch parliamentary elections today. What political party would you vote for?", "What is your date of birth?", and "Do you consider yourself as belonging to a religious denomination, and if so, which one?" The answers to these questions were used to create the experimental conditions (see below). Participants were told that, after filling out the questionnaire, they would be assigned to groups to participate in a decision-making experiment and that this might take a couple of days. To ensure anonymity, they were told that although certain characteristics of group members might be shown during the experiment, these characteristics would only be shown on the aggregate level of the group and that personal information and individual decisions would never be made public.

After a couple of days, each participant was sent an e-mail with a link to a personalized website on which he or she could participate in the experiment. After the public goods game was explained and some test questions were asked to ensure participants had understood the payoff structure of the game, each participant played four public goods games, each time with different people, so as to ensure that decisions across the games would be independent one-shot games. For each game, they were asked both to decide how much they would contribute to the public account and to provide an estimate of how much they thought the other five members of their group had contributed on average. These estimates might partly be projections of the participants' own contribution on the contribution of their group members (Dawes, McTavish, & Shaklee, 1977). However, other research suggests that an important part of it reflects the actual expected contribution of group members (Yamagishi & Sato, 1986). After playing the four games, subjects filled out a short postexperimental questionnaire. Each participant was then paid via Internet banking.

As we told participants, groups indeed consisted of six members. However, some deception was involved in the fact that we manipulated the information on the traits that were held by the other group members. This enabled the systematic variation of group composition along each of the three trait dimensions, which was an essential element in this experiment. It would have been impossible to find real participants with the precise combinations of characteristics necessary to implement each of the different conditions. Resorting to other solutions to this problem, like using the strategy method, would not have been a good alternative for two reasons. First, we wanted to make sure that we would pay our participants the amount of money they made in each of the four games, because we wanted our participants to be motivated by the monetary outcomes of each of the games, to avoid unwanted variation caused by differences between participants in, for instance, risk-taking. Second, it has been shown that using the strategy method influences choice behavior in several games (e.g., Brandts & Charness, 2000), possibly as a result of choices made on the basis of emotions instead of rationality (Cook & Yamagishi, 2008), suggesting that this method could have an important influence on the outcomes of a cooperative dilemma.

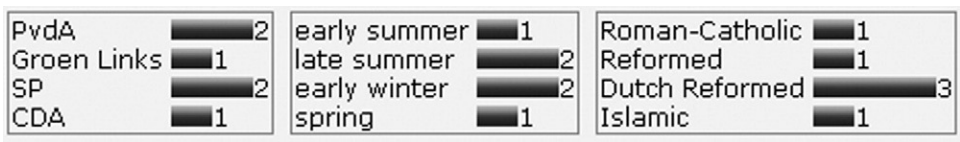
It was made clear to participants that they could only participate once in this study. In addition, we made sure that they did not participate more than once by checking name, e-mail address, bank account number, and social identification number.

2.2. Public goods game

Each individual received €1.00 per public goods game and could divide this over two accounts, an individual and a group account, in portions of €0.10. Each individual had his or her own individual account, and money put on this account would remain unchanged and would be returned to the participant after the game. Money put on the group account was multiplied by 1.8 and divided among all group members, irrespective of how much money (if any) each member had put on the group account. The public goods game is a good simulation of the dilemma of collective action, because the group receives the highest payoff if every group member contributes all his or her money to the public good, but on the individual level, the highest payoff is reached if an individual puts all his or her money on the private account, while sharing the benefits of contributions to the group account made by other group members.

2.3. Experimental conditions

In total, there were seven experimental conditions. The differences between the conditions resided in the composition of the groups. In each condition, participants were shown three graphs containing information on three sets of traits of the other group members. The trait categories were political party preference, birth season, and religion and were shown on the group level. An example is given in Graph 1.



Graph 1. An example of the three graphs participants were shown prior to decision making. From left to right, trait categories political party preference, birth season, and religion are shown. This example is drawn from the random group condition.

In the "ingroup" conditions, three of the other group members shared a trait with the focal participant for the relevant trait category. Of the other two group members, one had a randomly chosen trait that had a probability of one in six of being similar to that of the focal participant and one had a randomly chosen trait that was never similar to that of the focal participant. When the group was an "ingroup" with respect to one trait category (e.g., political party preference), the traits of the other two trait categories (e.g., birth season and religion) were randomly chosen, with the exception, of course, of the traits of the focal participant. In the "outgroup" conditions, four of the other group members had the same, randomly chosen trait that was not similar to the focal participant's trait. The fifth other group member had a randomly chosen trait that had a chance of one in six of being similar to that of the focal participant. Again, when the group was an "outgroup" with respect to one trait category, the traits of the other two categories were randomly chosen, again with the exception of the trait of the focal participant. The seventh condition, the random condition, was a condition in which all traits in all three categories, with the exception of those of the focal participant, were randomly chosen.

In the religion ingroup, the randomly chosen trait was replaced by the focal participant's trait because of a flaw in the script, causing the religion ingroups to be, on average, somewhat more homogenous than the politics and birth season ingroups. However, this does not pose a problem for the interpretation of our results. First, we do not want to compare the different ingroups among each other. Second, a Mann–Whitney *U* test showed no significant difference between contribution levels in the political or birth season ingroup conditions when three or four group members had a similar trait as the focal participant.

For each trait category, randomly chosen traits were chosen out of six possible traits. These traits were chosen based on the fact that they were the most often occurring traits among the people who had participated in two pilot experiments. For the politics conditions, the six traits were party preference for Groen links (green left), PvdA (social democrats), CDA (Christian democrats), VVD (liberals), SP (socialists), and CU (Orthodox Protestants). For the religion conditions, the traits were nonreligious, Roman Catholic, Reformed, Dutch Reformed, Islamic, or Hindu. For each trait category, one of these traits could be replaced by another trait in the case that a participant held a trait that was not among these most occurring traits. Birth season was defined according to 2-month periods, which do not coincide with astrological periods. We therefore assumed that participants would not have any prior systematic expectations regarding the behavior of individuals in these trait categories. Choosing the most frequently occurring traits increased the realism of the group compositions. Realism was also the reason why we chose not to make the ingroup and outgroup conditions fully homogenous. By including a random component in each of the conditions, we ensured that the group compositions did not seem artificially controlled by the experimenters. Indeed, none of the participants indicated in their answers to a postexperimental question about their ideas about the purpose of the experiment that they thought that the information they had been given about group composition was unrealistic. Each participant was randomly assigned to play four out of the seven conditions in a random order.

3. Hypotheses

From the theoretical discussion, we derive hypotheses based on the parochialism thesis of the cultural group selection perspective and hypotheses regarding the linkages between the parochialism thesis and the idea of conditional cooperation.

Hypotheses regarding ingroup favoritism and outgroup hostility ("parochialism"):

H1. Average contributions in the ingroup conditions will be higher than the average contributions in the outgroup conditions.

H2. The difference between the average contribution in ingroup and outgroup conditions will be higher for the two culturally inheritable traits (political party preference and religion) than for the minimal trait (birth season).

H3. The average contribution in ingroup conditions will be higher than average contributions in the random condition (ingroup favoritism).

H4a. The average contribution in outgroup conditions will be lower than the average contribution in the random condition (outgroup hostility).

H4b. The average contribution in outgroup conditions will be the same as the average contribution in the random condition (no outgroup hostility).

Hypotheses regarding the linkage between parochialism and conditional cooperation:

H5. Ingroup favoritism can be partly explained by higher expectations in ingroups. If Hypothesis 4a is confirmed, outgroup hostility can be partly explained by lower expectations in outgroups.

H6. Conditionally cooperative behavior will be more pronounced in ingroup conditions, meaning that expectations of other group members' behavior have a stronger effect on the own contribution in ingroups.

4. Results

4.1. Overview

Before starting analyses, we removed four contribution decisions (of four different participants) because an unintended error occurred in the group composition. Fig. 1 shows

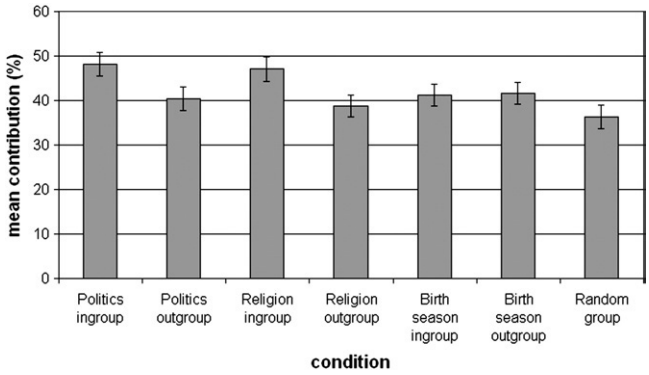


Fig. 1. The average contribution on the group account per condition, in percentage of the endowment.

the average contribution to the group account according to experimental condition. A first thing to note is that, on average, subjects contributed 42% of their endowment to the group account, which is within the range of 40–60% that is found in the large majority of one-shot public goods experiments. Also in other respects, our results across conditions are entirely in line with previous experiments. Only in 18% of the games did subjects contribute nothing to the public account. As in virtually all previous experiments, the modal strategy was to contribute exactly 50% of the endowment (Miller & Andreoni, 1991).

Fig. 1 shows that there is some variation across the experimental conditions, but even in the outgroup conditions, in which most group members differ from the deciding subject on one of the manipulated traits, the contribution levels are still substantial. Nevertheless, in the politics and religion ingroup conditions, people contribute about €0.08 more than in the respective outgroup conditions. It is further noteworthy that the lowest contribution levels are found in the random group condition, and not in the outgroup conditions, as the parochialism perspective would lead us to expect.

4.2. Ingroup favoritism and outgroup hostility

We start our analyses by testing the first hypothesis with multilevel regression models with contributions to the public good as the dependent variable. Multilevel regression analysis is appropriate in this case because of the interdependence of decisions taken by the same individual (see Raudenbush & Bryk, 2002). A two-level model had a significantly better fit than a one-level model (chi-square, $p<.01$). We therefore distinguished two hierarchical levels: the higher level of individual participants and the lower level of the four contribution decisions nested within the individual participants. We conducted three separate multilevel regressions, one for each trait category, with the relevant ingroup condition as independent variable and the relevant outgroup condition as the omitted reference category. Results can be found in Table 1.

Table 1
Results of multilevel regression analyses of own contribution level to the public good on ingroup condition

Trait category	*B*	*p*	R^2
Politics	7.385	.008	.061
Religion	5.249	.043	.004
Birth season	−1.472	.585	.012

Reference categories are outgroups, with ingroups as independent variables.

Table 2
Results of multilevel regression analysis of own contribution level to the public good on group composition

Trait category	Condition	*B*	*p*
Politics	Ingroup	7.901	.000
	Outgroup	3.000	.179
Religion	Ingroup	7.870	.000
	Outgroup	2.434	.267
Birth season	Ingroup	1.787	.422
	Outgroup	3.838	.091

Reference category: random group condition.

The table shows that participants contribute significantly more in the ingroup conditions than in the outgroup conditions for both cultural trait categories but not for the minimal trait conditions. Hypothesis 1 is therefore partly confirmed. Further, this finding automatically leads to a confirmation of Hypothesis 2: the difference between contributions in ingroups and outgroups is larger for cultural traits than for the minimal group trait.

The next multilevel regression analysis is aimed at testing Hypotheses 3, 4a, and 4b. Table 2 reports the results of the regression in which only the different compositions of groups are added as predictors, with the random group condition as the omitted reference category. In line with Hypothesis 3, we find evidence of ingroup favoritism in the politics and religion conditions, indicated by a significant positive influence of the ingroup condition on contributions. For the minimal groups, we find no ingroup favoritism. Further, in line with Hypothesis 4b, we find no indication for a significant tendency towards outgroup hostility, since there is no significant negative influence of outgroup conditions on contributions, indicating that subjects do not treat outgroup members worse than they do members of a randomly composed group.

4.3. Parochialism and conditional cooperation

Because the preceding analyses revealed no significant differences in the contribution levels among them, and to increase the power of the effects, we merged the two cultural ingroups, the two cultural outgroups, and the two birth season groups for the following regressions. To increase comparability, in Model I, we repeat the analysis we presented before in Table 2 but used the new merged group composition variables just described. Results are presented in Model I in the first column of Table 3. Again, we used the random group condition as the omitted reference category.

In Model II, we added the expectation participants had about other group members' contributions. In line with previous research, we find that expectations have a

Table 3
Results of multilevel regression analyses of own contribution level to the public good on group composition, expectations of others' contribution, and control variables (*B* coefficients with significance levels in parentheses)

Independent variable	Model I	Model II	Model III	Model IV
Cultural ingroup	7.898 (.000)	4.808 (.005)	0.072 (.980)	0.282 (.922)
Cultural outgroup	2.721 (.157)	1.095 (.518)	1.188 (.483)	1.164 (.491)
Birth season groups	2.761 (.155)	1.343 (.432)	1.473 (.389)	1.452 (.395)
Expectation of other group members' contribution		0.701 (.000)	0.663 (.000)	0.664 (.000)
Cultural ingroup×expectation interaction			0.104 (.042)	0.100 (.053)
Party preference: liberals				−7.982 (.020)
Party preference: green left				6.206 (.043)
R^2	.026	.240	.242	.243
n	927	927	927	927
N	233	233	233	233

Reference category: random group condition. *N* indicates number of subjects; *n* indicates number of contribution decisions.

significant influence on participants' contributions. Further, as the R^2 values indicate, the effect of expectations on contributions is much stronger than that of the ingroup condition: the explained variance rises from 2.6% to 24%. Although both the relevance of ingroup favoritism and conditional cooperation are confirmed by our data, the latter is by far the more powerful of the two. In addition, the results provide partial support for Yamagishi and Kiyonari's (2000) claim that ingroup biases are not so much due to identification with the group interest but to higher expectations regarding ingroup members' contributions. In line with this idea, the effect size of the ingroup condition is strongly reduced once we introduce expectations, suggesting that an important part of ingroup favoritism does not take the form of a direct bias but indirectly raises expectations about group members' contributions and thereby, through the mechanism of conditional cooperation, also raises the own contribution level. Hypothesis 5 is thus confirmed.

To test Hypothesis 6, we introduce in Model III the interaction effect between ingroup condition and expectations. In regression terms, this hypothesis implies that expectations have an additional (stronger) effect in the ingroup conditions than in the other conditions. This interaction term is indeed significant, which confirms Hypothesis 6. The direct effect of the ingroup condition now becomes insignificant.

The results thus show that the direct effect of ingroup conditions can be entirely explained away once we take the interaction between group composition and conditional cooperation into account. People do cooperate more with ingroups, but the effect is explained not by a direct ingroup bias but by the higher expectations that they have of ingroup members' contributions and by the fact that ingroup members condition their own contributions more strongly on those of other group members.

As a final step in our analysis, we add in the regression in the fourth column of the table (Model IV) control variables on the individual level in order to check the robustness of our results. We ran several regressions including as predictors the sex and age of respondents, as well as all political party and religious affiliations. We additionally included interaction effects between sex and group conditions, in order to check whether males and females reacted differently to ingroup and outgroup stimuli, as suggested by Van Vugt, De Cremer, and Janssen (2007). Most of these variables turned out not to have significant effects on contribution levels, with the exception of the two reported in the table. Adherents of the liberal VVD party contributed significantly less, and those of the green left party contributed significantly more. More importantly, the introduction of these control variables does not alter our main findings regarding the importance of expectations of other group members' contributions and of the stronger effect that these expectations have in ingroup conditions, although the significance level of the latter decreased slightly to .053.

4.4. The interplay between contributions, expectations, and group composition

We can get more insight into the relation between own contributions and expected contributions from others, depending on group composition, by looking at the differential between the two. Fig. 2 shows average expected

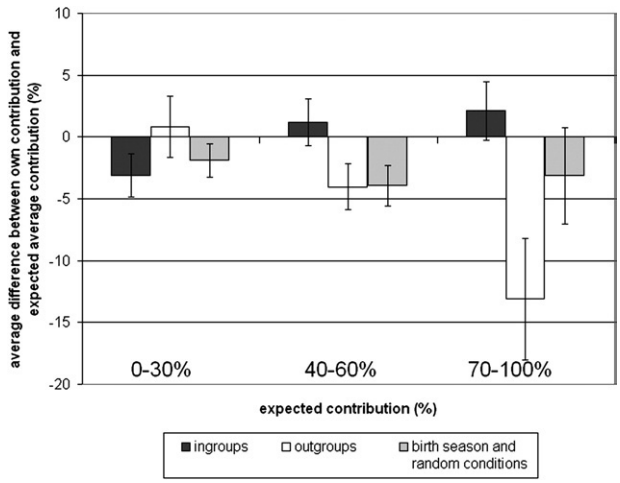


Fig. 2. Average difference between own contribution and average expected contribution for different expectation levels and group compositions.

contribution levels minus contribution levels in different experimental conditions as a function of the level of expectations: low (expected contribution level, 0–30% of the endowment), medium (expected contribution level, 40–60% of the endowment), or high (expected contribution level, 70–100% of the endowment). As the figure makes clear, contributing more than one expects others to contribute is most prevalent in ingroup conditions, especially when expectations about other group members' contributions are high. By contrast, when expectations are low, there is no tendency for ingroup members to give more than they expect others to—in fact, they give less in such conditions than people in outgroup or random/birth season conditions do, although the difference is not statistically significant. Mann–Whitney *U* tests of these differences show that there are no significant differences among the three conditions at low expectation levels. At medium and high expectation levels, the difference between ingroup and outgroup conditions is statistically significant at the .05 and the .01 level, respectively.

5. Conclusions and discussion

By analyzing the effects of parochialism and conditional cooperation simultaneously, our study aimed to throw more light on the possible linkages between these two perspectives on collective action. To our knowledge, such a combined assessment of these two perspectives has not been undertaken in previous experimental research. Our analyses focus on groups larger than two, which enables us to study cooperative behavior that, from an evolutionary view, cannot be explained by reciprocal altruism. Our study also improves over many earlier studies of group biases by combining two features. First, our design includes a trivial, "minimal" trait (a person's birth season) and two more meaningful, culturally inheritable traits (political party preference and religious affiliation). Second, for each of these three traits, we constructed an ingroup condition, in which most other group members shared a trait with the focal subject, and an outgroup condition, in which most other group members held a different trait than the focal subject. Unlike most previous studies, we also included a third type of group composition, in which other group members' traits were randomly chosen. This allows us to distinguish ingroup favoritism from outgroup hostility.

Like earlier studies, we found that subjects contributed significantly more to the group account in ingroup than in outgroup conditions. However, this was only true for the two culturally inheritable traits, whereas we found no significant difference between ingroup and outgroup conditions for the trivial birth season trait. The possible reason why our results deviate from earlier findings in the minimal group research tradition is that if only minimal group information is presented, the artificiality of the design motivates participants to behave in a way that they expect the experimenters to predict. However, these minimal group characteristics lose their significance for subjects' decisions once other more meaningful cultural cues are available (Gil-White, 2003).

By comparing both ingroups and outgroups to the random group condition, we further found that although there is consistent evidence for ingroup favoritism, this does not imply outgroup hostility, as subjects' contributions in the outgroup conditions were not significantly different from (and even a bit higher than) those they made in randomly composed groups. Our evidence thus confirms the view of authors who have argued that ingroup preferences do not need to imply outgroup hostility (e.g., Allport, 1954; Brewer, 1999; Mummendey et al., 1992).

This finding seems to run counter to the emphasis in much of the cultural group selection literature on outgroup hostility and intergroup conflict, most prominently in the form of war. It is important to realize, however, that neither intergroup hostility nor direct conflict is necessary for cultural group selection to operate. All that is strictly required is that individuals behave in such a way that they bestow more benefits on members of their own cultural group than on those of other groups. This may take the form of being nice to ingroup members and hostile to outgroup members, but this is only one possible scenario. As long as the behavior results in fitness advantages for the ingroup compared to the outgroup, it may entail indifferent or even mildly positive behavior towards outgroups (e.g., Wilson, 2002, p. 156), for instance, because such behavior reduces the costs of intergroup conflict (see, similarly, Gil-White, 2003). Cultural group selection arguments often do not sufficiently take into account the fact that conflicts tend to be extremely costly, even for the winning party (for a similar argument in chimpanzee intergroup conflicts, see Wilson, Hauser, & Wrangham, 2001). A group that can make another group adopt its cultural traits by way of persuasion and voluntary imitation will be more efficient than one that has to impose its cultural traits by force, both because of the costs of conflict and because voluntarily adopted traits will maintain themselves through socialization, whereas imposed traits may need costly policing.

This interpretation of cultural group selection in terms of differential cooperativeness fits the fact that we found strong evidence for conditionally cooperative behavior, even in our outgroup conditions. However, by analyzing group composition and conditional cooperation simultaneously, we were able to show in a series of multilevel regression analyses that almost half of the effect of the ingroup conditions on contribution levels was indirect, in the sense that subjects had higher expectations regarding fellow ingroup members' contribution and, consequently, contributed more themselves. The remaining direct effect of the ingroup conditions disappeared entirely when we introduced an interaction term capturing the fact that people condition their own contributions more strongly on those of ingroup members. All in all, then, our analysis thus revealed no evidence for a direct ingroup bias (and

even less for outgroup hostility). Indirectly, however, the ingroup bias is caused by higher expectations of ingroup members and by the fact that participants contribute more in ingroups than in outgroups when they have high expectations of ingroup members' contributions.

Why do people condition their contributions more strongly on their expectations in the case of ingroups? The reason cannot be greater trust, because we show that—at higher expectation levels—subjects contribute more to ingroups while we hold the expected contribution level (the best indicator for trust in other group members) constant. We suggest two plausible interpretations for this interaction effect. A first possible explanation is that people are more familiar with their ingroup than with outgroups. An affiliate of the Reformed churches knows better how his average fellow church member "ticks" than how an average Catholic or Muslim thinks and behaves. As a result, the degree of certainty with which subjects estimate other people's contributions is likely to be greater for ingroups than for outgroups. Assuming that given uncertainty about the estimated contributions of others, people following a conditionally cooperative strategy prefer to stay on the safe side, it follows that people will more closely match the (more certain) estimates of others' contributions in ingroups than they will the (less certain) estimates they make of outgroup members' behavior.

This, however, does not yet explain why, in interactions with ingroup members, participants contribute even more than they expect others to and why, at low expectation levels, they tend to react more severely by withholding their contributions than they do in outgroup conditions. This pattern can be interpreted within the theoretical framework of "strong reciprocity" (Bowles & Gintis, 2004b; Fehr, Fischbacher, & Gächter, 2004), which argues that human group action is enabled by a tendency to punish defectors, and reward cooperators, even at a cost to oneself. Indeed, in dealing with ingroups, our subjects seem to display such a pattern, by "rewarding" group members by contributing more, and "punishing" them by contributing less than they expect these other group members to contribute. If this interpretation is correct, behavior in ingroups is not so much characterized by a generally favoritist bias, but is more strongly normative, which entails not only more prosocial reactions to others' prosociality but also more severe reactions when fellow group members do not behave in the group interest. From an individual-level selection point of view, directed rewarding and punishing of co-operators and free riders, respectively, can indeed explain why individuals are more likely to contribute to the public good. However, in line with Richerson and Boyd (2005, pp. 100–101), we argue that a process of cultural group selection is needed to explain why individuals are willing to make costs in order to improve the group's payoff (in this specific case, by rewarding cooperation) and why these costs are made to reward cooperation and not another behavioral pattern. Therefore, we think that, although we did not find a direct ingroup effect after controlling for expectations and the interaction between expectations and ingroups, our findings can only be explained by a cultural group selection argument. To our knowledge, no mathematical simulation models on the evolution of the found relation between expectations, contributions, and the ingroup–outgroup differentiation exist. We hope that our findings can motivate models in this direction.

In our study, the differences found across the ingroup, outgroup, and random conditions were, although often significant, relatively small in magnitude. It is possible, however, that our realistic research design has partly mitigated group composition effects. To begin with, we added a random component to all our group conditions. This made the stimuli more realistic but the groups also less homogenous. For instance, an ingroup condition consisted of four to five ingroup members (including the subject) but also one to two members of other groups. In addition, we composed ingroups and outgroups always along only one of the three traits. Thus, subjects confronted with a political ingroup were simultaneously confronted with a randomly composed religious group, which might contain not only a few religious ingroup members but also several outgroup members. Thus, our design entails so-called "cross-cutting cleavages," which have been identified in political science (e.g., Almond & Verba, 1963; Lipset, 1960) and sociology (e.g., Coser, 1956; Oberschall, 1973) as mitigating influences on group mobilization. "Cross-categorization" has also been shown to weaken ingroup bias in experimental research (Brown & Turner, 1979; Deschamps & Doise, 1978; Hartstone & Augoustinos, 1995). In follow-up experiments, we want to check to what extent we get stronger effects of group composition on contributions to the public good if we remove cross-cutting cleavages from the design and give subjects information on only one trait of other group members simultaneously. In cross-national experiments, we additionally want to investigate whether our results hold when we compose groups according to other, perhaps more powerful cultural traits, such as language or nationality.

A further interesting avenue that can be explored in future research is how the different patterns of conditional cooperation across ingroups, outgroups, and randomly composed groups affect the stability of collective action over time. Several studies have shown that levels of contribution to public goods games decay over time. One proposed cause for this finding is that subjects react to free riding tendencies by reducing their own contributions, which, again, leads others to reduce their contributions, and so forth (Kollock, 1998). The strong reciprocity that characterizes ingroup interactions may make it easier to sustain high levels of collective action, because the tendency to overcompensate others' contributions may help to counteract downward defection spirals (for a related argument, see Carpenter, 2004). Such a mechanism would have a similar function as the "generous" variants of conditional cooperation that have been shown to be superior

strategies in dyadic cooperation environments under conditions of uncertainty and perception errors (Bendor, 1987, 1993; Wu & Axelrod, 1995).

The research design can also be expanded in the direction of actually competing groups. In the study reported in this article, public goods games were played in single groups, which were varyingly composed of ingroup, outgroup, and random group members, but the design included no real competition among these three types of groups. In future experiments, we will introduce a group contest component, in which the payoffs of the members of the one group partly depend on how much money it is able to mobilize on its group account in comparison to the other group (so-called intergroup public goods games, see Bornstein & Ben-Yossef, 1994; Bornstein, Gneezy, & Nagel, 2002). By varying the composition of the two groups along similar lines as we have done in this study, we hope to find out how conditional cooperation and group-specific strategies interact in situations where there is actual group competition for scarce resources.

Acknowledgments

We thank Jutta Höhne for her assistance in performing multilevel regressions and Jacquelien van Stekelenburg for her comments and suggestions. We also thank two anonymous reviewers for very helpful comments on an earlier version of this article.

Appendix A. Supplementary data

Supplementary data associated with this article can be found, in the online version, at doi:10.1016/j.evolhumbehav.2009.01.003.

References

Allport, G. W. (1954). *The nature of prejudice.* Cambridge, MA: Addison-Wesley.

Almond, G., & Verba, S. (1963). *The civic culture.* Princeton: Princeton University Press.

Amichai-Hamburger, Y. (2005). Internet minimal group paradigm. *CyberPsychology & Behavior, 8*, 140–142.

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

Bendor, J. (1987). In good times and bad: Reciprocity in an uncertain world. *American Journal of Political Science, 31*, 531–558.

Bendor, J. (1993). Uncertainty and the evolution of cooperation. *Journal of Conflict Resolution, 37*, 709–734.

Bernhard, H., Fischbacher, U., & Fehr, E. (2006). Parochial altruism in humans. *Nature, 442*, 912–915.

Bornstein, G., & Ben-Yossef, M. (1994). Cooperation in intergroup and single-group social dilemmas. *Journal of Experimental Social Psychology, 30*, 52–67.

Bornstein, G., Gneezy, U., & Nagel, R. (2002). The effect of intergroup competition on group coordination: An experimental study. *Games and Economic Behavior, 41*, 1–25.

Bowles, S., & Choi, J. K. (2003). The co-evolution of love and hate. In P. Van Parijs (Ed.), *Cultural diversity versus economic solidarity. Proceedings of the seventh francqui colloquium, Brussels, 28 February–1 March 2003* (pp. 81–97). Brussels: De Boeck.

Bowles, S., & Gintis, H. (2004a). Persistent parochialism: Trust and exclusion in ethnic networks. *Journal of Economic Behavior & Organization, 55*, 1–23.

Bowles, S., & Gintis, H. (2004b). The evolution of strong reciprocity: Cooperation in heterogeneous populations. *Theoretical Population Biology, 65*, 17–28.

Boyd, R., & Richerson, P. J. (1982). Cultural transmission and the evolution of cooperative behavior. *Human Ecology, 10*, 325–351.

Boyd, R., & Richerson, P. J. (1985). *Culture and the evolutionary process.* Chicago: The University of Chicago Press.

Boyd, R., & Richerson, P. J. (1988). The evolution of reciprocity in sizable groups. *Journal of Theoretical Biology, 132*, 337–356.

Brandts, J., & Charness, G. (2000). Hot vs. cold: Sequential responses and preference stability in experimental games. *Experimental Economics, 2*, 227–238.

Brewer, M. B. (1999). The psychology of prejudice: Ingroup love or outgroup hate? *Journal of Social Issues, 55*, 429–444.

Brown, R. J., & Turner, J. C. (1979). The criss-cross categorization effect in the intergroup discrimination. *British Journal of Social and Clinical Psychology, 18*, 371–383.

Carpenter, J. P. (2004). When in Rome: Conformity and the provision of public goods. *Journal of Socio-Economics, 33*, 395–408.

Choi, J. -K., & Bowles, S. (2007). The coevolution of parochial altruism and war. *Science, 318*, 636–640.

Cook, K. S., & Yamagishi, T. (2008). A defense of deception on scientific grounds. *Social Psychology Quarterly, 71*, 215–221.

Coser, L. A. (1956). *The functions of social conflict.* Glencoe, IL: Free Press.

Dawes, R. M. (1980). Social dilemmas. *Annual Review of Psychology, 31*, 169–193.

Dawes, R. M., McTavish, J., & Shaklee, H. (1977). Behavior, communication, and assumptions about other people's behavior in a commons dilemma situation. *Journal of Personality and Social Psychology, 35*, 1–11.

Deschamps, J. C., & Doise, W. (1978). Crossed category memberships in intergroup relations. In H. Tajfel (Ed.), *Differentiation between social groups.* London: Academic Press.

Falk, A., & Fischbacher, U. (2002). "Crime" in the lab—Detecting social interaction. *European Economic Review, 46*, 859–869.

Fehr, E., Fischbacher, U., & Gächter, S. (2004). Strong reciprocity, human cooperation, and the enforcement of social norms. *Human Nature, 13*, 1–25.

Fischbacher, U., Gächter, S., & Fehr, E. (2001). Are people conditionally cooperative? Evidence from a public goods experiment. *Economics Letters, 71*, 397–404.

Gil-White, F. J. (2003). Ultimatum game with an ethnicity manipulation: Results from Khovdiin Bulgan Sum, Mongolia. In J. Henrich, R. Boyd, S. Bowles, H. Gintis, E. Fehr, & C. Camerer (Eds.), *Foundations of human sociality: Ethnography and experiments in 15 small-scale societies.* Oxford: Oxford University Press.

Gil-White, F. J., & Richerson, P. (2003). Large-scale human cooperation and conflict. In L. Nadel (Ed.), *Encyclopedia of cognitive science* London: Nature Publishing/MacMillan.

Goette, L., Huffman, D., & Meier, S. (2006). The impact of group membership on cooperation and norm enforcement: Evidence using random assignment to real social groups. *American Economic Review, 96*, 212–216.

Halevy, N., Bornstein, G., & Sagiv, L. (2008). "In-group love" and "out-group hate" as motives for individual participation in intergroup conflict. A new game paradigm. *Psychological Science, 19*, 405–411.

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

Hartstone, M., & Augoustinos, M. (1995). The minimal group paradigm: Categorization into two versus three groups. *European Journal of Social Psychology, 25*, 179–193.

Henrich, J. (2004). Cultural group selection, coevolutionary processes and large-scale cooperation. *Journal of Economic Behavior & Organization, 53*, 3–35.

Kiyonari, T., & Yamagishi, T. (2004). Ingroup cooperation and the social exchange heuristic. In R. Suleiman, D. V. Budescu, J. Fischer, & D. M.

Messick (Eds.), *Contemporary psychological research on social dilemmas.* Cambridge: Cambridge University Press.

Klandermans, B. (1984). Mobilization and participation: Social–psychological expansions of resource mobilization theory. *American Sociological Review, 49*, 583−600.

Klandermans, B. (1997). *The social psychology of protest.* Oxford: Wiley-Blackwell.

Kollock, P. (1998). Social dilemmas: The anatomy of cooperation. *Annual Review of Sociology, 24*, 183−214.

Lipset, S. M. (1960). *Political man: The social bases of politics.* Garden City, NY: Doubleday & Co.

Marwell, G., & Ames, R. E. (1979). Experiments on the provision of public goods. I. Resources, interest, group size, and the free-rider problem. *The American Journal of Sociology, 84*, 1335−1360.

Messick, D. M., & Brewer, M. B. (1983). Solving social dilemmas. In L. Wheeler, & P. Shaver (Eds.), *Review of personality and social psychology No. 4* (pp. 11−44). Beverly Hills: Sage Publications.

Miller, J. H., & Andreoni, J. (1991). Can evolutionary dynamics explain free riding in experiments? *Economics Letters, 36*, 9−15.

Mummendey, A., Simon, B., Dietze, C., Grunert, M., Haeger, G., Kessler, S., et al. (1992). Categorization is not enough: Intergroup discrimination in negative outcome allocations. *Journal of Experimental Social Psychology, 28*, 125−144.

Oberschall, N. (1973). *Social conflict and social movements.* Englewood Cliffs: Prentice-Hall, Inc.

Olson, M. (1965). *The logic of collective action. Public goods and the theory of groups.* Cambridge, MA: Harvard University Press.

Raudenbush, S. W., & Bryk, A. S. (2002). *Hierarchical linear models. Applications and data analysis methods.* Number 1 in advanced quantitative techniques in the social sciences. Thousand Oaks, CA.

Richerson, P., & Boyd, R. (2005). *Not by genes alone.* How culture transformed human evolution. Chicago.

Tajfel, H. (1982). Social psychology of intergroup relations. *Annual Review of Psychology, 33*, 1−39.

Tajfel, H., Billig, M. G., Bundy, R. P., & Flament, C. (1971). Social categorization and intergroup behavior. *European Journal of Social Psychology, 1*, 149−178.

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

Van Vugt, M., De Cremer, D., & Janssen, D. P. (2007). Gender differences in cooperation and competition. The male-warrior hypothesis. *Psychological Science, 18*, 19−23.

Vaughan, G. M., Tajfel, H., & Williams, J. (1981). Bias in reward allocation in an intergroup and an interpersonal context. *Social Psychology Quarterly, 44*, 37−42.

Wilson, D. S. (2002). *Darwin's cathedral.* Chicago: The University of Chicago Press.

Wilson, D. S., & Sober, E. (1994). Reintroducing group selection to the human behavioral sciences. *Behavioral and Brain Sciences, 17*, 585−608.

Wilson, M. L., Hauser, M. D., & Wrangham, R. W. (2001). Does participation in intergroup conflict depend on numerical assessment, range location, or rank for wild chimpanzees? *Animal Behaviour, 61*, 1203−1216.

Wu, J., & Axelrod, R. (1995). How to cope with noise in the repeated prisoner's dilemma. *Journal of Conflict Resolution, 39*, 183−189.

Yamagishi, T., Jin, N., & Kiyonari, T. (1999). Bounded generalized reciprocity. Ingroup boasting and ingroup favoritism. *Advances in Group Processes, 16*, 161−197.

Yamagishi, T., & Kiyonari, T. (2000). The group as container of generalized reciprocity. *Social Psychology Quarterly, 63*, 116−132.

Yamagishi, T., Makimura, Y., Foddy, M., Matsuda, M., Kiyonari, T., & Platow, M. J. (2005). Comparisons of Australians and Japanese on group-based cooperation. *Asian Journal of Social Psychology, 8*, 173−190.

Yamagishi, T., & Sato, K. (1986). Motivational bases of the public goods problem. *Journal of Personality and Social Psychology, 50*, 67−73.