Medial prefrontal cortex reacts to unfairness if this damages the self: a tDCS study

Claudia Civai,[1] Carlo Miniussi,[2,3] and Raffaella I. Rumiati[4]

[1]Radboud University, Donders Institute for Brain, Cognition and Behaviour, Nijmegen, The Netherlands, [2]Neuroscience Section, Department of Clinical and Experimental Sciences, University of Brescia, Italy, [3]Cognitive Neuroscience Section, IRCCS Centro San Giovanni di Dio Fatebenefratelli, Brescia Italy, and, and [4]Neuroscience Area, SISSA, Trieste, Italy

Neural correlates of unfairness perception depend on who is the target of the unfair treatment. These previous findings suggest that the activation of medial prefrontal cortex (MPFC) is related to unfairness perception only when the subject of the measurement is also the person affected by the unfair treatment. We aim at demonstrating the specificity of MPFC involvement using transcranial direct current stimulation (tDCS), a technique that induces cortical excitability changes in the targeted region. We use a modified version of the Ultimatum Game, in which responders play both for themselves (myself—MS condition) and on behalf of an unknown third-party (TP condition), where they respond to unfairness without being the target of it. We find that the application of cathodal tDCS over MPFC decreases the probability of rejecting unfair offers in MS, but not in TP; conversely, the same stimulation increases the probability of rejecting fair offers in TP, but not in MS. We confirm the hypothesis that MPFC is specifically related to processing unfairness when the self is involved, and discuss possible explanations for the opposite effect of the stimulation in TP.

INTRODUCTION

In recent years a lot of attention has been devoted to the investigation of the cognitive, emotional and neural substrates of social behavior in economic contexts. The concept of human rationality, considered as the axiom according to which people always act to pursuing their own self-interest and to maximizing their resources, has been challenged by the findings in behavioral economics that show how prosocial behavior, altruism and, more in general, other-regarding preferences often overcome self-interest in guiding human decisions. The Ultimatum Game (UG) is a widely used laboratory tool to investigate economic decision-making, and it is one of the most straightforward examples of non-utilitarian behavior. In this game, one player (proposer) is given an amount of money, e.g. 10 euros, and is required to share this money with a second player (responder); the responder can either accept or reject the offers made by the proposer, knowing that if he/she accepts, the money will be divided as the proposer decided, whereas if he/she rejects, no one gets anything. Although the rational self-interested response would be to accept even the lowest offer, because 1 euro is always better than zero, people, on average, tend to reject offers that are considered unfair (Güth et al., 1982). Different theories have been put forward for explaining this phenomenon: reciprocity theories (Rabin, 1993; Falk and Fischbacher, 2006) interpret rejections as a punishment for unfair behavior; inequality aversion (Bolton, 1991; Fehr and Schmidt, 1999; Bolton and Ockenfels, 2000) assumes that people are intrinsically adverse to inequality, in particular the disadvantageous type; negative emotional reaction theory (Pillutla and Murnighan, 1996; Sanfey et al., 2003; van't Wout et al., 2006) relies on psychological, psychophysiological and neuroimaging evidence of negative emotional involvement to explain rejections as driven by an impulsive emotional negative reaction to unfair treatment.

As far as the neural correlates are concerned, the activation of areas such as anterior insula (Sanfey et al., 2003; Güroğlu et al., 2010; Chang and Sanfey, 2013), anterior cingulate (Sanfey et al., 2003; Boksem and

De Cremer, 2010), medial (Güroğlu et al., 2010; Campanhã et al., 2011; Corradi-Dell'Acqua et al., 2013;), ventromedial (Koenigs and Tranel, 2007; Moretti et al., 2009), ventrolateral (Tabibnia et al., 2008) and dorsolateral (Sanfey et al., 2003; Knoch et al., 2006) prefrontal cortex (PFC) has repeatedly been associated with UG rejections. Patient studies have shown that a lesion of the ventromedial PFC (VMPFC) increases the rate of rejections of unfair offers in the UG (Koenigs and Tranel, 2007), although only under certain circumstances, i.e. when the monetary reward is not immediately available to the patient (Moretti et al., 2009); non-invasive brain stimulation (NIBS; i.e. repetitive transcranial magnetic stimulation—rTMS, or transcranial direct current stimulation—tDCS) studies found that interfering with the activity in the dorsolateral PFC (DLPFC) decreases the rate of rejections (van't Wout et al., 2005; Knoch et al., 2006, 2008). Combining functional magnetic resonance imaging (fMRI) and TMS, Baumgartner et al. (2011) showed that both DLPFC and posterior VMPFC (pVMPFC), and their connectivity, contribute to the evaluation of unfair offers and to the subsequent costly decision to reject them; specifically, they found that a decrease in the activation of the pVMPFC is associated with a decrease in the rejection rate of the unfair offers. This finding seems to be in contrast with the results from the patient studies mentioned above, which showed that the lesion of the VMPFC increases the rejection rate for unfair offers. The two studies use different methodologies and caution needs to be used when comparing them; however, the discrepancies in findings indicate that the specific role of each area is still under debate.

One of the reasons why the cognitive mechanisms underpinning the reaction to unfairness in the UG are not yet clear, is the fact that the traditional UG carries some confounds that prevent to isolate the basic mechanisms that shape this complex behavioral response. One important theoretical limitation of this game is that it is a self-centered task, in that the responder is always the target of unfairness; therefore, it is not clear whether rejections are driven by pure fairness considerations or they are an impulsive reaction to an unfair treatment that targeted the responder (see Civai, 2013 for a detailed discussion on the limitations of the traditional UG and the consequent possible misinterpretation of the results). To overcome this issue, we developed a modified version of the UG, in which we asked our participants to play as the

responders both for themselves (myself—MS condition), as in the traditional UG, and on behalf of an unknown third-party (third-party—TP condition): because in TP, the responder has to evaluate unfairness without being the target of the unfair treatment, this condition helps shed light on the preference for equality by ruling out the confound of the self-serving bias that is present in the MS condition, i.e. the traditional UG. Results strongly suggest that people actually care about fairness, and they are ready to sacrifice resources to establish an equal distribution; in fact, no difference between MS and TP conditions is found, at least when disadvantageous unfairness is considered (Civai et al., 2010; Corradi-Dell'Acqua et al., 2013). However, despite no difference in the behavioral patter, the process has different neural correlates: when participants are the targets of the unfair treatment (MS condition), a stronger negative emotional arousal (Civai et al., 2010) and a higher activation in the MPFC (Civai et al., 2012; Corradi-Dell'Acqua et al., 2013) are observed when rejecting unfair offers. These effects are not observed in TP. Interestingly, a recent event-related potential study supports the idea of a different neural sensitivity for MS and TP: the authors found that the amplitude of the medial frontal negativity is modulated by the level of fairness of the offers only when the subject is the target of the unfair treatment, but not when the target is a TP (Alexopoulos et al., 2012). Thus, although previous accounts have related the increase in emotional arousal and in the activation of MPFC to a more general perception of unfairness (e.g. Sanfey et al., 2003; Güroğlu et al., 2010), the findings just discussed strongly suggest that these results are better ascribed to the personal involvement of the participant in the unfair exchange, as they are specific to the MS condition, rather than to a general evaluation of unfairness.

In this study, we aim at further testing whether the role of the MPFC is specific to the personal damage derived from unfair treatment rather than to a general reaction to unfairness. To do so, we used tDCS, aiming to induce changes in the activity of MPFC, with the final goal to alter the subject performance during of UG. tDCS is a technique where a direct current of low-level intensity is applied for a few minutes via electrodes placed on the subject's scalp. This current reaches the cortex and modulates the membrane polarity of neurons within a region of underlying neural tissue. tDCS-induced changes during stimulation result from changes in the permeability of the neural membrane, which is depolarized by anodal stimulation and hyperpolarized by cathodal stimulation (Nitsche and Paulus, 2000; Utz et al., 2010; Stagg and Nitsche, 2011). Therefore, tDCS can transiently influence behavior by altering neuronal activity, which may have facilitatory or inhibitory behavioral effects. Since our previous fMRI findings suggest that an increase in the MPFC activation is specifically associated with rejections of unfair offers when the participant's own outcome is at stake, this study aims at confirming the involvement of this area: specifically, our hypothesis is that the stimulation of MPFC by cathodal tDCS, compared with sham (placebo), would significantly reduce the probability of rejecting unfair offers in MS but not in TP by decreasing the neural activation in the area, and we chose to use cathodal stimulation because we were interested in the effects of reducing the functionality of MPFC. Moreover, because a previous study investigated the effect of cathodal tDCS applied over the DLPFC on the UG behavior (Knoch et al., 2008), we wanted to keep the stimulation consistent in order to be able to compare and integrate the findings from both works.

MATERIALS AND METHODS
Participants
Forty-eight participants (all females, between 18 and 25 years), with no history of neurological disease or psychiatric disorder, were recruited among undergraduate and graduate student population at the University of Trieste. They were all paid 10 euros for their participation. All participants signed an informed consent before starting the experiment, and according to the Declaration of Helsinki, all procedures were approved by SISSA Ethics Committee. People with history of mental illness, neurological disorders, such as epilepsy, under psychopharmacological medications, or with previous knowledge of the UG were excluded. Safety procedures were used in accordance with NIBS indications (Poreisz et al., 2007; Rossi et al., 2009), The data of eight participants were excluded from the analysis: two of them did not believe the cover story, one had previous knowledge of the game but did not inform the experimenters beforehand, and for five of them there were technical difficulties with the tDCS. Overall, the data of 40 subjects were taken into consideration for the analysis.

Task and procedure
Task
All participants were required to play as responders in the modified TP UG, very similar to the one described in Corradi-Dell'Acqua et al. (2013; Figure 1A). They had to accept or reject offers by pressing two different keys on a computer keyboard, using the index and the middle fingers of the right hand. Participants were told that the offers presented had been decided by a proposer in a previous session, thus making clear that their choices could not have influenced the course of the game. Participants were told that they, and their proposers, would have been paid a percentage of all the money gained throughout the whole game; actually, the proposer did not exist, participants were facing offers established a priori by the experimenter, and they were paid a fixed amount of 10 euros. Participants responded to 84 UG offers: 42 offers for the MS and 42 for the TP condition. They knew that, in the MS condition, their choice would have affected only their final payoff, whereas in the TP condition their choice would have affected exclusively the payoff of the responder in the subsequent experimental session. They were told that the TP would have received a percentage of the money that they gained on his/her behalf; in addition, they were told that part of their own final payoff depended on what the previous responder had decided on their behalf. The amount offered varied, both for MS and for TP, between 1 and 5 euros out of 10: 1–2 euros were considered unfair offers (16 trials), 3 euros mid-value offers (10 trials) and 4–5 euros fair offers (16 trials). To control for the social interactive nature of the UG, a control task was introduced, called Free Win (FW): here, participants had to accept or reject money given them 'for free' by the computer. The amounts were the same as in the UG (42 offers between 1 and 5 euros), but the rules were different, and much simpler: in the FW, accepting meant getting the money, without further implications, and rejecting meant not getting the money; for this reason, when referring to the FW, we use 'low offers' and 'high offers' as the equivalent for 'unfair offers' and 'fair offers'. For each experimental trial, participants were presented with the offer, and they were required to make a choice within 4 s. Examples of trials for each game condition: UG_MS: 'I offer you 2 euros out of 10'; UG_TP: 'I offer the next participant 2 euros out of 10'; FW: 'The computer gives you 2 euros' (Figure 1B).

Procedure.
Participants were screened for exclusion criteria over the phone or via e-mail, and informed on the nature of the experiment, in particular as far as the tDCS methodology was concerned. The day of the experiment, they were tested individually. They read the instructions of the game and then underwent the preparation.

A constant current flow of 2 mA was delivered by a battery-driven stimulator (Eldith, NeuroConn) through a pair of a saline-soaked sponge electrodes (35 cm^2; current density: 0.057 mA/cm^2).
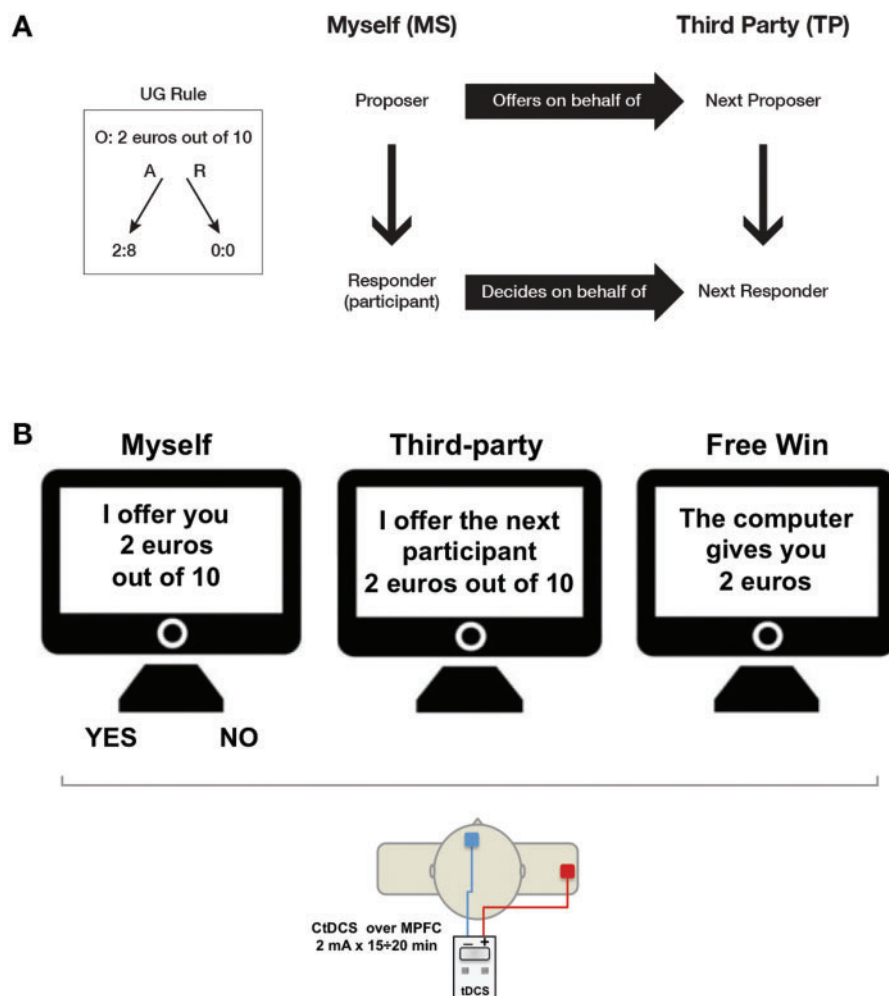
Fig. 1 (**A**) The structure of the modified third-party UG is described, in the same fashion as Corradi-Dell'Acqua *et al.* (2013). The UG rule is described: 'O' stands for Offer of the proposer. 'A' and 'R' stands for Accept and Reject, and the small arrows indicate the two UG outcomes: if the responder accept, the money is divided as the proposer decided, which is, in this case, 2 euros to the responder and 8 euros to the proposer; if the responder rejects, they both get zero. The same rule applies for the TP condition, where the outcome affects the next participants. (**B**) The tDCS placement (cathodal electrode over the Fpz-Fp1 site, and extra-encephalic reference on the right shoulder) is shown, together with three example trials representing the three different conditions: MS, when the outcome of the participant's decision influenced the participant's payoff; TP, when the outcome of the participant's decision influenced an anonymous third-party's payoff; FW, a control condition in which the computer assigned some money to the participant for free.

The 'active' cathodal electrode was placed over the MPFC (reference MNI coordinates: −2, 58, 8), on an area between Fpz and Fp1 in the international 10–20 nomenclature for Electroencephalography (EEG) electrode positioning. These coordinates were chosen according to Corradi-Dell'Acqua *et al.* (2013) findings, in which the activation of the MPFC was higher for rejections rather than acceptances, in UG_MS offers, but not in UG_TP. The 'reference' electrode was fixed extra-cephalically on the right arm (Figure 1B). It is important to specify that the tDCS is not focal, and thus it is possible that the effects of the stimulation are more widespread and not clearly confined to the area identified by an imaging study; however, it can be assumed that the area under the electrode is likely to be the most affected by the stimulation. We used an extra-cephalic reference to avoid interference effects from brain areas beneath the reference electrode. The electrodes were kept in place with elastic bands, and an electro-conductive gel was applied under the electrodes to help reduce impedance to the electrical current. In the cathodal tDCS condition, the current was applied from 2 min before the starting of the task, and lasted until the end of it; overall, each participant received 15–20 min of stimulation. The current was ramped up and down over the first and last 10 s of stimulation

in the cathodal tDCS condition, but only during the first and last 30 s (10 fade-in phase, 10 at level and 10 fade-out phase) in the sham tDCS condition (Gandiga *et al.*, 2006). After the experimental session, all participants completed a questionnaire on the tDCS-induced sensations (Fertonani *et al.*, 2010) to determine whether the stimulation protocol (cathodal *vs* sham) affected the sensations experienced that could potentially influence the subject performance; no differences were found between groups.

RESULTS

A generalized linear mixed model (GLMM) was used to analyze the effects of the within-subject factors Target (MS, TP) and Fairness Level (unfair, mid-value, fair), and the between-subject factor Stimulation (cathodal and sham) on the binary response (accept/reject), with Subjects as a random factor. A GLMM was chosen to model the relationship between a categorical response and the explanatory variables as a logistic regression, accounting for random effects elicited by the interdependence of responses that come from the same subject. The parameter estimates of the model indicate logarithmic odds of the logit model. The analysis was implemented in RStudio, version 0.98.501,

Table 1 Marginal effects for the model in which fairness level, stimulation, target and their interactions are the predictors of the likelihood of rejection

	Marginal Effects (SE)
Fairness Level	−0.527 (0.03)***
Stimulation	−0.146 (0.07)*
Target	−0.027 (0.03)
Fairness Level × Stimulation	0.150 (0.04)***
Fairness Level × Target	0.123 (0.04)***
Target × Stimulation	0.152 (0.04)***
Fairness Level × Target × Stimulation	−0.020 (0.06)
N	3355

The values in brackets refer to the standard error (SE), and the stars refer to the level of significance in the linear model (*$P < 0.05$, ***$P < 0.0001$). N is the number of observations.

'lme4' package (Bates *et al.*, 2014; lme4: Linear mixed-effects models using Eigen and S4. R package version 1.1-6. http://CRAN.R-project.org/package=lme4). To clarify the magnitude of the increment in the probability of rejection, Table 1 shows the standard coefficients of a linear mixed model, which indicate how much, in percentage, the probability of rejection increases when each predictor increases of one unit. Because the magnitude and the direction of the effects for mid-value and fair offers overlap, we have pooled the data and consider these conditions as one, resulting in unfair *vs* mid-value + fair offers.

A significant main effect of Fairness Level ($\beta = -3.299$, $P < 0.001$) was found, indicating that an increase in fairness decreases the probability of rejecting the offer. Moreover, three significant positive two-way interactions were found: (1) Fairness Level × Stimulation ($\beta = 0.992$, $P < 0.001$; Figure 2A) and (2) Target × Stimulation ($\beta = 0.870$, $P < 0.001$) (Figure 2B), indicating that the likelihood of rejection for fairer offers (interaction 1) TP (interaction 2) increases in cathodal tDCS, but not in sham and (3) Fairness Level × Target ($\beta = 0.952$, $P < 0.001$), showing that the chance of rejecting fair offers increases in TP, but not in MS.

Given that no three-way interaction was found, we are unable to determine, from this first analysis, which are the specific effects that drive the interaction, i.e. whether it is a decrease in the likelihood of rejection in MS or an increase in TP to determine the effects. Therefore, to investigate these effects more in depth, MS and TP conditions were considered separately: a positive Fairness Level × Stimulation interaction was found both in MS, where it approaches significance ($\beta = 0.642$, $P = 0.058$), and in TP ($\beta = 0.874$, $P < 0.01$), indicating that, for both targets, the likelihood of rejection increases for mid-value and fair offers, with respect to unfair offers, more during cathodal tDCS than during sham. Linear parameter estimates are presented in Table 2. Figure 3A,B report the rejection rate for each Target condition, clustered by Fairness Level: as the bar plots show, the significant interaction in MS is driven by the difference between stimulations in unfair offers, where cathodal stimulation decreases the rejections with respect to sham, whereas in TP, the interaction is driven by the difference in mid-value and fair offers, where cathodal stimulation increases rejections with respect to sham. We have also analysed separately the two simple effects of stimulations, and we found a significant main effect of Fairness Level ($\beta = -3.317$, $P < 0.001$) and a significant interaction Target × Fairness Level ($\beta = 0.956$, $P < 0.001$) for the sham condition, indicating that participants were more likely (14% increase in probability) to reject TP fair offers compared with MS, an effect that has never been found in our previous studies; in the cathodal stimulation, we found a significant main effect of Fairness Level ($\beta = -2.299$, $P < 0.001$) and a significant
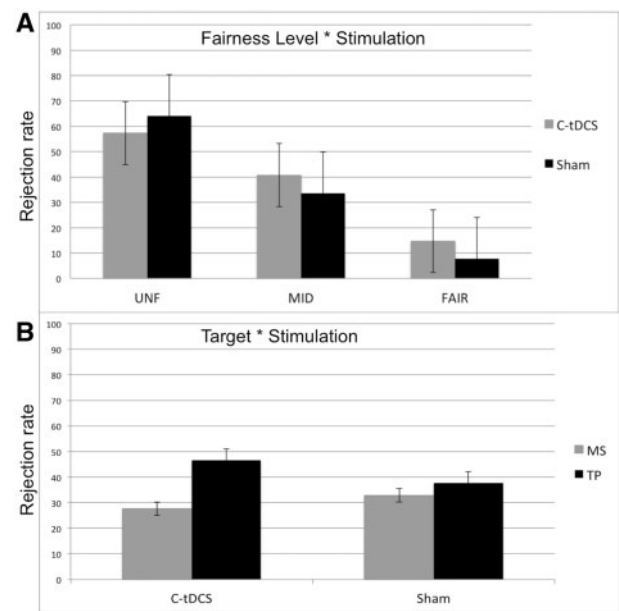
Fig. 2 (**A**) This bar plot shows the rejection rate and the standard error (*y*-axis) for the significant interaction Fairness Level × Stimulation: the gray bars refer to cathodal stimulation (C-tDCS), and the black bars to Sham. The Fairness Level is plotted on the x-ax, (UNF = unfair offers, MID = mid-value offers, FAIR = fair offers). (**B**) This bar plot shows the rejection rate and the standard error (*y*-axis) for the significant interaction Target × Stimulation: the gray bars refer to MS condition, and the black bars to TP. The Stimulation is plotted on the x-as, (C-tDCS = cathodal stimulation, Sham = sham)

Table 2 Marginal effects for both myself (MS) and third-party (TP) models in which fairness level, stimulation and their interaction are the predictors of the likelihood of rejection

	Marginal Effects for MS (SE)	Marginal Effects for TP (SE)
Fairness Level	−0.526 (0.03)***	−0.404 (0.03)***
Stimulation	−0.146 (0.07)*	0.006 (0.08)
Fairness Level*Stimulation	0.149 (0.04)***	0.130 (0.04)**
N	1679	1676

The values in brackets refer to the standard error (SE), and the stars refer to the level of significance in the linear model (*$P < 0.05$, **$P < 0.01$, ***$P < 0.001$). N is the number of observations.

main effect of Target ($\beta = 0.694$, $P < 0.001$), suggesting that the likelihood of rejections was higher (13% increase in probability) for TP as opposed to MS not only for all types of offers but also for the fair offers in the sham condition. In addition, we performed a linear mixed model analysis considering the trial onset time as a predictor, and we found no significant effects.

An analysis to control for the effect of the social interactive nature of the UG was also conducted, considering Task (2 levels: UG, and the control task FW) as an additional variable. We found a significant negative main effect of Task ($\beta = -2.949$, $P < 0.001$), indicating that the probability of rejecting an offer was higher in UG compared with FW, and a significant negative main effect of Fairness Level ($\beta = -3.226$, $P < 0.001$), indicating that unfair/low offers were more likely to be rejected than mid-value and fair/high offers; moreover, significant interactions Task × Stimulation ($\beta = 0.712$, $P < 0.05$), Task × Fairness Level ($\beta = 0.831$, $P < 0.05$), and Stimulation × Fairness Level ($\beta = 0.713$, $P < 0.05$) were found, indicating a higher probability of rejections, respectively, for the UG in the cathodal stimulation, for the UG in the mid-value and fair offers, and for the mid-value and fair/high offers in the cathodal stimulation.
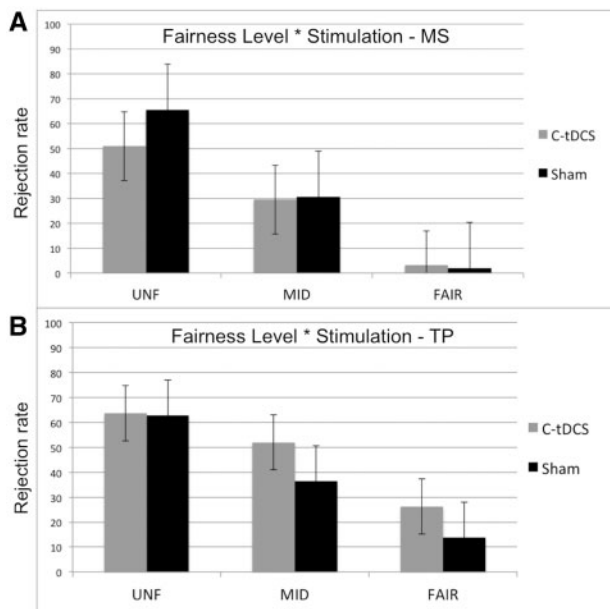
Fig. 3 These bar plots report, for the simple effects of target, the rejection rate and the standard error (*y*-axis) for cathodal (C-tDCS, gray) and sham (black) stimulations, for each Fairness Level (*x*-axis: UNF = unfair offers, MID = mid-value offers, FAIR = fair offers), plotted on *x*. The upper bar plot (**A**) reports the results for MS; the lower bar plot (**B**) reports the results for TP.

No significant three-way interaction was found. Figure 4A shows the rejection rate for the significant interaction Task × Stimulation. Considering the FW alone, a main effect of Fairness Level ($\beta = -2.748$, $P < 0.001$) was found, indicating that low offers were more likely to be rejected than higher offers (Figure 4B). This effect was found in a previous study (Civai *et al.*, 2012) and has been interpreted as a possible carry-on effect from the UG. No significant difference between stimulations was found for FW.

DISCUSSION

In a previous fMRI study (Corradi-Dell'Acqua *et al.*, 2013), an increase in the activation of MPFC, within a UG context, was specifically associated with the rejection unfair offers when the self was targeted, rather than to a more general unfairness evaluation. In this work, we find converging evidence using a complementary technique (i.e. tDCS), by means of which we modulate the behavior of UG responders through a stimulation applied over the MPFC.

First of all, we shall briefly discuss the effects of cathodal tDCS. While the concept of cathodal stimulation worsening performance seems well established for tDCS in the motor system (Stagg and Nitsche, 2011), the same concept is not so directly applicable in the cognitive neuroscience field and the relation between type of stimulation and final behavior is often quite complex (e.g. Jacobson *et al.*, 2012; Miniussi *et al.*, 2013). We should specify that, in the case of the UG, it is quite hard to determine if a deviation from the normal behavior consists in an impaired response or in an improved performance: we do know the normal baseline response, and thus, we can recognize a deviation from the norm, but depending on what theory is embraced, the optimal response varies. For example, for the classic utility theory, the optimal response is 0% of rejection, while this would represent a dysfunctional response in a theoretical frame that considers rejections to be triggered by a normal negative emotional reaction. Here we assumed that, based on previous literature on the motor functions, but not only, cathodal tDCS would lead to a temporary
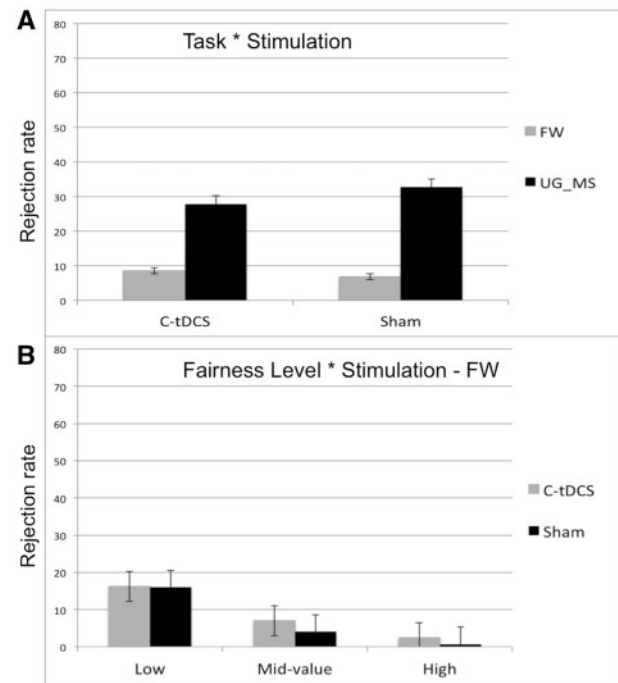
Fig. 4 (**A**) This bar plot shows the rejection rate and the standard error (*y*-axis) for the significant interaction Task × Stimulation: the gray bars refer to the control task FW, and the black bars the UG, Myself condition (UG_MS). The Stimulation is plotted on the x-ax (C-tDCS = cathodal stimulation, Sham = sham stimulation). (**B**) This bar plot shows the rejection rate (RR) and the standard error (SE) (*y*-axis) for the significant interaction Fairness Level*Stimulation, when considering the FW alone: the gray bars refer to the cathodal stimulation (C-tDCS) and the black bars to the Sham stimulation. The Fairness Level is plotted on the x-ax (Low = low offers, Mid = mid-value offers, High = high offers).

decreasing of the functionality of the MPFC; however, the only way to determine for sure what happens to the functionality of this area would have been to collect imaging data during, or soon after, the stimulation. Further studies should combine these techniques to give more detailed answers.

The current data show that a cathodal stimulation over the MPFC leads to a significant difference in the likelihood of rejection between the condition in which the participant is the target of the unfair treatment, and the one in which the target is a TP, being the likelihood of rejection lower for the former. Interestingly, when the two simple effects of Target (myself and TP) were investigated separately to understand whether the difference during cathodal tDCS was driven by a decrease in the rejections for MS or by an increase in the rejection for TP, the results show that when the participant is the target of the unfair treatment, there is a decrease of the likelihood of rejection for unfair offers, whereas when the target is a TP, there is an increase in rejections for mid-value and fair offers. Importantly, these findings support our initial hypothesis that the MPFC plays a specific role in evaluating the self-damaging unfairness rather than the broad concept of unfairness, as the stimulation decreased rejections of unfair offers in MS but not in TP. This result suggests that, when the first person is involved, the MPFC is crucial for integrating emotional, rational and social information (Amodio and Frith, 2006), particularly in the presence of contrasting self- and social interests (Koban *et al.*, 2014); when the cortical excitability of this area is reduced, it is more difficult to overcome self-interest, leading to a decrease in the probability of rejection of the unfair offers.

It is also important to consider the low spatial focality of tDCS due to heterogeneous tissue conductivities (Nitsche and Paulus, 2011); we

believe that the effects of the tDCS in this study should be interpreted in terms of effects on the MPFC, as this is the area under the electrode, but because no stimulation was applied on other areas of the PFC, we cannot exclude whether the tDCS effect that we find might be valid also for other regions. In fact, Ruff et al. (2013), who show in a recent study that anodal and cathodal stimulations over the right lateral PFC have opposite effects on voluntary and sanction-induced social norm compliance-related behavior, claim that the norm compliance mechanism is 'probably not restricted to neural activity within this brain area [right lateral PFC, Ed.], given that the PFC is involved in many aspects of behavioral control and that brain stimulation can affect areas interconnected with the stimulation site' (p. 484). Interestingly, a study by Knoch et al. (2008) investigated the effects on UG rejections of cathodal tDCS applied over the right DLPFC, and reported results similar to ours, i.e. UG responders accepted more unfair offers with respect to sham stimulation. The overlapping of our findings with Knoch et al.'s findings suggests that the effect of tDCS on the reaction to unfair UG offers might not be confined to a specific area of PFC. However, the findings overlap only as far as the MS condition (or traditional UG) is concerned. Our study shows that cathodal tDCS over MPFC differently influences MS and TP, decreasing the likelihood of rejections for unfair offers in MS while increasing it for fairer offers in TP. This is important because it shows that, as far as the MPFC is concerned, the temporary reduction of the cortical excitability of this area causes people to accept unfairness that damages their own pockets not because they completely fail to apply a social norm, as shown by the normal acceptance rate of unfair offers in TP, but possibly because it is more difficult to overcome self-interest when self-interest and social norms collide. This is the same conclusion reached by Knoch et al. (2008). However, because these authors did not have a condition in which self-unrelated fairness concerns were investigated, i.e. the current TP condition, we cannot say whether the DLPFC, when temporarily altered via cathodal tDCS, fails to integrate self-interest and social norms when self-interest is at stake, as the MPFC, or it favors the pursuing of resources maximization rather than equality norm as the default response, which would be the case if cathodal tDCS on the DLPFC decreased rejections of unfair offers for TP as well as for MS. Given that the main goal of this study, was to investigate the differential role of MPFC on behavioral reaction to unfairness when either the self (MS) or an unknown TP were involved, rather than understanding the specific involvement of different PFC and non-PFC regions in this behavior, we choose not to target alternative areas. Further studies are needed to better address the functional specificity of the different areas of PFC.

There is also a possible alternative explanation for the involvement of the MPFC in UG rejections, according to which the stimulation over this area decreases the negative emotional reaction, elicited by unfairness, that accompanies rejections in MS (van't Wout et al., 2006; Civai et al., 2010), rather than decreasing the ability to control the greediness in favor of social rule compliance. Further studies are needed to disentangle these two interpretations. One possibility would be using physiological techniques, such as skin conductance response and self-reports, in combination with tDCS, to get a measure of emotional valence and arousal: a decrease in the physiological response during tDCS stimulation would suggest an explanation based on the decrease in negative emotional reaction.

The unpredicted result regarding the difference between stimulations in the TP condition, i.e. a higher probability of rejection of mid-value and fair offers in the cathode stimulation compared with the sham stimulation, might be ascribed to the well-known role of the MPFC in the self-other perception and judgment; in particular, dorsal MPFC has been repeatedly found to be associated with other-related judgment and perspective taking (Decety and Jackson, 2004;

D'Argembeau et al., 2007; see Denny et al., 2012, for a meta-analysis). The difference between stimulations in TP could be due to the fact that, in the cathodal stimulation, participants fail to relate to the TP and to identify him/her as a member of their group of responders. In our previous studies, the lack of difference in the rejection rate between MS and TP has been attributed to the in-group effect that was created by the awareness of the participants of being part of the same responders' group together with the TP (Zamir, 2001; Civai et al., 2010; Corradi-Dell'Acqua et al., 2013;). This interpretation has been supported by the findings presented in a recent study (Civai et al., 2013), where we describe an experiment in which the identity of the TP was manipulated, in that he/she was not presented as 'the next responder', like in this study, but as a participant who would not have taken part in the game: in this case, our participants tended to favor less the TP on behalf of whom they were deciding, possibly because they lacked the trigger for identifying him/her as an in-group member, and, as such, more similar to themselves. In this study, we show how cathodal tDCS applied over the MPFC seems to decrease this process of identification of the participant with the TP. Interestingly, a (possible) decrease in perspective taking lead to an increase in the likelihood of rejections, especially for the mid and fair offers. This suggests that rejecting unfairness, especially when it is not very severe, such as in the case of mid-value offers, is less hard when there is no conflict with self-interested motivations; when the participant is directly involved or when he/she is more prone to identify him/herself with the TP, the likelihood of rejections decreases for fairer offers. However, this interpretation is speculative, as we do not have enough data to support it; further investigations using the TP manipulation as an ingroup or outgroup member with respect to the responder, and assessing perspective-taking, are needed to shed light on these findings and to draw stronger conclusions.

An alternative explanation, which could also account for the increase in the probability of rejections of fair offers in the sham condition, is that some of the participants might have been more competitive than others, and might have rejected fair offers on behalf of the TP fearing that, otherwise, the TP would have ended up with more money than them; however, this is a speculation, as we did not asked our participants to provide a description of the decision-making strategy, if any, that was used. In fact, this result might also be ascribed to the small sample size, common to many social cognitive neuroscience studies, which prevent to perform an analysis of individual differences that could be very informative, in particular when analyzing complex behaviors, such as unfairness reactions. In the future, we are planning to continue the investigation of fairness perception and equality norms using larger samples of participants and by collecting more measurements of individual differences, to build more precise models of social decision-making.

In conclusion, we carried out a study in which we aimed at confirming the involvement of MPFC in the evaluation of personal damaging unfairness, using the tDCS, a technique that allowed the modulation of cortical activation: we confirmed our hypothesis that tDCS of the MPFC decreases the likelihood of rejections of unfair offers in MS but not in TP. Moreover, we found interesting and unpredicted results as far as the TP condition is considered, suggesting, although speculatively, that tDCS of this area also modulates the ability to take other's perspective, as extensively reported in previous literature. Future research is needed in order to incorporate individual differences, as, for example, personality traits, strategic thinking and prosocial/individualistic tendencies to develop better social decision-making models.

Conflict of Interest

None declared.

REFERENCES

Alexopoulos, J., Pfabigan, D.M., Lamm, C., Bauer, H., Fischmeister, F.P.S. (2012). Do we care about the powerless third? An ERP study of the three-person ultimatum game. *Frontiers in Human Neuroscience*, *6*, 59.

Amodio, D.M., Frith, C.D. (2006). Meeting of minds: the medial frontal cortex and social cognition. *Nature Reviews of Neuroscience*, *7*, 268–77.

Bates, D., Maechler, M., Bolker, B., Walker, S. (2014). lme4: linear mixed-effects models using Eigen and S4. R package version 1.1-6. http://CRAN.R-project.org/package=lme4. (accessed 19 May 2014).

Baumgartner, T., Knoch, D., Hotz, P., Eisenegger, C., Fehr, E. (2011). Dorsolateral and ventromedial prefrontal cortex orchestrate normative choice. *Nature Neuroscience*, *14*, 1468–74.

Boksem, M.A., De Cremer, D. (2010). Fairness concerns predict medial frontal negativity amplitude in ultimatum bargaining. *Social Neuroscience*, *5*, 118–28.

Bolton, G.E. (1991). A comparative model of bargaining: theory and evidence. *American Economic Review*, *81*, 1096–136.

Bolton, G.E., Ockenfels, A. (2000). ERC: a theory of equity, reciprocity, and competition. *American Economic Review*, *90*, 166–93.

Campanhā, C., Minati, L., Fregni, F., Boggio, P.S. (2011). Responding to unfair offers made by a friend: neuroelectrical activity changes in the anterior medial prefrontal cortex. *The Journal of Neuroscience*, *31*, 15569–74.

Chang, L.J., Sanfey, A.G. (2013). Great expectations: neural computations underlying the use of social norms in decision-making. *Social cognitive and affective neuroscience*, *8*(3), 277–84.

Civai, C. (2013). Rejecting unfairness: emotion-driven reaction or cognitive heuristic? *Frontiers in Human Neuroscience*, *7*, 126.

Civai, C., Corradi-Dell'Acqua, C., Gamer, M., Rumiati, R.I. (2010). Are irrational reactions to unfairness truly emotionally-driven? Dissociated behavioural and emotional responses in the Ultimatum Game task. *Cognition*, *114*, 89–95.

Civai, C., Crescentini, C., Rustichini, A., Rumiati, R.I. (2012). Equality versus self-interest in the brain: differential roles of anterior insula and medial prefrontal cortex. *Neuroimage*, *62*, 102–12.

Civai, C., Rumiati, R.I., Rustichini, A. (2013). More equal than others: equity norms as an integration of cognitive heuristics and contextual cues in bargaining games. *Acta Psychologica*, *144*, 12–8.

Corradi-Dell'Acqua, C., Civai, C., Rumiati, R.I., Fink, G.R. (2013). Disentangling self-and fairness-related neural mechanisms involved in the ultimatum game: an fMRI study. *Social Cognitive and Affective Neuroscience*, *8*, 424–31.

D'Argembeau, A., Ruby, P., Collette, F., Degueldre, C., Balteau, E., Luxen, A., Maquet, P., Salmon, E. (2007). Distinct regions of the medial prefrontal cortex are associated with self-referential processing and perspective taking. *Journal of Cognitive Neuroscience*, *19*, 935–44.

Decety, J., Jackson, P.L. (2004). The functional architecture of human empathy. *Behavioral and Cognitive Neuroscience Reviews*, *3*, 71–100.

Denny, B.T., Kober, H., Wager, T.D., Ochsner, K.N. (2012). A meta-analysis of functional neuroimaging studies of self-and other judgments reveals a spatial gradient for mentalizing in medial prefrontal cortex. *Journal of Cognitive Neuroscience*, *24*, 1742–52.

Falk, A., Fischbacher, U. (2006). A theory of reciprocity. *Game and Economic Behavior*, *54*, 293–315.

Fehr, E., Schmidt, K.M. (1999). A theory of fairness, competition, and cooperation. *Quarterly Journal of Economics*, *114*, 817–68.

Fertonani, A., Rosini, S., Cotelli, M., Rossini, P.M., Miniussi, C. (2010). Naming facilitation induced by transcranial direct current stimulation. *Behavioural Brain Research*, *208*, 311–18.

Gandiga, P.C., Hummel, F.C., Cohen, L.G. (2006). Transcranial DC stimulation (tDCS): a tool for double-blind sham-controlled clinical studies in brain stimulation. *Clinical Neurophysiology*, *117*, 845–50.

Güroğlu, B., van den Bos, W., Rombouts, S.A.R.B., Crone, E.A. (2010). Unfair? It depends: neural correlates of fairness in social context. *Social Cognitive and Affective Neuroscience*, *5*, 414–23.

Güth, W., Schmittberger, R., Schwarze, B. (1982). An experimental analysis of ultimatum bargaining. *Journal of Economic Behavior and Organization*, *3*, 367–88.

Jacobson, L., Koslowsky, M., Lavidor, M. (2012). tDCS polarity effects in motor and cognitive domains: a meta-analytical review. *Experimental Brain Reserch*, *216*, 1–10.

Knoch, D., Nitsche, M.A., Fischbacher, U., Eisenegger, C., Pascual-Leone, A., Fehr, E. (2008). Studying the neurobiology of social interaction with transcranial direct current stimulation—the example of punishing unfairness. *Cerebral Cortex*, *18*, 1987–90.

Knoch, D., Pascual-Leone, A., Meyer, K., Treyer, V., Fehr, E. (2006). Diminishing reciprocal fairness by disrupting the right prefrontal cortex. *Science*, *314*, 829–32.

Koban, L., Pichon, S., Vuilleumier, P. (2014). Responses of medial and ventrolateral prefrontal cortex to interpersonal conflict for resources. *Social Cognitive and Affective Neuroscience*, *9*, 561–9.

Koenigs, M., Tranel, D. (2007). Irrational economic decision-making after ventromedial prefrontal damage: evidence from the Ultimatum Game. *The Journal of Neuroscience*, *27*, 951–6.

Miniussi, C., Harris, J.A., Ruzzoli, M. (2013). Modelling non-invasive brain stimulation in cognitive neuroscience. *Neuroscience Biobehavioural Rewiew*, *37*, 1702–12.

Moretti, L., Dragone, D., Di Pellegrino, G. (2009). Reward and social valuation deficits following ventromedial prefrontal damage. *Journal of Cognitive Neuroscience*, *21*, 128–40.

Nitsche, M.A., Paulus, W. (2000). Excitability changes induced in the human motor cortex by weak transcranial direct current stimulation. *The Journal of Physiology*, *527*, 633–9.

Nitsche, M.A., Paulus, W. (2011). Transcranial direct current stimulation—update 2011. *Restorative neurology and neuroscience*, *29*, 463–92.

Pillutla, M.M., Murnighan, J.K. (1996). Unfairness, anger, and spite: emotional rejections of ultimatum offers. *Organizational Behavior and Human Decision Processes*, *68*, 208–24.

Poreisz, C., Boros, K., Antal, A., Paulus, W. (2007). Safety aspects of transcranial direct current stimulation concerning healthy subjects and patients. *Brain Research Bulletin*, *72*, 208–14.

Rabin, M. (1993). Incorporating fairness into game theory and economics. *American Economic Review*, *83*, 1281–302.

Rossi, S., Hallett, M., Rossini, P.M., Pascual-Leone, A., & the Safety of TMS Consensus Group (2009). Safety, ethical considerations, and application guidelines for the use of transcranial magnetic stimulation in clinical practice and research. *Clinical Neurophysiology*, *120*, 2008–39.

Ruff, C.C., Ugazio, G., Fehr, E. (2013). Changing social norm compliance with noninvasive brain stimulation. *Science*, *342*, 482–4.

Sanfey, A.G., Rilling, J.K., Aronson, J.A., Nystrom, L.E., Cohen, J.D. (2003). The neural basis of economic decision-making in the ultimatum game. *Science*, *300*, 1755–8.

Stagg, C.J., Nitsche, M.A. (2011). Physiological basis of transcranial direct current stimulation. *The Neuroscientist*, *17*, 37–53.

Tabibnia, G., Satpute, A.B., Lieberman, M.D. (2008). The sunny side of fairness preference for fairness activates reward circuitry (and disregarding unfairness activates self-control circuitry). *Psychological Science*, *19*, 339–47.

Utz, K.S., Dimova, V., Oppenländer, K., Kerkhoff, G. (2010). Electrified minds: transcranial direct current stimulation (tDCS) and galvanic vestibular stimulation (GVS) as methods of non-invasive brain stimulation in neuropsychology—a review of current data and future implications. *Neuropsychologia*, *48*, 2789–810.

van't Wout, M., Kahn, R.S., Sanfey, A.G., Aleman, A. (2005). Repetitive transcranial magnetic stimulation over the right dorsolateral prefrontal cortex affects strategic decision-making. *Neuroreport*, *16*, 1849–52.

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*, 564–8.

Zamir, S. (2001). Rationality and emotions in ultimatum bargaining. *Annals of Economics and Statistics/Annales d'Économie et de Statistique*, *61*, 1–31.