Are irrational reactions to unfairness truly emotionally-driven? Dissociated behavioural and emotional responses in the Ultimatum Game task

Claudia Civaia, Corrado Corradi-Dell'Acqua, Matthias Gamer, Raffaella I. Rumia

A R T I C L E   I N F O

Article history:
Received 8 October 2008
Revised 21 August 2009
Accepted 1 September 2009

Keywords:
Economical decision-making
Unfairness
Altruistic punishment
Frustration
Skin conductance response
Utilitarian rejections

A B S T R A C T

The “irrational” rejections of unfair offers by people playing the Ultimatum Game (UG), a widely used laboratory model of economical decision-making, have traditionally been associated with negative emotions, such as frustration, elicited by unfairness (Sanfey, Rilling, Aronson, Nystrom, & Cohen, 2003; van't Wout, Kahn, Sanfey, & Aleman, 2006). We recorded skin conductance responses as a measure of emotional activation while participants performed a modified version of the UG, in which they were asked to play both for themselves and on behalf of a third-party. Our findings show that even unfair offers are rejected when participants’ payoff is not affected (third-party condition); however, they show an increase in the emotional activation specifically when they are rejecting offers directed towards themselves (myself condition). These results suggest that theories emphasizing negative emotions as the critical factor of “irrational” rejections (Pillutla & Murnighan, 1996) should be re-discussed. Psychological mechanisms other than emotions might be better candidates for explaining this behaviour.

© 2009 Elsevier B.V. All rights reserved.

1. Introduction

In recent years the study of the role of emotions in decision-making has become an increasingly prominent issue in cognitive neuroscience. A wealth of studies have hypothesized an emotional pathway in the brain that seems to operate in many types of decisional processes, including moral judgment (Moll, Zahn, de Oliveira-Souza, Krueger, & Grafman, 2005, for a review) and economical decision-making (Pillutla & Murnighan, 1996; Sanfey et al., 2003), that have traditionally been linked to rational thinking and choices (Kohlberg, 1969; von Neumann & Morgestern, 1947).

The Ultimatum Game (UG), a model of economical decision-making employed in the laboratory, has always been thought of as a classical example of emotionally-driven behaviour. 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 offers. If the responder accepts, the money will be divided as the proposer has decided, otherwise both players will receive nothing. Classical economical theories posit that, to maximize his/her own gain, the proposer should always offer the smallest amount of money, whilst the responder, following the principle that “few is better than nothing”, should accept every offer. However, the behavioural findings clearly show that the proposer typically divides the money equally, and that the responder rejects offers which favor the proposer too much, and those that he/she considers unfair (Bolton & Zwick, 1995).

0010-0277/$ - see front matter © 2009 Elsevier B.V. All rights reserved.
doi:10.1016/j.cognition.2009.09.001
Individuals’ “irrational” choices have been explained in terms of altruistic punishment (Fehr & Gachter, 2002). According to this theory, the punishment, even if costly and yielding no direct benefit for the punisher (as in the case of single-shot UG), is used to penalize selfish behaviour of others, as it leads them to cooperate in future interactions (Fehr & Gachter, 2002). It has been suggested that irrational rejections might be best explained by negative emotions, such as frustration, that drive participants to punish rather than making an utilitarian choice (Fehr & Gachter, 2002; Pillutla & Murnaghan, 1996). Consistent 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 et al., 2006), a measure of emotional activation (Bouscain, 1992).

It has also been argued that only self-centered emotions, for instance anger and frustration, play some role in the UG, as individual payoff is heavily involved in this task (Moll & de Oliveira-Souza, 2007). However, it has also been proved that individuals choose an act of punishment even though their payoffs are not directly affected by a violation of fairness and cooperation norms (i.e. the third-party punishment). Fehr and Fischbacher (2004), for instance, found that participants decided to give up some of their own money to punish the unfair behaviour of one player towards another. Thus, altruistic punishment also occurs in conditions in which unfairness should not elicit, at least in principle, any self-centered emotion. This raises the question of whether, in the UG task, the “irrational” punishing behaviour and negative emotions are always causally related, or whether they can operate separately depending on the myself/third-party distinction.

In the present study, we investigated the role of emotions in the UG by measuring skin conductance responses (SCR) while participants played as responders in a modified version of the UG and by collecting emotional ratings after they completed the task to measure the valence of the hypothetical arousal. Participants carried out both the classical version of the UG and a modified version of the task in which any putative monetary income was not going into the participants’ own pocket, but into a third-party’s (see Section 2). Indeed, in the latter condition the proposer’s offer did not directly address the participant's payoff, unfairness should in principle, elicit neither self-centered emotions (Moll & de Oliveira-Souza, 2007), nor, as consequence, SCR increases when the offer is about to be rejected (van’t Wout et al., 2006). Thus, the account according to which the punishing behaviour and the negative emotions are causally related (Fehr & Gachter, 2002; Pillutla & Murnaghan, 1996), also predicts that such emotional decrease should be associated with a similar decrease in the amount of punishing choices (rejections) (van’t Wout et al., 2006). However, based on previous studies of altruistic punishment (e.g. Fehr & Fischbacher, 2004), we predicted that participants should reject unfair offers addressing a third-party: if this were indeed the case, we expected a significant increase in SCR for offers about to be rejected even in the third-party condition as well.

2. Methods

2.1. Participants

Thirty-four healthy Italian volunteers (22 females), who ranged in age from 18 to 35 years (M = 23.56, SD = 3.90), took part in the experiment. They all were paid for participating in the study, the scientific goal of which was unknown to them. The study was approved by the local ethics committee and conducted in accordance with the Declaration of Helsinki.

2.2. Task

Participants were required to play as responders in a modified version of the UG and had either to accept or reject 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 to make offers on how to split each of them. Consistent with previous studies (e.g. Polezzi et al., 2008), offers in each trial could be either 1, 2, 3, 4 or 5 euros out of 10. Furthermore, participants were informed that, in one condition, they and their opponent would play for themselves (consistent with the classical UG), whereas, in another condition, they would play on behalf of those players acting as proposer and responder in the upcoming testing session (see Fig. 1). In order to make our task compatible to the single-shot UG, participants were told that the opponent would receive feedback only at the 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.

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, 2, 3, 4 or 5 euros). As in the case of the UG, they could decide for themselves or on behalf of the next participant. If they accepted the offer, they/the third-party would receive that amount, otherwise they/the third-party would receive nothing. This yielded to a $2 \times 2 \times 5$ design, with TASK (UG vs. FW), TARGET (myself vs. third-party) and GAIN (1, 2, 3, 4 or 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 in 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. Although we did not systematically investigate whether participants had doubts about the authenticity of the situation, the majority of them, when informally interviewed afterwards, said they believed they had played against a human opponent. Only a few reported having doubts at the end of the experimental session.

2.3. Apparatus and procedure

All the participants were tested in a quiet room at SISSA using a PC and a 15-in monitor (OliData s.p.a.). Presentation® 12.0 Software (http://www.neurobs.com) was used to construct and deliver the experimental stimuli. The offer appeared on the screen for 5 s, followed by a 6 s 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 2 s (see Fig. 2). The inter-trial interval was around 11 s on average, to allow skin conductance to return to its baseline. All 20 conditions, each of which was repeated four times, were presented in a randomized order. The whole experiment (80 trials + 24 s of trial duration) including a short break of 1 min after half of the trials lasted approximately 33 min.

2.4. Skin conductance recordings

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 conductance was recorded using a DC amplifier with a low-pass filter set at 64 Hz and a sample frequency of 256 Hz. Values of skin conductance were automatically transformed to microsiemens values by the Procomp Infinity System (Bio-Medical Instruments, Inc., Warren, MI, USA). Before starting the task, 1 min of baseline was recorded. We measured the artifact-free amplitude of the skin conductance response that began between 1 and 3 s after the presentation of the offer and exceeded a threshold of

---

**Fig. 1.** Illustration of the task as it was presented to participants when giving the instructions. There are four conditions: the first and the second refer to the Ultimatum Game and the third and the fourth refer to the control task (Free Win situation). In the first and in the third conditions participants are asked to decide for themselves, whereas in the second and in the fourth they are asked to choose on behalf of a third-party (next participant).

**Fig. 2.** Time line for each single trial of the Ultimatum Game. Each trial lasted 24 s. First, participants saw the offer on the screen for 5 s, followed by 6 s of blank screen. Next, the question “do you accept/do you accept on his behalf” appeared on the screen for 2 s, within which participants had to answer by button press. An average of 11 s inter-trial interval followed the question.
0.05 μS. In the case of overlapping responses, the inflection point between the two responses served as the baseline or peak, depending on the latency criterion. The resulting amplitudes were z-transformed within each participant in order to eliminate individual differences in responsivity.

2.5. Emotional ratings

To further investigate the emotional reactions in our study, participants rated their feelings in the most crucial conditions (i.e. 1, 3 and 5 euros of gain when playing the UG in both the myself and the third-party condition) at the end of the experimental session. Since we were interested in detecting the presence of a perceived emotion, here we included only the extreme conditions (and not those associated with 2 and 4 euros) on the assumption that they were those more likely to elicit a stronger, and thus the most easily detectable, response. The mid-value offer (3 euros) was chosen as a baseline, since it could reflect the emotionally neutral condition. Participants used a 12-point Likert-scale for each condition ranging from −6, corresponding to strong negative emotions, to +6, indicating strong positive emotions.

3. Results

3.1. Rejection rates

For each subject and condition, the rejection rates were calculated across all 4 repetitions, and used in a 2 (TASK: UG, FW) × 2 (TARGET: myself, third-party) × 5 (GAIN: 1, 2, 3, 4, 5 euros) Repeated Measures ANOVA. Statistical Analysis was carried out using SPSS 11.5 Software (SPSS Inc., Chertsey UK). Results indicated a significant main effect of TASK (F(1, 33) = 76.24, p < .001, ηp² = .69), with the UG eliciting a larger amount of rejections than the FW (see Table 1 and Fig. 3), as well as a main effect of GAIN (F(4, 132) = 52.7, p < .001, ηp² = .61), with low offers being rejected more than high offers. This effect is however driven by the TASK × GAIN interaction, which was found to be significant as well (F(4, 132) = 49.89, p < .001, ηp² = .60), suggesting that low offers are rejected significantly more often than high offers in the UG but not in the FW. None of the remaining effects of the ANOVA were found to be statistically significant.

3.2. Emotional ratings

We analyzed the emotional ratings for the most unfair offer (1 euros out of 10), the fairest offer (5 euros out of 10) and the mid-value offer (3 euros out of 10), both for the myself and the third-party conditions. One-sample two-tailed T-tests showed that for the mid-value offer the ratings did not differ significantly from zero (i.e. the neutral emotion), while for both targets, the ratings for the unfair offer were significantly different from 0 towards the negative emotion (UG (1:9) myself: t (33) = −9.79, p < .001 UG (1:9) third-party: t (33) = −4.37, p < .005), and so were those for the fair offer towards the positive emotion (UG (5:5) myself: t (33) = 22.29, p < .001; UG (5:5)

Table 1
Rejection rates (RR) (%) and skin conductance response amplitudes (SCR amp) [z-transformed μS] for the four conditions collapsed by gain.

<table>
<thead>
<tr>
<th></th>
<th>UG</th>
<th>FW</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Myself</td>
<td>Third-party</td>
</tr>
<tr>
<td>RR (SEM)</td>
<td>35.73 (5.59)</td>
<td>38.09 (4.93)</td>
</tr>
<tr>
<td>SCR amp (SEM)</td>
<td>.0887 (.0157)</td>
<td>-.0073 (.0164)</td>
</tr>
</tbody>
</table>

Note: Corresponding standard errors of the mean are printed in brackets.
third-party: \( t(33) = 5.63, p < .005 \). Moreover, an ANOVA, considering TARGET (myself and third-party) and GAIN (1, 3, 5 euros out of 10) as factors, showed a significant effect of TARGET (\( F(1, 33) = 4.328, p < .05, \eta^2_p = .116 \)), a significant effect of GAIN (\( F(2, 66) = 101.82, p < .001, \eta^2_p = .75 \)), and a significant TARGET × GAIN interaction (\( F(2, 66) = 12.662, p < .001, \eta^2_p = .277 \)). A paired-samples T-test demonstrated that there was a significant difference between targets for fair (\( t(33) = 4.01, p < .001 \)) and unfair (\( t(33) = -2.742, p < .01 \)) offers, while no difference between targets were found for the mid-value offer; in particular, both the reported positive and the negative emotions were rated as stronger in the *myself* than in the *third-party* condition (see Fig. 4).

3.3. Skin conductance response amplitude

For each subject and condition, the average of z-standardized skin conductance response amplitudes were calculated across all 4 repetitions, and used in a 2 (TASK) × 2 (TARGET) × 5 (GAIN) Repeated Measures ANOVA. We found a significant main effect of TASK (MSE = 0.23, \( F(1, 33) = 4.91, p < .05, \eta^2_p = .13 \)) and a significant main effect of TARGET (MSE = 0.28, \( F(1, 33) = 7.93, p < .01, \eta^2_p = .19 \)), suggesting that participants were more aroused whilst playing the UG than the FW, and when their own interest, and not the third-party’s, was at stake (see Table 1). None of the remaining effects were found to be significant.

In addition, we investigated the relation between SCR and rejections. Following van’t Wout et al. (2006), we focused our analysis on small offers (1 euro), as in the UG they were associated with the largest negative emotional arousal (see our analysis of emotional ratings above). Consistent with the emotional ratings, that was collected only for the 1-euro offers, the 2-euros unfair offers were not considered in this analysis. We used a Linear Mixed Model (Neuhaus, McCulloch, & Shayle, 2008) which is more robust against missing cells, as few subjects scored in all conditions. The model included RESPONSE (accept/reject), TARGET and TASK as fixed factors, and SUBJECTS as random factor. A compound symmetry covariance structure was specified. We found a significant main effect of TARGET (\( F(1, 138.67) = 7.36, p < .01 \)), indicating a stronger emotional arousal when offers were directed to oneself (\( 0.23 \pm 0.11 \) z-transformed SCR) rather than to a third-party (\( 0.085 \pm 0.085 \) z-transformed SCR), and a significant TARGET × RESPONSE interaction (\( F(1, 144.91) = 4.28, p < .05 \)), reflecting participants’ higher SCR amplitudes when rejecting small offers for themselves than when rejecting for a third-party. No target difference was found for the acceptances (see Table 2 and Fig. 5).

### Table 2

<table>
<thead>
<tr>
<th></th>
<th>Myself</th>
<th>Third-party</th>
</tr>
</thead>
<tbody>
<tr>
<td>Rejections</td>
<td>.560 (.151)</td>
<td>-.085 (.146)</td>
</tr>
<tr>
<td>Acceptances</td>
<td>.043 (.115)</td>
<td>-.043 (.128)</td>
</tr>
</tbody>
</table>

Note: Corresponding standard errors of the mean are printed in brackets.
dressed oneself or a third-party; however they exhibited an increased negative emotional arousal when about to reject the most unfair offer addressing oneself (but not a third-party). The account according to which rejections in the UG are irrational responses driven exclusively by negative emotions should therefore be reconsidered.

We replicated the well-documented pattern of accepting fair offers and increasing the rate of rejection as offers become less fair (Bolton & Zwick, 1995; Guth, Huck, & Muller, 2001; Roth, 1995; Sanfey et al., 2003). In keeping with what predicted by Fehr and Fischbacher (2004), participants showed the same behaviour even when playing on behalf of a third-party. This pattern was not found in the control task, in which participants had to either accept or reject money given by the computer. This allows us to conclude that, even though the responder’s personal gain is the same in both UG and control task, in the UG only the perception of an unfair division of money drives him/her to reject these offers choosing the so called non-utilitarian or “irrational” solution.

The analysis of the electrodermal activity revealed an increase of offer-related SCR amplitudes whilst playing the UG, relative to the FW, and when one’s own interest, relative to a third-party’s, was at stake. No significant effects associated with the factor GAIN were found thus suggesting an equal amount of emotional arousal irrespective of the magnitude of the offers. However, the analysis of emotional ratings revealed a significant increase of negative emotions associated with the most unfair condition (1 euro out of 10), a significant increase of positive emotions associated with the most fair condition (5 euros out of 10) and no significant emotional activation during mid-value offers (3 euros out of 10). As they were not collected online, it is not possible to make a trial-by-trial correlation between the emotional ratings and the increase in SCR. However as it is well established that SCR can reflect an unspecified emotional arousal, the association of the physiological response with those conditions reported subjectively as emotionally arousing suggests that the significant increase of SCR amplitudes paired with both fair and unfair conditions in the UG (as opposed of the FW) could reflect an increased emotional arousal irrespective of valence. As for the case of the mid-value offers, which were not evaluated as emotionally arousing by participants, SCR might instead reflect cognitive effort (Bouscein, 1992), as mid-value offers in the UG are usually associated with the longest response times, and with a larger N350 after the presentation of the offer (Polezzi et al., 2008), which usually occurs when ambiguous stimuli are processed (e.g. Schendan & Kutas, 2003).

Finally, when we focused on the trials associated with small offers (1 euro), which in the case of the UG are the most unfair and are associated with largest negative emotional activation, we found a significant increase of SCR when rejecting (rather then accepting) offers addressing oneself. Such an increase (reminiscent of the one first described by van’t Wout et al. (2006)) was not found when the offers were directed at a third-party. Thus, if rejections are emotionally-driven, as they are not utilitarian in nature (Fehr & Gachter, 2002; Pillutla & Murningham, 1996), we would expect to find an increase in the electrodermal activity when participants reject (compared with when they accept) small offers, also when these effect a third-party. Instead, our data suggest that participants’ rejections and their emotional reaction are independent, although co-occurring when participants play the UG for themselves.

An alternative explanation for the responder’s behaviour 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 interaction dynamic, and the nature of this dynamic. For instance, Winter and Zamir (2005), 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 behaviour 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 (Bolton & Zwick, 1995; Forsythe, Horowitz, Savin, & Safton, 1994), 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' behaviour is directly affected by the tolerance to unfairness he expects in the responder. Even though rejections in the UG are irrational from an individualistic perspective, in that the money loss does not increase the responder's chance of having better offers in the remaining part of the experimental session, they can be considered rational from a collectivistic point of view, because they are supposed to lead the proposers to play fairly and, consequently, to an increase in the overall gain for the population of responders (Zamir, 2001). The account according to which the responder's rejections are utilitarian is in agreement with our behavioural 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' rejection of the offers on behalf of the third-party, which are considered unfair, might reflect the will of preventing a bargain which, if accepted, would be detrimental for the population of the responders (Zamir, 2001). Critically, this account does not necessarily predict that rejection should be associated with an increased negative emotional arousal.

That emotions do play a role in the UG is demonstrated by previous studies (e.g. Harlé & Sanfey, 2007; Sanfey et al., 2003; van't Wout et al., 2006) as well as by the present study, when participants played in the UG in the myself condition. In fact we do not exclude that other emotional responses might have entered in this social interaction. It is plausible, for instance, that altruistic feelings and motivations similar to those described by Moll et al. (2006) with regard to charitable donation, contribute to act in the same way both for oneself and on behalf of another person, by rejecting the unfair behaviour. What our findings seem to suggest is that negative emotions are not always the key-mechanism underlying the responder's rejections. These emotions might be triggered whenever one's own interest is at stake, and are not the ultimate cause of this behaviour. Future research is necessary to further understand this phenomenon. Imaging techniques, in particular, may help to disentangle between areas associated with the rejections in the myself and in the third-party conditions.

Acknowledgments

We would like to thank Gareth Bland for resuming the manuscript and Prof. Rino Rumiati for his precious insights and suggestions.

References


