

Rejection of unfair offers in the ultimatum game is no evidence of strong reciprocity

Toshio Yamagishi^{a,1}, Yutaka Horita^b, Nobuhiro Mifune^c, Hirofumi Hashimoto^d, Yang Li^e, Mizuho Shinada^d, Arisa Miura^e, Keigo Inukai^f, Haruto Takagishi^g, and Dora Simunovic^e

^aBrain Science Institute, Tamagawa University, Machida, Tokyo 194-8610, Japan; ^bDepartment of Psychology, Sophia University, Chiyoda-ku, Tokyo 102-0094, Japan; ^cDepartment of Psychology, Kobe University, Nada, Kobe, Hyogo 657-0013, Japan; ^dDepartment of Social Psychology, Graduate School of Humanities and Sociology, The University of Tokyo, Bunkyo-ku, Tokyo 113-0033, Japan; ^eDepartment of Behavioral Science, Hokkaido University, Sapporo, Hokkaido 060-0810, Japan; ^fDepartment of Economics, Hokkaido University, Sapporo 060-0810, Japan; and ^gDepartment of Child Neuropsychiatry, Graduate School of Medicine, The University of Tokyo, Bunkyo-ku, Tokyo 113-8655, Japan

Edited* by Karen S. Cook, Stanford University, Stanford, CA, and approved November 1, 2012 (received for review July 21, 2012)

The strong reciprocity model of the evolution of human cooperation has gained some acceptance, partly on the basis of support from experimental findings. The observation that unfair offers in the ultimatum game are frequently rejected constitutes an important piece of the experimental evidence for strong reciprocity. In the present study, we have challenged the idea that the rejection response in the ultimatum game provides evidence of the assumption held by strong reciprocity theorists that negative reciprocity observed in the ultimatum game is inseparably related to positive reciprocity as the two sides of a preference for fairness. The prediction of an inseparable relationship between positive and negative reciprocity was rejected on the basis of the results of a series of experiments that we conducted using the ultimatum game, the dictator game, the trust game, and the prisoner's dilemma game. We did not find any correlation between the participants' tendencies to reject unfair offers in the ultimatum game and their tendencies to exhibit various prosocial behaviors in the other games, including their inclinations to positively reciprocate in the trust game. The participants' responses to postexperimental questions add support to the view that the rejection of unfair offers in the ultimatum game is a tacit strategy for avoiding the imposition of an inferior status.

inequity aversion | other regarding preferences | assertiveness

Whether cooperation in human society can be explained by reputation-based reciprocity alone or whether strong reciprocity (1–4) is also required to explain cooperation is a currently disputed issue in attempts to solve the puzzle of cooperation in human societies (5). The main issue at stake is whether individual propensities to punish noncooperators are necessary to support cooperation beyond the level afforded by the inclinations of individuals to care about their reputations as good members of their community. The strong reciprocity model of cooperation was proposed as a means of overcoming the limitations of reputation-based reciprocity models that only explain the evolution of cooperation over a narrow range of conditions, namely, among societies that are characterized by the presence of long-lasting small groups in which the behavioral histories of the individual group members are transparent to the other members. Strong reciprocity theorists argue that the limitations of reputation-based reciprocity models can be overcome by assuming that strong reciprocators who stabilize cooperation by punishing noncooperators are present within a given community. A strong reciprocator is defined as an individual who is willing to “sacrifice resources for rewarding fair and punishing unfair behavior *even if this is costly and provides neither present nor future material rewards for the reciprocator*” (3). In other words, strong reciprocators reciprocate both positively and negatively (3, 6)—positive reciprocity promotes cooperation, and negative reciprocity stabilizes it. Although the current study does not directly address the issue of whether strong reciprocity is needed to explain the evolution of cooperation, it addresses this issue indirectly by examining the legitimacy of

using findings from ultimatum game experiments as evidence of strong reciprocity.

Findings from economic game experiments, particularly findings regarding the rejection of unfair offers in the ultimatum game (7), have provided the major impetus for the development of the strong reciprocity theory of cooperation (3, 8, 9). The ultimatum game (7) (UG) is an economic game that is played by two individuals, a proposer and a responder. The proposer is given a certain amount of money from the experimenter and is asked to propose a way of dividing the money between himself and the responder. The responder then has to choose to either accept or reject the offer. If the responder accepts the offer, both the proposer and the responder earn money as specified by the proposer. However, if the responder rejects the offer, neither player receives any money. The game is usually played with complete anonymity, and it is generally played between strangers. Under these conditions, rational and self-regarding proposers should offer as little as possible to the responder, and responders who are similarly rational and self-regarding should accept any nonzero offers. However, the findings of UG experiments do not support this prediction. In many UG experiments, the most common offer among the proposers is a 50–50 split, and approximately half of the responders reject unfair offers in which they would receive less than 30% of the total sum (10). The rejection of unfair offers that is frequently observed in UG experiments has been regarded as evidence of strong reciprocity that is driven by a preference for reciprocal fairness and inequity aversion (11) on the part of the responder.

In the present study, we argue and provide experimental evidence that interpreting the observed rejection of unfair offers in a UG experiment as a form of strong reciprocity driven by a preference for fairness may not be as straightforward an interpretation as has been assumed. As stated in the aforementioned definition (3), a strong reciprocator should exhibit both positive and negative reciprocity. It is generally assumed that the behavioral tendencies toward both positive and negative reciprocity are driven by both a preference for reciprocal fairness (12) and inequity aversion (6). A reciprocally fair individual is motivated to respond in kind—he will respond to kindness with kindness and will meet hostility with hostility. Similarly, an inequity-averse individual is motivated to avoid inequity and to implement equitable outcomes. Regardless of the specific motivation

Author contributions: T.Y. designed research; Y.H., N.M., H.H., Y.L., M.S., A.M., K.I., H.T., and D.S. performed research; T.Y., Y.H., N.M., and H.H. analyzed data; and T.Y., Y.H., N.M., and D.S. wrote the paper.

The authors declare no conflict of interest.

*This Direct Submission article had a prearranged editor.

Freely available online through the PNAS open access option.

¹To whom correspondence should be addressed. E-mail: yamagishitoshio@gmail.com.

This article contains supporting information online at www.pnas.org/lookup/suppl/doi:10.1073/pnas.1212126109/-DCSupplemental.

from which an act of strong reciprocity is derived, a strong reciprocator who rejects an unfair offer in the UG (an example of negative reciprocity) should act in a positively reciprocal manner in other situations insofar as his or her inclination to act in accordance with strong reciprocity is driven by a relatively stable set of social preferences.

We challenge the interpretation of the rejection of unfair offers in the UG as a form of strong reciprocity by determining whether individuals who reject unfair offers actually do behave in a positively reciprocal manner in other situations. Specifically, we will first demonstrate that the rejection of unfair offers in the UG (hereafter referred to as the rUG) is not correlated with the degree to which positive reciprocity is exhibited by the same individual when playing the role of trustee in the trust game. Furthermore, we will demonstrate that the rUG is not correlated with either the prosocial dispositions or prosocial value orientations (SVOs) (13, 14) of the responders, nor with their prosocial behaviors in other games such as cooperation in the prisoner's dilemma game ($cPDG$), trust as a truster (tTG), or allocation of resources to a recipient as a dictator in the dictator game (aDG). These findings will provide evidence against the strong-reciprocity-theory-based interpretation of the rUG as a reflection of a disposition toward practicing both positive and negative reciprocity on the part of the responder.

If not based on a social preference for reciprocal fairness or inequity aversion, where does rUG come from? We argue that the rUG is a psychological response to a challenge to the integrity or status of the responder. This alternative interpretation of the rUG , called the "wounded pride hypothesis" by Straub and Murnighan (15), received partial support from studies demonstrating that individuals who rejected unfair offers had higher levels of testosterone than individuals who accepted such offers (16). Other studies found that administering testosterone increases rUG among men (17), although not among women (18). Xiao and Houser (19) provided further evidence for this interpretation by showing that their participants used an option to express their anger to the proposer instead of rejecting an unfair offer as a means to assert that they were not willing to accept an inferior position. Another piece of evidence that supports the account of unfair offers as a means of asserting strength came from a study of the impunity game that was conducted by Yamagishi et al. (20). The impunity game (21) is similar to the ultimatum game, but the amount of money proposed to a responder is reduced to zero upon her rejection of an unfair offer, whereas the proposer keeps his money intact. Thus, in the impunity game, a responder who is given an unfair offer can neither achieve equality nor punish the unfair proposer by rejecting the offer because the proposer keeps his money regardless of the responder's decision. Instead, rejecting an unfair offer in the impunity game causes the outcome of the distribution to be even more unfair to the responder. Yamagishi et al. (20) found that responders in the impunity game (or in the private impunity game in which the proposer was not informed of the responder's choice) rejected 30–40% of the unfair offers.

According to this alternative interpretation (22), the rUG is not a prosocial behavior oriented toward restoring or enforcing a fair outcome distribution. Thus, it leads to a prediction that an individual's tendency toward the rUG will not be correlated with her prosocial behaviors in other games. The second goal of the present study is to provide further support for this alternative account of the rUG by demonstrating the lack of correlation between the rUG and the aforementioned prosocial behaviors in other games. We also sought to provide evidence that there was a positive relation between the rUG and a psychological scale that we constructed to measure the degree to which a player was unwilling to submit to other players, whereas there was no relation between a player's rUG and her concern for positive or negative reciprocity.

Predictions Based on Strong Reciprocity. The logic by which the rUG is interpreted as a form of strong reciprocity suggests that the player's rUG should be positively correlated with his prosocial choices in the other games. That is, if the rUG is a product of a social preference for fairness, the tendency to reject unfair offers in the UG should be positively correlated with other game behaviors that are caused by the same preference for fairness. Specifically, a player's rUG will be positively correlated with the level of positive reciprocity exhibited by the trustee to the trusting action taken by the truster. Furthermore, the player's rUG is expected to be positively correlated with both his aDG and tTG (the proportion of the entrusted money returned by the trustee) because these behaviors are expected to reflect the player's preference for fairness. Similarly, the rUG is predicted to be positively correlated with the tTG because the trusting choice in the TG increases the total amount of money that will be shared between the two players. Finally, the level of $cPDG$ that a player exhibits, especially in the one-shot version of the game, is known to be correlated with the prosocial preference (23) that includes a preference for the fair distribution of outcomes (24). Thus, according to the strong reciprocity interpretation of the rUG , the rUG is predicted to be positively correlated with $cPDG$.

Predictions Based on Assertiveness. If the interpretation that rUG is a means of asserting strength is correct, no correlation between a player's rUG and the other aforementioned game behaviors is predicted because challenges to a player's determination to avoid being subjugated by the other player are irrelevant in games that do not involve threats to the status of either player. For example, the dictator in the dictator game has total control of the outcome distribution, so he will not be threatened by any challenger. Similarly, neither the trusters nor the trustees (toward whom the truster has shown trust) in the trust game face any threat to their integrity. A negative correlation between a player's rUG and $cPDG$ is predicted on the basis of experimental evidence that suggests that defectors in the PDG interpret cooperation as a sign of weakness, whereas cooperators interpret cooperation as a sign of moral quality (25). This finding suggests that some PDG players, especially defectors, perceive the game in terms of a status competition. PDG players who see defection as a reflection of their strength and defect would also tend to reject unfair offers in the UG to show their strength. Thus, a player's rUG will be negatively correlated with his $cPDG$.

Results

Hypotheses Testing. Fig. 1 depicts the presence of positive reciprocity among the responses of the responders in the trust game. That is, in general, the trustee returned a greater proportion of the entrusted money when the truster handed her a larger amount of money. An analysis of variance of the returned proportion revealed a significant main effect of the entrusted amount [$F(5, 460) = 23.44, P < 0.001$]. We constructed a measure of positive reciprocity by subtracting the average proportion of the returned money for the two lowest values of entrusted money (100 yen and 200 yen) from the average proportion of returned money for the two highest amounts of entrusted money (500 yen and 600 yen). We found no significant correlation between this measure of positive reciprocity and the rUG ($r = 0.07$) and rejected the strong reciprocity prediction.

We report both Pearson and Spearman correlations between the rUG and the behaviors in the other games that are relevant to testing the major hypotheses of this article in Table 1. Correlations among the various game behaviors are provided in *SI Appendix, Table S1*. No significant correlations between rUG and any of the other game behaviors were found, with the exception of a significant *negative* correlation between the rUG and the $cPDG$. We further examined possibilities of nonlinear relationships and found no significant relationships between rUG and

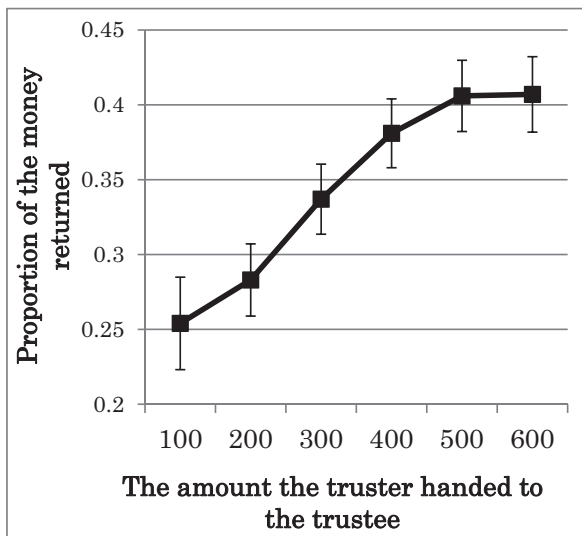


Fig. 1. The relationship between the entrusted amount and the proportion of money that was returned in the trust game. Error bars represent standard errors.

the other game behaviors. Scattergrams for the relationships are provided in *SI Appendix*. These findings consistently reject the predictions of the strong reciprocity hypothesis and provide support for the predictions that were derived from the assertiveness account of the rejection behavior.

Social Value Orientation. The prosociality and proselfness measures of SVO were negatively correlated ($r = -0.75, P < 0.001$), so we decided to use the difference between the prosociality measure and the proselfness measure as an overall measure of prosociality. The competitiveness measure was highly skewed—92% of the participants were never classified as being competitive by any of the four SVO measures—and it did not correlate with rUG ($r = -0.02$). Because so few participants were scored as being competitive, we decided to exclude the competitiveness measure from the analysis. Table 1 lists the correlations between the overall measure of prosociality and various game behaviors. With the exception of the rUG and aDG , all of the game behavior measures were positively correlated with the overall SVO measure, which suggests that the game behaviors of the participants were guided by their prosocial/proself preferences. However, the correlation between the prosocial SVO and rUG was *negative*, which implies that the rejection of unfair offers in the UG is difficult to justify as being “altruistic.”

Reciprocity and Assertiveness. The positive reciprocity scale and the negative reciprocity scales were not significantly correlated with each other ($r = -0.07$), which is consistent with the findings that have been presented in the preceding paragraphs. The vengefulness scale scores were positively correlated with the negative reciprocity scores ($r = 0.66, P < 0.001$) but not with the positive reciprocity scores ($r = -0.03$). Furthermore, no significant correlations between rUG and any of these three reciprocity scales were identified ($r = 0.18$ with vengefulness, -0.07 with positive reciprocity, and 0.09 with negative reciprocity). A significant correlation was found between the assertiveness scale and rUG ($r = 0.36, P < 0.001$), whereas the assertiveness scale was not correlated with any of the three reciprocity scales or other game behaviors. These findings suggest that the rUG does not reflect either vengeance or negative reciprocity per se; rather, it is more likely a reflection of a participant’s aversion to being regarded as someone who lacks the strength to assert himself in the face of the proposer.

Discussion

The findings of the present experiment do not support the view that the tendency to reject unfair offers in the UG represents strong reciprocity. A strong reciprocator is characterized as having a disposition toward both positive *and* negative reciprocity, so it follows that strong reciprocators who reject unfair offers in the UG should behave in a fair manner and should reciprocate positively in other games. This prediction was clearly rejected by our findings.

If the rejection of unfair offers in the UG is not an altruistic punishment of norm-violating behavior, what is it? Our answer to this question is that it reflects a tendency on the part of the responder to avoid being subjugated to the proposer. This alternative account of the rejection of unfair offers in the UG suggests that the rUG should be either unrelated or negatively related to the other game behaviors and to the prosociality SVO measure. The results of our experiments consistently supported the predictions of this alternative account. First, the rejection of unfair offers in the UG was *negatively* correlated with our measure of prosocial SVO. Second, the responder’s decision to reject unfair offers in the UG was found to be independent of his or her inclination to act in prosocial ways in other behavioral games, with the exception of the tendency to cooperate in the PDG ; there was a *negative* correlation between the rUG and this behavior. Third, the rejection of unfair offers in the UG was not correlated with positive reciprocity in the TG . Fourth, the rejection of unfair offers in the UG was not correlated with either the positive or negative reciprocity scales, nor was it correlated with the general vengefulness scale. Finally, we found a positive correlation between the rejection of unfair offers in the UG and a self-report scale that measured their assertiveness against the intentions of another player to exert control.

Table 1. Correlations of the six game behaviors with the rUG , the prosocial value orientation measure, the positive and negative reciprocity scales, the vengefulness scale, and the assertiveness scale

	<i>n</i>	Mean	SD	Correlations with row variables					
				rUG	oUG	$cPDG$	tTG	rTG	aDG
rUG	86	0.43	0.43	1.00	0.14	-0.27*	-0.17	-0.02	-0.18
rUG (Spearman)	86	—	—	—	0.21	-0.26*	-0.19	-0.03	-0.18
Prosociality	106	0.27	0.60	-0.20	0.21*	0.28**	0.28**	0.30**	0.12
Positive reciprocity	86	5.28	0.68	-0.07	0.02	0.08	0.12	0.11	-0.03
Negative reciprocity	86	3.33	1.19	0.09	-0.21	-0.15	-0.25*	-0.31**	-0.23*
Vengefulness	86	3.89	1.24	0.18	-0.12	-0.16	-0.20	-0.20	-0.13
Assertiveness	105	4.00	0.68	0.36***	0.17	-0.09	0.03	0.07	0.07

* $P < 0.05$, ** $P < 0.01$, *** $P < 0.001$.

The most important implication of the findings described above is that the UG is not an appropriate tool for measuring an individual's disposition toward strong reciprocity. The findings of the present study should encourage strong reciprocity theorists to examine the possibility that there may be similar alternative explanations for other game behaviors that have been regarded as evidence of strong reciprocity more carefully. For example, in a third-party punishment experiment (8), a dictator makes an unfair allocation to another person who is not the participant. The third-party participant knows that the dictator has performed an unfair behavior while being aware that the participant, who has the capability of punishing unfair individuals, was watching him. From the participant's point of view, a dictator who behaves in an unfair manner in this situation is essentially saying "You are a coward, and you won't have the nerve to punish me, ha ha!" As in the UG, the punishment of an unfair dictator in the third-party punishment experiment may result from a participant's inclination to assert his strength and his unwillingness to be treated as a weak and acquiescent person. In short, the third-party participant's response can be considered akin to saying "I'll teach you a lesson— don't take me lightly!"

One important theoretical implication of the present work concerns whether or not the relationship between positive and negative reciprocity is critical for the evolution of strong reciprocity. Evolutional models of positive reciprocity have been proposed on the basis of the future reputational gains of positive reciprocal behavior (26–28). The unique feature of the strong reciprocity model is that it explains the evolution of reciprocity in a manner that does not require a long time horizon. This model of the evolution of reciprocity, however, requires that positive and negative reciprocity are inextricably linked within a given individual (2). Despite the importance of this requirement for the strong reciprocity account of the evolution of reciprocity and cooperation, the relationship between positive and negative reciprocity has not been the target of systematic investigation. In social dilemma experiments in which participants have the option of punishing defectors, it has been shown that cooperators are likely to punish noncooperators (29–31). However, these findings may not provide clear evidence of the positive relationship between the two types of reciprocity because the cooperation–punishment relationship that is often observed in social dilemma experiments may be a product of frustration aggression instead of an expression of a fairness-related preference. More rigorous tests of the consistencies of both positive and negative reciprocity and the link between them can be conducted via a strategy that is similar to the one adopted by the current study: the use of across-game behavioral consistency. These types of rigorous tests will provide a firmer empirical foundation for the strong reciprocity model. One possible drawback of the current methodology could come from influences of various factors that vary with time for each individual as well as social and individual events that take place between games. Given the relatively strong correlations between game behaviors that took place with long intervals, for example, $r = 0.55$ for cPDG and tTG conducted almost 3 y apart (*SI Appendix*), it is difficult to attribute the lack of correlations between rUG and the other game behaviors to the influence of such factors and events, at least in the current study.

Methods

Sample and Data Collection. The study presented in this paper is part of a larger research project that was initiated in February 2008 and concluded in November 2011. During the three-and-a-half-year duration of the study, the same set of participants was asked to play a variety of experimental games, including the games we tested in this paper. We inserted long time intervals between different experimental games to minimize possible carry-over effects. An initial sample that included 108 participants was recruited from the population of nonstudent residents of Sapporo via newspaper advertisements. We later added a second sample of an additional 108 participants. In the present article, we only report findings from the first sample

because the participants in the second sample did not play the ultimatum game. Approximately equal numbers of male (51) and female (57) participants were included, and their ages were approximately evenly distributed within the age range of 21–69 y (as of December 2007). The research presented in this paper was approved by the ethics committee of the Center for Experimental Studies in Social Sciences, Hokkaido University. All participants read and signed the consent form each time they participated in the study. A more detailed description of this project is presented in *SI Appendix*.

Games. Although the same set of individuals participated in all four games, some participants failed to participate in some games. Detailed descriptions of the procedures for all of the games are presented in *SI Appendix*.

Ultimatum game. The proposer proposed a division of 1,500 yen ($\$1 \approx 80$ yen) between the two players, and the responder decided whether to accept or reject the proposed division. If the responder accepted the proposed division, each participant would earn the money that had been specified in the proposed division of the 1,500 yen. However, if the responder rejected the proposal, neither participant would earn any money. At the start of the experiment, all of the participants were assigned the role of responder and were asked to decide whether to accept or reject each of the offers made by the proposers with whom they were partnered using a strategy method. Specifically, the participants were given a booklet that contained several pages, each of which described a particular proposed division of the 1,500 yen. The proposed divisions ranged from 100 yen to the recipient and 1,400 yen to the proposer to 1,400 yen to the recipient and 100 yen to the proposer, and the order of the pages in each booklet was randomized. After the participants had finished indicating their responses as recipients, they were told that they would play the same game with a different participant. However, the participants played as the proposers in this round, and they were therefore instructed to propose divisions of 1,500 yen to a randomly matched recipient. In the analysis, we use the relative frequency with which the recipients rejected offers of 100, 200, 300, 400, and 500 yen (7–33% of the total money) as a measure of the rUG. The percentages of offer rejections were 53, 53, 49, 40, 21, 14, 2, 2, 3, 6, 5, 6, 6, and 7% for the offers from 1,400/100 through 100/1,400.

Dictator game. All of the participants first played the role of dictator, in which they were instructed to decide how much of a sum of 2,500 yen that had been provided by the experimenter to keep for themselves and how much to give to a randomly matched recipient. Before the participants made their decisions, they were informed that a lottery would be used to randomly match them with the other participants and to randomly assign the two roles (dictator and recipient) after they had made their decisions. After the random assignments of both partners and roles, the participants who had been randomly selected as recipients (half of the participants) received the amount that the dictator with whom they were matched had allocated to a recipient. The remaining half of the participants received the shares that they allocated to themselves.

Trust game. The trust game we used took the form of an investment game (32). All of the participants were first assigned the role of truster, and they were each given 600 yen by the experimenter. Then, the participants decided upon a portion of this amount to give to a matched responder, who would then receive three times the allotted money and would be given the opportunity to decide upon a method of dividing the tripled amount between herself and the truster. After the participants made this trusting decision, they were given instructions for a second game in which they were partnered with a different participant with whom they would play the role of trustee. The participants were asked to determine the amount of money from the allotment that they received (after it was tripled) that they would give to the trusters using the strategy method. That is, they were asked to indicate the amount of money that should be returned to the truster when the tripled value of the allotted money was 300, 600, 900, 1,200, 1,500, and 1,800 yen. We used the average proportion of the money that the trustee returned to the truster when she was trusted with totals of 1,200, 1,500, and 1,800 yen (after tripling).

Prisoner's dilemma game. The version of the PDG that was used in the present study involved the use of an exchange format. Each participant was given an endowment of 500 yen and was asked to decide how much of it to give to his or her randomly matched partner. The money that had been provided by the participant was doubled in value and was then given to his or her partner. Each participant played the game three times and was given a different partner for each game. The participants earned the total sum of money that they had gained during the three games. No feedback regarding the outcome of any individual game was provided. We used the average of the three amounts that the participants had allocated to their partners in the three games as a measure of their cPDG.

Social Value Orientation. During the three-and-a-half-year period over which the culture and game study was conducted, the social value orientations of the participants were measured four times; two of these measures used the triple dominance method (13), and two of them used the ring method (14). Each individual measure was used to classify each participant as a cooperater, a competitor, or an individualist, and the proportions in which the participants were assigned to each classification were used as measures of prosociality, competitiveness, and proselfness, respectively. Participants for whom any type of SVO measure was completely missing were dropped from the analysis.

Reciprocity and Assertiveness. We administered the Personal Norm of Reciprocity scale developed by Perugini et al. (33), which was composed of positive and negative reciprocity subscales. In addition, we included a third scale that

was designed to measure the level of vengefulness of each participant. Finally, we constructed and administered an attitude scale that was designed to measure the degree to which a participant was inclined to assert himself when other people attempted to impose their influence. The items that were included in the vengefulness and assertiveness scales are presented in *SI Appendix, Tables S2 and S3*. The vengefulness measure was administered during the experimental wave in which participants played the UG (August 2010), and the remaining scales were administered approximately 2 y before or after participation in the UG.

ACKNOWLEDGMENTS. The research reported in this study was funded by Grant-in-Aid for Scientific Research 19046005 from the Japan Society for the Promotion of Science.

- Boyd R, Gintis H, Bowles S, Richerson PJ (2003) The evolution of altruistic punishment. *Proc Natl Acad Sci USA* 100(6):3531–3535.
- Bowles S, Gintis H (2004) The evolution of strong reciprocity: Cooperation in heterogeneous populations. *Theor Popul Biol* 65(1):17–28.
- Fehr E, Fischbacher R, Gächter S (2002) Strong reciprocity, human cooperation and the enforcement of social norms. *Hum Nat* 13:1–25.
- Gintis H (2000) Strong reciprocity and human sociality. *J Theor Biol* 206(2):169–179.
- Guala F (2012) Reciprocity: Weak or strong? What punishment experiments do (and do not) demonstrate. *Behav Brain Sci* 35(1):1–15.
- Fehr E, Schmidt KM (1999) A theory of fairness, competition, and cooperation. *Q J Econ* 114(3):817–868.
- Güth W, Schmittberger R, Schwartz B (1982) An experimental analysis of ultimatum bargaining. *J Econ Behav Organ* 3(4):367–388.
- Fehr E, Fischbacher U (2003) The nature of human altruism. *Nature* 425(6960):785–791.
- Gintis H, Boyd R, Bowles S, Fehr E (2003) Explaining altruistic behavior in humans. *Evol Hum Behav* 24:153–172.
- Camerer C (2003) *Behavioral Game Theory: Experiments in Strategic Interaction* (Princeton Univ Press, Princeton).
- Fehr E, Gintis H (2007) Human motivation and social cooperation: Experimental and analytical foundations. *Annu Rev Sociol* 33:43–64.
- Rabin M (1993) Incorporating fairness into game theory and economics. *Am Econ Rev* 83(5):1281–1302.
- Liebrand WBG, McClintock CG (1988) The ring measure of social values: A computerized procedure for assessing individual differences in information processing and social value orientation. *Eur J Pers* 2(3):217–230.
- Van Lange PAM, Otten W, De Bruin EM, Joireman JA (1997) Development of prosocial, individualistic, and competitive orientations: Theory and preliminary evidence. *J Pers Soc Psychol* 73(4):733–746.
- Straub PG, Murnighan JK (1995) An experimental investigation of ultimatum games: Information, fairness, expectations, and lowest acceptable offers. *J Econ Behav Organ* 27(3):345–364.
- Burnham TC (2007) High-testosterone men reject low ultimatum game offers. *Proc Biol Sci* 274(1623):2327–2330.
- Zak PJ, et al. (2009) Testosterone administration decreases generosity in the ultimatum game. *PLoS ONE* 4(12):e8330.
- Eisenegger C, Naef M, Snodgrass R, Heinrichs M, Fehr E (2010) Prejudice and truth about the effect of testosterone on human bargaining behaviour. *Nature* 463(7279):356–359.
- Xiao E, Houser D (2005) Emotion expression in human punishment behavior. *Proc Natl Acad Sci USA* 102(20):7398–7401.
- Yamagishi T, et al. (2009) The private rejection of unfair offers and emotional commitment. *Proc Natl Acad Sci USA* 106(28):11520–11523.
- Bolton GE, Katok E, Zwick R (1998) Dictator game giving: Rules of fairness versus acts of kindness. *Int J Game Theory* 27(2):269–299.
- Frank RH (1988) *Passions Within Reason: The Strategic Role of the Emotions* (W. W. Norton, New York).
- Balliet D, Parks C, Joireman J (2009) Social value orientation and cooperation in social dilemmas: A meta-analysis. *Group Process Intergroup Relat* 12(4):533–547.
- Eek D, Gärling T (2006) Prosocials prefer equal outcomes to maximizing joint outcomes. *Br J Soc Psychol* 45(Pt 2):321–337.
- Liebrand WBG, Jansen RWTL, Ruken WM, Suhre CJM (1986) Might over morality: Social values and the perception of other players in experimental games. *J Exp Soc Psychol* 22:203–215.
- Axelrod R (1984) *The Evolution of Cooperation* (Basic Books, New York).
- Nowak MA, Sigmund K (2005) Evolution of indirect reciprocity. *Nature* 437(7063):1291–1298.
- Trivers RL (1971) The evolution of reciprocal altruism. *Q Rev Biol* 46:35–57.
- Falk A, Fehr E, Fischbacher U (2005) Driving forces of informal sanctions. *Econometrica* 73:2017–2030.
- Fehr E, Gächter S (2000) Cooperation and punishment in public goods experiments. *Am Econ Rev* 90(4):980–994.
- Gürerck O, Irlenbusch B, Rockenbach B (2006) The competitive advantage of sanctioning institutions. *Science* 312(5770):108–111.
- Berg J, Dickhaut J, McCabe K (1995) Trust, reciprocity, and social history. *Games Econ Behav* 10(1):122–142.
- Perugini M, Gallucci M, Presaghi F, Ercolani AP (2003) The personal norm of reciprocity. *Eur J Pers* 17(4):251–283.