DO I CARE FOR OTHERS' MONEY AS MUCH AS FOR MY OWN? PLAYING THE ULTIMATUM GAME TASK IN BEHALF OF A THIRD-PARTY

CORRADO CORRADI-DELL'ACQUA,
CLAUDIA CIVAI, RAFFAELLA I. RUMIATI

Cognitive Neuroscience Sector, Scuola Internazionale Superiore di Studi Avanzati (SISSA/ ISAS), Trieste, Italy

corradi@sissa.it

ABSTRACT
Classical economical theory sees rejections of unfair offers by people playing the Ultimatum Game (UG) as "irrational". Recent studies suggested that these are triggered by negative emotions, such as frustration (Sanfey et al., 2003; van't Wout et al., 2006) and by the urge to punish those who made the offers (Fehr & Gachter, 2002). Another account postulates that rejections are instead "rational" according to the rules of social exchange reasoning, in that they will increase the chance of future players to receive fair offers (Zamir, 2001). We tested these two accounts by employing healthy participants in modified version of the UG in which players knew that their putative rejections were not harming those who made offers. The analysis of skin conductance responses shows that this task was significantly less emotionally arousing than the traditional UG game. However, unfair offers were rejected at a comparable rate in both the classical and modified versions of the Ultimatum Game. In light of these results, theories holding rejections as triggered by emotional arousal and by the urge to punish who made the offers should be re-discussed; in fact, our data suggest that the emotional response might be triggered whenever one's own interest is at stake, and is not the ultimate cause of this behavior. We believe instead that any offer leading to an unfair distribution of money within the group is sufficient to trigger a rejection and, therefore, that psychological mechanisms which account for social exchanges dynamics might be better candidates for explaining this behavior.

Keywords: Economical Decision-Making; unfairness; altruistic punishment; frustration; skin conductance.

1. INTRODUCTION

Social interpersonal exchange is an ancient and cross-culturally universal feature of many species including insects, high-level primates and humans (e.g., Hamilton, 1964a, 1964b; Trivers, 1971). In most of its expressions, interpersonal exchange results in the ability to assign social credit or blame for shared outcomes and to act appropriately according to these assignments (Tolmin et al., 2006; Trivers, 1971): for instance, people might identify and encourage those behaviors which maximize the aggregate welfare and discourage those who lead to unfair distributions of goods.

In recent years, the study of interpersonal exchanges has become a prominent issue in both experimental psychology and cognitive neuroscience. Indeed, studies employing behavioral techniques have shown that reasoning in a context of social interpersonal exchanges diverge from reasoning in other domains. For instance, studies investigating conditional reasoning revealed healthy volunteers as uneasy at understanding conditional rules (e.g., "if *P* then *Q*"), as less that 30% of the interviewed people were able to identify their possible violation ("*P* and non-*Q*" – e.g., Watson & Johnson-Laird, 1972). However, more than 65% of the interviewed people were able to solve the task if the conditional rules were expressed as social contracts (e.g. "If John takes this *benefit*, he must pay this *price*") and their violations as a cheating behavior (e.g. "John takes this *benefit*, without paying this *price*" – e.g., Cosmides, 1989; Gigerenzer & Hug, 1992; Platt & Griggs, 1993; see also Stone et al., 2002).

Another example of reasoning in a context of social interpersonal exchanges is provided by the Ultimatum Game (UG) task. In this task, one player (the *proposer*) makes offers to a second player (the *responder*) of how to split an amount of money given by the experimenter; the *responder*, in turn, can either accept or reject the offer. If the responder accepts, the money will be divided as the *proposer* has decided, otherwise both players will end up with nothing. Classical economical theories posit that, to maximize his/her own gain, the *responder*, following the principle that "few is better than nothing", should always accept every offer. However, the behavioral findings clearly show that he rejects offers which favor the *proposer* too much, and that he/she considers unfair (Bolton & Zwick, 1995). Importantly, this behavioural pattern has also been observed in both the single-shot UG, in which the two players interact only once, and in the covered UG, in which the *proposer* is not informed about the *responder*'s reaction (Abbik, Sadrieh, & Zamir, 1999; Zamir, 2001), and therefore rejections lose their role as negotiating tools.

One plausible interpretation of the behavioural pattern is that people behave irrationally, as they are driven by negative emotions and by the desire to punish the *proposer*. According to one

of these theories, punishment, even if costly and yielding no direct benefit for the *responder* (as in the case of covered UG), is used to penalize the *proposer*'s selfish behavior (Fehr & Gachter, 2002). It has been suggested that the irrational rejection might be caused by negative emotions, such as frustration, that drive participants to punish rather than making an utilitarian choice (Pillutla et al., 1996; Fehr & Gachter, 2002). Consistently with this view, Sanfey et al. (2003) have recently associated the rejection of unfair offers with an increase of both the neural activity in anterior insula, traditionally correlated with feelings of anger and disgust (Calder, Lawrence, & Young, 2001; Phillips et al., 1997), and the skin conductance response (van't Wout, Kahn, Sanfey, & Aleman, 2006), as a measure of emotional activation (Boucscein, 1992).
Another plausible interpretation states that the *responder*'s rejections, although irrational according to the classical economical theory as they do not increase the chance of having better offers in the remaining part of the experiment, are rational according to the rules of social exchange reasoning, in that they discourage unfair behavior of future *proposers* and, therefore, increase the overall gain of the population of the *responders* (Zamir, 2001).
These two accounts – although never mentioned in literature as antithetic one to another – generate different predictions. Indeed, the former account postulates that rejections are driven by (1) *proposer*'s unfair treatment enhancing the *responder*'s emotional arousal and (2) the *responder*'s being aware that his choices are harmful for the *proposer*. Thus, no rejections should be seen in a modified version of the UG in which none of these assumptions are met. This is not the case of the second account, according to which any offer leading to an unfair distribution of money within the group should be rejected. We tested our prediction by measuring skin conductance response while participants played as *responders* in a modified version of the UG. In one condition, they carried out the classical UG, whereas, in a second condition, they were told that both *proposers* and *responders* were playing in behalf of a third-party. Thus, in the latter condition, neither the *proposer*'s offers addressed directly the *responder*'s payoff, nor *responder*'s choices addressed directly the *proposer*'s payoff. Our assumption (confirmed by the analysis of skin conductance responses) is that responder's emotional arousal should be significantly lower when playing in behalf of a third-party than in the classical UG. Thus, the account according to which rejections and negative emotions are causally related (Fehr & Fischbacher, 2002; Sanfey et al., 2003;

van't Wout et al., 2006) predicts that such a putative decrease in skin conductance response during the third-party, together with the awareness that the *proposer* cannot be punished by the *responder*, should lead a similar decrease in the rate of rejections. On the other hand, the account positing rejections as rational according to the rules of social exchange reasoning predicts that, irrespective of the amount of emotional arousal measured by the skin conductance response, participants should exhibit the same amount of rejections when playing in behalf of a third-party than in the classical UG.

2. METHODS

Thirty-one healthy Italian volunteers (20 females), who ranged in age from 18 to 35 years (M=23.56, SD=3.90), took part in the experiment. They were required to play as *responders* in a modified version of the UG in which they either accepted or rejected the offers the *proposer* made, following the classical rules explained above. Before starting the game, they were introduced to a collaborator of the experimenter, who pretended to play as the *proposer*, in order to strengthen the illusion of playing against a human adversary, whereas they were actually playing against a computer. They were told that the opponent had been given a number of 10 euros bank notes and would have made offers on how to split each of these bank notes. At each trial offers could range from 1 to 5 euros out of 10. Furthermore, participants were informed that, in one condition, they and their opponent would play for themselves (consistently with the classical UG), whereas, in a second condition, they would play on behalf of those players acting as *proposer* and *responder* in the upcoming testing session. In order to make our task compatible to the single-shot UG, participants were told that the opponent would not have received any feedback until the very end of the experiment, when they have both been informed on how much each of them had gained, depending on the choices they had made; in this way, they knew rationally that they could not affect the opponent's behaviour through their rejections. In addition, they knew that a percentage of the money split on behalf of third parties would be given to next players; they were also informed that, following the same principle, their starting stakes were percentages of the money that previous players have split on their behalf. To control for the social interactive nature of the UG, participants performed a control task (Free Win [FW] task) in which they either accepted or rejected a variable amount of money given by the computer (1-5 euros). As in the case of the UG, they could have decided for themselves or on behalf of the next participant. If they accepted the offer, they/the third party would have received that amount, otherwise they/the third party would receive nothing. This yields to a 2*2*5 design with task (UG *vs.* FW), target *(myself vs. third-party)* and gain (1-5 euros) as within-subjects factors. Participants were informed that their compensation for participating in the experiment would be proportional to the amount of money gained during the myself condition. Moreover, they knew that a percentage of the money split on behalf of third parties would be given to next players; they were also informed that, following the same principle, their starting stakes were percentages of the money that previous players had split on their behalf. Irrespective of their performance on the task, participants received the same amount of money as compensation. During the experimental session the offer appeared on the screen for five seconds, followed by a six-second blank screen. Participants were required to respond by button press, highlighted on the computer keyboard, as soon as the question "Do you accept?" appeared on the screen, where it lasted for two seconds. The inter-trial interval was averaged around 11 seconds, to allow skin conductance to return to its baseline. All 20 conditions, each of which were repeated four times, were randomized through the experiment, with duration of approximately 32 minutes (80 trials * 24 seconds of trial duration). Skin conductance was recorded during the whole experiment using a pair of prewired 8 mm Ag/AgCl electrodes, attached to the distal phalanx surfaces of the index and little finger of the non-dominant hand. The electrode pair was excited with a constant voltage of 0.5 V and the conductance was recorded using a DC amplifier with a low pass filter set at 64 Hz and a sample frequency of 256. Values of skin conductance were automatically transformed to microsiemens values by the Procomp Infinity System and further analyzed using Ledalab 2.1.3 freeware software (Benedek & Kaernbach, 2008). A phasic increase in conductance of 0.05 μS or more was counted as a Skin Conductance Response (SCR). Our analysis focused on the average phasic increase in the four seconds prior to the moment in which subjects were instructed to provide a response.

3. RESULTS

For each subject, and for each condition, the rejection rate and the average phasic increase in conductance were calculated across all 4 repetitions, and used in a 2 (TASK: UG, FW) x 2, (TARGET:

myself, third party) x 5 (GAIN: 1, 2, 3, 4, 5 Euros) Repeated Measures ANOVA as implemented in SPSS 11.5 Software. The analysis of Rejection Rates revealed a significant main effect of TASK (MSE = 2336.51, $F(1, 30) = 58.96$, $p < .001$, $\eta_p^2 = .66$), with the UG eliciting a larger amount of rejections (myself: 37.26 ± 3.89%;third-party: 40.54 ± 4.64%) than the FW (myself: 5.32 ± 2.12%; third-party: 12.83 ± 3.64%). The TASK x GAIN interaction was also found significant (MSE = 544.065, $F(4, 120) = 39.67$, $p < .001$, $\eta_p^2 = .57$), reflecting low offers being rejected significantly more than the high offers in the UG and not in the FW. Polynomial contrasts revealed that the difference between UG and FW changed linearly across all gain levels (MSE = 1030.95, $F(1, 30) = 78.09$, $p < .01$, $\eta_p^2 = .72$).
The analysis of SCRs revealed a significant TASK x TARGET interaction (MSE = .012, $F(1, 30) = 6.24$, $p < .05$, $\eta_p^2 = .17$), reflecting the largest phasic increase in skin conductance whilst playing the UG for oneself (0.14 ± 0.02 μS), with respect to the UG on behalf of a third-party (0.11 ± 0.02 μS), or to the FW (myself: 0.11 ± 0.02%; third-party: 0.13 ± 0.02%). None of the remaining effects of the ANOVAs were found to be significant

4. DISCUSSION

We have investigated the nature of "irrational" rejections during the Ultimatum Game by having participants perform a modified version of the paradigm in which they were asked to play for themselves or on behalf of a third party. Our analysis of the conditions in which participants played for themselves confirmed all previous findings associated with the classical UG task: first, we replicated the well-documented pattern of accepting fair offers and increasing the rate of rejection as offers become less fair (Bolton & Zwick, 1995; Roth, 1995; Guth, Huck, & Muller, 2001; Sanfey et al., 2003); this pattern was not found in the control task, in which participants had to either accept or reject money given by the computer, supposing that, even though the *responder*'s personal gain is the same, the perception of an unfair division drives him/her to reject unfair offers choosing the so called non-utilitarian or "irrational" solution. Secondly, skin conductance data show that, when participants played for themselves, the UG elicited a larger emotional arousal than when played the FW task.
More importantly, the analysis of SCRs revealed a TARGET*TASK interaction, which reflects a significant increase of SCRs associated with the UG (but not the FW), when their own interests (but not

that of others) were at stake. The fact that playing the UG for oneself is more emotionally arousing than the other three conditions, confirm our initial assumption that the amount of negative emotions elicited by the UG is smaller when one's own payoff is not directly at stake.

If rejections are not utilitarian in nature, as they are driven by negative emotions and by the desire to punish the *proposer* (Sanfey et al., 2003; van't Wout et al., 2006), these should not be found (or should be found in a lesser degree) whilst playing the UG in behalf of a third party as (1) participants are aware that their putative rejection is not affecting the *proposer*'s payoff and (2) this condition is less emotionally arousing. Our analysis of the rejection rates go against this prediction: indeed no TARGET*TASK interaction was found to be significant, as in the case of the analysis of SCRs, but the TASK main effect was, reflecting the amount of rejections associated with playing the UG for both targets (myself: 37%; third-party: 40%) being significantly larger than the amount of rejection associated with the FW task (myself: 5%, third-party: 12%). In light of our results, the account according to which rejections are irrational responses, driven exclusively by emotions and by the desire to punish the proposer, needs to be reconsidered. An alternative explanation for the *responder*'s behavior can be related to the notion of context dependent fairness proposed by Zamir and colleagues (Winter & Zamir, 2005; Zamir, 2001), according to which the sense of equity may change depending on both the person engaged in the social exchanges dynamic, and the nature of this dynamic. Winter and Zamir (2005), for instance, reported a modified version of the UG in which the *proposer* played with virtual-*responders* which could be either much more tolerant or unforgiving to unfair offers than real human *responders*. They found that the *proposers* quickly adapted their behavior to the virtual-*responders*, by behaving unfairly with the tolerant and fairly with the unforgiving *responders*. This is similar to what happens in the Dictator Game (Forsythe, Horowitz, Savin, & Safton, 1994; Bolton & Zwick, 1995), in which the *proposer* cannot have his offers rejected by the *responder* and, therefore, behaves far less fairly than in the UG. All these observations suggest that in the UG *proposers*' behavior is directly affected by the tolerance to unfairness he expects in the *responder*. Thus, rejections in the UG are indeed rational in that, although they do not increase the *responder*'s chance of having better offers in the remaining part of the experimental session, they lead the *proposers* to play fairly

and, in turn, to an increase the overall gain of the population of the *responders* (Zamir, 2001). Such behavior is not predicted by classical economical theory, which is based on the assumption (present in the experimental instructions of many experiments, but less frequently believed by participants) that player's choices have effects which are limited to the experimental session (Zamir, 2001). The account according to which the *responder*'s rejections are utilitarian is in agreement with our behavioral results. In our study, participants were told prior to the experiment that their starting stakes depended on how previous players had decided to split the money; it is therefore likely that they felt part of a group in which cooperation led to a maximization of everyone's gain. Thus the participants' prosocial behavior showed when they rejected the unfair offers on behalf of the third party, might reflect the will to strengthen the public good for the population of the *responders* (Zamir, 2001). This is also consistent with studies of social psychology in which participants decided give up some money in order to punish other's unfair behavior, even when their payoffs are not directly affected by a violation of fairness (Fehr & Fischbacher, 2004).
This account is in agreement also with our psychopshysiological data, in that it does not predict that the rejection rates are associated with an increased emotional response. Emotions do play a role in the UG, as demonstrated by previous studies (Sanfey et al., 2003; van't Wout et al., 2006) as well as by the present study when participants played in the myself condition (i.e. the classical version of the UG). However, the dissociation we have reported here between the physiological and the behavioral pattern when participants performed the task in the third-party condition, implies that emotions are not always the key mechanism underlying the *responder*'s rejections. The emotional response might be triggered whenever one's own interest is at stake, and not the ultimate cause of this behavior.

REFERENCES

– Abbink, K, Sadrieh, A, & Zamir, S. (1999). The covered response Ultimatum Game. Discussion Paper #191, The Hebrew University, Center for Rationality and Interactive Decision Theory.
– Benedek, M., & Kaernbach, C. (2008). Ledalab: Leipzig electro-dermal

activity laboratory (Version 2.1.3). Available at http://www.ledalab.de.
- Bolton, G. E., & Zwick, R. (1995). Anonymity versus punishment in ultimatum bargaining. *Games and Economic Behavior*, 10, 95-121.
- Bouscein, W. (1992). *Electrodermal activity*. New York, NY: Plenum.
- Calder, A. J., Lawrence, A. D., & Young, A. W. (2001). Neuropsychology of fear and loathing. *Nature Reviews Neuroscience*, 2, 352-363.
- Cosmides, L. (1989). The logic of social exchange: has natural selection shaped how humans reason? Studies with the Watson selection task. *Cognition*, 31, 187-276.
- Fehr, E., & Gachter, S. (2002). Altruistic punishment in humans. *Nature*, 415, 137-140.
- Fehr, E., & Fischbacher, U. (2004). Third-party punishment and social norms. *Evolution and Human Behavior*, 25, 63–87.
- Forsythe, R., Horowitz, J. L., Savin, N. E., & Safton, M. (1994). Fairness in simple bargaining experiments. *Games and Economic Behavior*, 6, 347-369.
- Gigerenzer, G., & Hug, K. (1992). Domain-specific reasoning: social contracts, cheating, and perspective change. *Cognition*, 43, 127-171.
- Guth, W., Huck, S., & Muller, W. (2001). The relevance of equal splits in Ultimatum Games. *Games and Economic Behavior*, 37, 161–169.
- Hagen, E. H., & Hammerstein, P. (2006). Game theory and human evolution: A critique of some recent interpretations of experimental games. *Theoretical Population Biology*, 69, 339-348.
- Hamilton, W. D. (1964a). The genetical evolution of social behavior I. *Journal of Theoretical Biology*, 7, 1-16.
- Hamilton, W. D. (1964b). The genetical evolution of social behavior II. *Journal of Theoretical Biology*, 7, 17-52.
- Kohlberg, L. (1969). Stage and sequence: the cognitive-developmental approach to socialization. In D.A. Goslin (Ed.), *Handbook of Socialization Theory and Research* (pp. 347-480). Chicago, IL: Rand McNally.
- Moll, J., Zahn, R., de Oliveira-Souza, R., Krueger, F., & Grafman, J. (2005). The neural basis of human moral cognition. *Nature Reviews Neuroscience*, 6, 799-809.
- Moll, J., & de Oliveira-Souza, R. (2007). Moral judgments, emotions and the utilitarian brain. *Trends in cognitive science*, 11, 319-321.
- Phillips, M. L., Young, A. W., Senior, C., Brammer, M., Andrew, C., Calder, A. J., Bullmore, E. T., Perrett, D. I., Rowland, D., Williams, S. C., Gray, J. A., & David, A. S. (1997). A specific neural substrate for perceiving facial expressions of disgust. *Nature*, 389, 495-498.
- Pillutla, M. M., & Murninghan, J. K. (1996). Unfairness, anger, and spite: emotional rejections of ultimatum offers. *Organizational Behavior and Human Decision Processes*, 68, 208-224.
- Platt, R. D. & Griggs, R. A. (1993). Darwinian algorithms and the Wason selection task: a factorial analysis of social contract selection task problems. *Cognition*, 48, 163-192.
- Polezzi, D., Daum, I., Rubaltelli, E., Lotto, L., Civai, C., Sartori, G.,

& Rumiati, R. (2008). Mentalizing in economic decision-making. *Behavioural Brain Research*, 190, 218-223.

- Roth, A.E. (1995). Bargaining experiments. In: Kagel, J., Roth, A. (Ed.). *Handbook of experimental economics* (253-342). Princeton, NJ: Princeton University Press.
- Sanfey, A. G., Rilling, J. K., Aronson, J. A., Nystrom, L. E., & Cohen, J. D. (2003). The neural basis of economical decision-making in the Ultimatum Game. *Science*, 300, 1755-1758.
- Stone, V. E., Cosmides, L., Tooby, J., Kroll, N., & Knight, R. T. (2002). Selective impairment of reasoning about social exchange in a patient with bilateral limbic system damage. *Proceedings of the National Academy of Sciences of the United States of America*, 99, 11531-11536.
- Tolmin D., Amin Kayali, M., King-Casas, B., Anen, C., Camerer, C. F., Quartz, S. R., & Read Montague, P. (2006). Agent-Specific Responses in the Cingulate Cortex During Economic Exchanges. *Science*, 5776, 1047-1050.
- Trivers, R. (1971). The Evolution of Reciprocal Altruism. *The Quarterly Review of Biology*, 46, 35-57.
- van 't Wout, M., Kahn R. S., Sanfey, A. G., & Aleman, A. (2006). Affective state and decision-making in the Ultimatum Game. *Experimental Brain Research*, 169, 564-568.
- Von Neumann, J., & Morgestern, O. (1947). *Theory of games and economic behavior.* Princeton, NJ: Princeton University Press.
- Watson, P., & Johnson-Laird, P. (1972). *Psychology of Reasoning: Structure and Content.* Cambridge, MA: Harvard University Press.
- Winter, E., & Zamir, S. (2005). An Experiment with Ultimatum Bargaining in a Changing Environment. *Japanese Economic Review* , 56, 363-385.
- Zamir, S. (2001). Rationality and emotions in Ultimatum Bargaining. *Annales D'Economie et De Statistique*, 61, 1-31.