SOCIAL NEUROSCIENCE, 2016
http://dx.doi.org/10.1080/17470919.2016.1252795

Routledge
Taylor & Francis Group

Accepting unfairness by a significant other is associated with reduced connectivity between medial prefrontal and dorsal anterior cingulate cortex

Ramzi Fatfouta [a,b*], Dar Meshi[a,b], Angela Merkl[b,c] and Hauke R. Heekeren[a,b]

[a]Department of Education and Psychology, Freie Universität Berlin, Berlin, Germany; [b]Cluster "Languages of Emotion", Berlin, Germany; [c]Department of Neurology with Experimental Neurology, Charité-Universitätsmedizin, Berlin, Germany

ABSTRACT
Conflict is a ubiquitous feature of interpersonal relationships, yet many of these relationships preserve their value following conflict. Our ability to refrain from punishment despite the occurrence of conflict is a characteristic of human beings. Using a combination of behavioral and neuroimaging techniques, we show that prosocial decision-making is modulated by relationship closeness. In an iterated social exchange, participants were more likely to cooperate with their partner compared to an unknown person by accepting unfair exchanges. Importantly, this effect was not influenced by how resources were actually being shared with one's partner. The medial prefrontal cortex (MPFC) was activated when the partner, rather than the unknown person, behaved unfairly and, in the same context, the MPFC demonstrated greater functional connectivity with the dorsal anterior cingulate cortex (DACC). MPFC–DACC connectivity was inversely associated with participants' tendency to "forgive" their partner for unfairness as well as performance outside the scanner on a behavioral measure of forgiveness. We conclude that relationship closeness modulates a neural network comprising the MPFC/DACC during economic exchanges.

ARTICLE HISTORY
Received 14 June 2016
Revised 18 October 2016
Published online
14 November 2016

KEYWORDS
Decision-making; interpersonal relationships; ultimatum game; social cognition; medial prefrontal cortex (MPFC)

In a world where conflict pervades social life, humans must be able to successfully negotiate with others. Essential to adaptive behavior is our ability to maintain interpersonal relationships even after conflict (De Waal & Pokorny, 2005). This capacity to refrain from retaliation and, instead, to motivate prosocial behavior toward conspecifics despite their harmful prior actions is a unique human feature (McCullough, Kurzban, & Tabak, 2013). From the perspective of natural selection, such behavior likely evolved because it allowed the repair of valuable relationships—that is, relationships whose repair would be anticipated to yield potential long-term benefits (McCullough, 2008). Research on nonhuman primates, for instance, indicates that prosocial behavior is most common among genetic relatives and close associates, given their potential contributions to each other's fitness (De Waal & Pokorny, 2005; Koski, Koops, & Sterck, 2007). Similarly, in humans, behavioral studies suggest that we are more inclined to act prosocially toward someone to whom we feel close and committed (Karremans & Aarts, 2007; Karremans et al., 2011). This is exemplified in socioeconomic games, where prosocial behavior has been defined as the tendency to forgo one's desire to punish noncooperative behavior (Li & Chen, 2012); people are more likely to tolerate unfairness when playing with a close friend (Campanhã, Minati, Fregni, & Boggio, 2011), or when primed with the concept of closeness (Fatfouta, Schulreich, Meshi, & Heekeren, 2015).

Neuroimaging studies investigating neural responses to transgressions (Farrow et al., 2005, 2001; Hayashi et al., 2010; Ricciardi et al., 2013; Strang, Utikal, Fischbacher, Weber, & Falk, 2014; Young & Saxe, 2009) have found activation in areas previously implicated in mental-state reasoning—known as mentalizing—and cognitive conflict (medial prefrontal and anterior cingulate cortical areas). However, none of these studies examined how these areas influence decision-making when faced with actual transgressions—in particular, in the context of close relationships. Furthermore, dealing with transgressions is a complex psychological function (Worthington, 2006) and, hence, unlikely to be implemented by a single brain area, but rather by a network of interacting brain areas. Explaining such a neurobiological mechanism remains an outstanding challenge.

CONTACT Ramzi Fatfouta r.fatfouta@gmail.com Department of Psychology, University of Potsdam, Karl-Liebknecht-Str. 24/25 14476 Potsdam, Germany
*Present address: Department of Psychology, University of Potsdam.
Supplemental data for this article can be accessed here

To fill this gap, we extend previous work in three important ways: First, we examined brain regions involved when humans face transgressions by a close other (i.e., their partner) and by a non-close other (i.e., an unknown person). Second, we investigated individuals' changes in brain activity as they interact with these people in a real, consequential social scenario. Third, we analyzed the relation between brain activation and individual differences in forgiveness, as assessed by a commonly used forgiveness scale and a newly developed behavioral test (Fatfouta, Schröder-Abé, & Merkl, 2014). We hypothesized that relationship closeness would modulate behavioral and neural responses to transgressions, resulting in greater acceptance of unfairness during interactions with one's partner compared to interactions with an unknown person. More specifically, the anterior medial prefrontal cortex (MPFC) is sensitive to relationship closeness in the context of unfair behavior (Campanhã et al., 2011) and responds to personally relevant (i.e., close) others (Krienen, Tu, & Buckner, 2010). We therefore expected increased MPFC activity in a comparison of unfair treatment by one's partner versus an unknown person. Notably, conflict is detected in the dorsal anterior cingulate cortex (DACC), with the MPFC playing a role in regulating such conflict (Nakao et al., 2010). We hypothesized that refraining from punishing unfair treatment would involve interactions between the MPFC and regions modulated by conflict between competing responses (i.e., accept vs. reject), particularly the DACC.

Method

Overview

To test these hypotheses, we adopted a game-theoretic approach that allows the simulation of an interactive social exchange. More specifically, we had 23 romantically involved couples (46 participants; see later) play a version of the Ultimatum Game (UG; Güth, Schmittberger, & Schwarze, 1982). In the UG, one individual (the proposer) proposes how to split a sum of money with another individual (the responder). If the responder accepts, the sum is divided according to the proposal. If the responder rejects, neither individual receives anything. In such a social-exchange game, transgressions are encoded as inequitable divisions of monetary rewards and, hence, entail real consequences for each individual. Typically, responders reject offers less than 30% of the proposer's endowment. In other words, they retaliate for the low offers they receive (Camerer, 2003; Crockett et al., 2013). By contrast, responders can also accept those offers—that is, they refrain from retaliation (i.e., "forgive") and act prosocially toward the proposer (McCullough, Worthington, & Rachal, 1997). A similar conceptualization has been employed in recent a neuroimaging study (Will, Crone, & Güroğlu, 2015). In our task, participants decided whether to accept or reject offers from their partner and an unknown person (Figure 1(a)). We also included a third condition, in which participants received offers from a computer (not relevant to the present study's research question). All offers were actually predetermined so as to control the number and size of offers made, which ranged from fair (proposing 5€ or 4€ out of 10€) to unfair (proposing 3€, 2€, or 1€ out of 10€) (see Koenigs & Tranel, 2007; Mehta & Beer, 2010). While playing the UG, one member of each couple (n = 23) was scanned using functional magnetic resonance imaging (fMRI), and we were specifically interested in comparing brain activations in the context of unfair offers from one's partner relative to unfair offers from an unknown person.

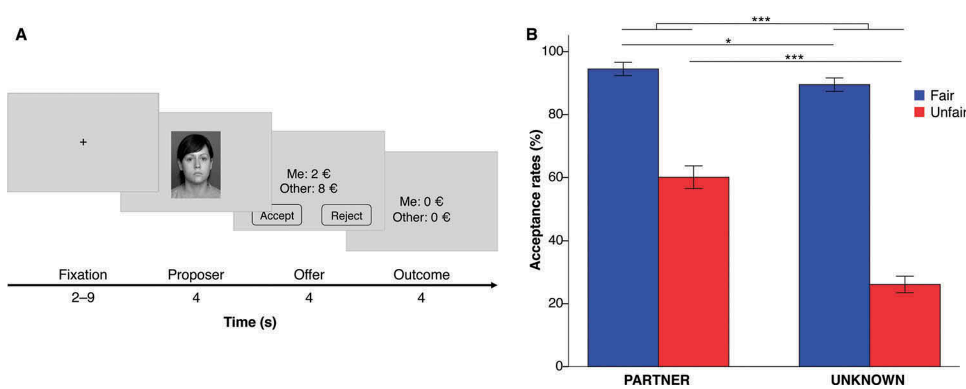

Figure 1. Experimental design and behavioral results. (a) Timeline of a single trial in the ultimatum game. In each trial, participants first viewed a fixation cross, then a photograph of the proposer, the offer, and then the outcome. While the offer was on screen, participants responded whether they accepted or rejected it by button press. (b) Mean acceptance rates (% of accepted offers) for fair and unfair offers from the partner and the unknown person. Error bars indicate one standard error of the mean, calculated within-subjects. ***P < 0.001 (two-tailed), *P < 0.05 (two-tailed).

Upon completion of the UG, all participants reported their subjective impressions by rating how fair, attractive, and trustworthy these proposers were. Participants also rated another person whom they had not interacted with but who had participated in the previous experimental session. This was done to confirm that the partners and the unknown persons were comparable regarding the aforementioned characteristics (see Supplementary Figure 1). Furthermore, we assessed participants' affective responses to unfairness, and additionally, individual differences in forgiveness (i.e., trait forgiveness), which were used to determine whether these individual differences were predicted by participants' brain activation. To explain, in addition to using a standard self-report (explicit) measure of forgiveness, we employed a computerized response-time task designed to assess individuals' automatic (implicit) associations regarding forgiveness. This task was an adaptation of the Implicit Association Test (IAT; Greenwald, McGhee, & Schwartz, 1998), in which participants had to categorize different combinations of self-related words and forgiveness- and retaliation-related words, as fast as possible. Automatic mental associations were inferred from participants' behavior (i.e., response times) on different trial types of this task (Fatfouta et al., 2014). The rationale for including such a behavioral test is that spontaneous brain responses are predictive of individuals' automatic associations rather than their explicit (i.e., self-reported) attitudes and beliefs (Stanley, Phelps, & Banaji, 2008). Finally, we assessed resource sharing among couples, so as to account for individual differences in how resources were actually being shared in real life.

Participants

Twenty-five heterosexual non-cohabiting couples took part in the study (N = 50). On average, couples had been together for 2.22 years (SD = 1.99); they maintained separate bank accounts, were unmarried, and had no children. All participants completed the experiment but only one member of each couple was scanned while playing the UG (see later). Four participants accepted every offer from both their partner and the unknown person and, hence, used a rule-based (vs. fairness-based) strategy (i.e., accepting every offer). These participants were removed from all analyses. Thus, behavioral analyses included 46 participants (M_{age} = 24.02 years, SD = 4.20; 22 male) and neuroimaging analyses included 23 participants (M_{age} = 24.35 years, SD = 3.80; 15 male). All participants gave written informed consent in accord with local ethics.

Experimental procedure

The experiment was conducted on two separate days. On the first day, each member of the couple separately provided a biometric photograph and rated the perceived level of closeness to his/her romantic partner by using the "Inclusion of Other in the Self" scale (Aron, Aron, & Smollan, 1992) (1 = *not at all close*, 7 = *very close*). Around two weeks later, couples returned to play a version of the UG. The experimenter explained the rules of the UG, and informed participants that the roles of proposer and responder within the couple would only be revealed when the experiment began. Participants were told they would be interacting with three different players via an online network: (a) their romantic partner, (b) an unknown person, and (c) a computer. Participants were further told that one round out of all trials would be randomly selected and paid out to them. After the experimenter made sure that all participants understood the instructions, participants were taken to two separate but adjacent rooms to perform the experiment; one participant went into the scanner and the other to a behavioral testing room. To enhance the interactive aspect of our UG task, it was emphasized that both the scanned and the non-scanned participant must start the game simultaneously. While in separate rooms, participants were told to prepare to play the game, and they then heard the experimenter give them both the countdown sequence "ready, set, go". On "go", participants were requested to commence the game by button press, after which they believed a computer algorithm determined their role in the UG. In reality, all participants received an on-screen message that they had been selected to play the game as the responder, and all offers were predetermined similar to other studies using the UG (Sanfey, Rilling, Aronson, Nystrom, & Cohen, 2003). Participants were then cued to press the button again, at which time the phrase "Connection is established ..." was presented, followed by the first experimental trial.

Participants played 180 UG trials in the role of responder. Each trial began with a fixation cross for an average of 5.5 s (variable range: 2–9 s in 1 s increments), followed by the proposer's photograph (4 s), the offer (4 s), and the outcome for that trial (4 s). While the offer was displayed, participants decided whether to accept or reject it by button press (left/right side counterbalanced across participants). Offers varied across three conditions: In the PARTNER condition (75 trials), participants received offers from their romantic partner. In the UNKNOWN condition (75 trials), participants received offers from an unknown person. The unknown person's gender was matched to that of the participants' romantic partner and this identity was held constant across participants. In the COMPUTER condition (30 trials), participants responded to offers from a computer proposer. Each condition contained equal proportions of fair (5€ [offered]: 5€ [kept], 4€:6€) and unfair (3€:7€, 2 €:8€, 1€:9€) offers. Offers were equally divided across three functional runs (60 trials/run, 18 min 18 s each) and each run had an equivalent number of trials from the three conditions (25 PARTNER trials, 25 UNKNOWN trials, 10 COMPUTER trials). Trials were presented in a pseudo-random order, with the restriction that no proposer and no offer would be presented on three consecutive trials. Note that we took several steps to control for potential confounds related to the photographs (see Supplementary Methods).

Post-scan measures

Upon completion of the UG task, both participants, while still in separate rooms, completed the following measures (in the order specified).

Fairness, attractiveness, and trustworthiness ratings

Participants rated the perceived fairness, attractiveness, and trustworthiness of both their romantic partner and the unknown person (1 = *not at all fair/attractive/trustworthy*, 7 = *extremely fair/attractive/trustworthy*). Consistent with research on positive illusions in romantic relationships (Murray & Holmes, 1997), we expected participants' ratings of their romantic partner to be positively biased (i.e., more favorable than those made by objective observers). Therefore, a third (control) condition required rating the perceived fairness, attractiveness, and trustworthiness of another gender-matched person (OTHER) who had participated in the previous session. We used these other-ratings as a reality benchmark to demonstrate that, on average, the romantic partners and the unknown persons were comparable regarding the above characteristics.

Affective responses to unfairness

Participants also rated their subjective affective state for unfair offers made by the romantic partner and the unknown person (in a counterbalanced order across participants). This was performed by asking participants about the following emotions: surprise, anger, happiness, disgust, revenge, pain, confusion, forgiveness, benevolence, and disappointment (randomized within participants; 1 = *not at all*, 7 = *very much*).

Implicit forgiveness

Participants performed a behavioral test of forgiveness using a variant of the IAT (Greenwald et al., 1998), the forgiveness IAT (Fatfouta et al., 2014). The forgiveness IAT is a computerized, speeded categorization task that measures mental associations of self with forgiveness. Participants were asked to categorize stimuli (words) into one of two categories, which varied across five blocks of trials. The first two blocks consisted of a simple discrimination task, in which participants practiced correctly categorizing stimuli from the target category (me–others) and attribute category (forgiving–vengeful). The third block combined both discrimination tasks (i.e., me–forgiving share one response button; others–vengeful share the other response button). In block 4, the labels of the attribute category were reversed (vengeful–forgiving). Block 5 consisted of the reversed combined discrimination task (i.e., me–vengeful share one response button; others–forgiving share the other response button). Automatic mental associations between "forgiving" and "me" were computed using the recommended scoring algorithm (so-called D_1; Greenwald, Nosek, & Banaji, 2003); higher scores reflect stronger associations between self and forgiveness (i.e., faster responses to the me–forgiving block relative to the me–vengeful block).

Explicit forgiveness

Participants' explicit forgiveness was measured using the Tendency to Forgive Scale (Brown, 2003), a four-item self-report questionnaire (α = 0.71). An example item: "I tend to get over it quickly when someone hurts my feelings" (1 = *strongly disagree*, 7 = *strongly agree*).

Resource sharing

Finally, we asked participants to indicate how much they share resources in their relationship on a single-item Likert scale (1 = *not at all*, 7 = *extremely*).

MRI data acquisition

Imaging was performed on a 3T scanner (Trio; Siemens, Erlangen, Germany). Anatomical images were acquired using a T1-weighted MPRage protocol (256 × 256 matrix, 176 sagittal slices of 1 mm thickness). Functional images were acquired using T2*-weighted echoplanar images (EPI) (TR = 2.0 s, TE = 30 ms, 64 × 64 matrix, flip angle = 70°, field of view = 192 mm, interslice gap = 0.6 mm). A total of 37 axial slices (3 × 3 × 3 mm voxels) parallel to the anterior and posterior commissure were collected per volume. A total of 545 volumes were collected per run. The total scanning time amounted to 54 min 54 s.

fMRI data analysis

Data were preprocessed and analyzed using FMRIB's Software Library, version 5.0.6 (FSL; Smith et al., 2004). Preprocessing included brain extraction (Smith, 2002), motion correction (Jenkinson, Bannister, Brady, & Smith, 2002), spatial smoothing with a Gaussian kernel of 6-mm full-width at half-maximum, and high-pass temporal filtering with a cutoff of 80 s. EPI images were registered to each participant's structural image using boundary-based registration (Greve & Fischl, 2009), and then normalized into Montreal Neurological Institute (MNI) space using linear registration with 12 degrees of freedom (Jenkinson et al., 2002; Jenkinson & Smith, 2001).

We estimated a general linear model (GLM) of the blood oxygen level dependent (BOLD) signal using the following 13 regressors (R):

- R1–R3 were indicator functions for the period of the trial when participants discovered the proposer (i.e., PARTNER, UNKNOWN, or COMPUTER);
- R4–R6 were indicator functions denoting a fair offer in each condition;
- R7–R9 were indicator functions denoting an unfair offer in each condition;
- R10–12 were indicator functions denoting the outcome for each condition;
- R13 was an indicator function denoting missed trials where no decision was made within the 4 s time window.

All regressors of interest were convolved with a double-gamma hemodynamic response function (HRF). Individual contrast images were calculated and then submitted to a higher-level mixed-effects analysis using voxel-wise one-sample *t*-tests. To identify brain regions showing greater activation for unfair offers from humans (vs. fair offers from humans), we computed the contrast [unfair PARTNER + unfair UNKNOWN > fair PARTNER + fair UNKNOWN]. To identify brain regions showing greater activation for unfair offers from one's partner (vs. unfair from the unknown person), we computed the contrast [unfair PARTNER > unfair UNKNOWN]. For all whole-brain analyses, *Z*-statistic images were thresholded at $z = 2.3$ and cluster corrected to $P < 0.05$.

Psychophysiological interaction (PPI) analysis

We performed a PPI analysis to identify brain regions showing context-specific changes in the relationship with the MPFC as a function of receiving unfair offers from one's partner versus unfair offers from the unknown person. The model was estimated in three steps (O'Reilly, Woolrich, Behrens, Smith, & Johansen-Berg, 2012). First, we identified the group peak response of the [unfair PARTNER > unfair UNKNOWN] contrast [MNI peak coordinates (*x, y, z*): −2, 66, 12] and created a sphere (5 mm radius) at this location. Second, we extracted individual average time-series of BOLD signal within the seed region. Third, for each participant, we estimated a GLM of the BOLD responses with the following three regressors:

- R1 was a psychological regressor denoting the main effect of task, convolved with a double-gamma HRF;
- R2 was a physiological regressor denoting the activation time course of the MPFC seed region;
- R3 was a PPI regressor denoting the element-by-element product of the previous two (i.e., the PPI term).

Individual contrast images were calculated and then submitted to a higher-level mixed-effects analysis using voxel-wise one-sample *t*-tests. For all whole-brain analyses, *Z*-statistic images were thresholded at $z = 2.3$ and cluster corrected to $P < 0.05$.

Results

Behavior

Acceptance rates (% of accepted offers) were analyzed using a 2 × 2 repeated-measures analysis of variance (using Bonferroni post-hoc tests) with player (PARTNER vs. UNKNOWN) and fairness (fair vs. unfair) as within-subject factors. The analysis revealed a main effect of fairness [$F(1, 45) = 183.30$, $P < 0.001$; $\eta_p^2 = 0.80$], indicating that participants were significantly less likely to accept unfair than fair offers [$M = 43.1\% \pm SEM = .037\%$ vs. $M = 92.0\% \pm .020\%$ *SEM* ($P < .001$), respectively; Figure 1(b)]. Importantly, there was also a main effect of relationship closeness [$F(1, 45) = 46.40$, $P < 0.001$; $\eta_p^2 = 0.51$]; participants were significantly more likely to accept offers from their partner than offers from the unknown person ($M = 77.3\% \pm .028\%$ *SEM* vs. $M = 57.8\% \pm .027\%$ *SEM* ($P < .001$), respectively). Furthermore, relationship closeness interacted significantly with fairness [$F(1, 45) = 44.77$, $P < 0.001$; $\eta_p^2 = 0.50$]; the difference in acceptance rates between partner and unknown was significant for unfair [$M = 60.1\% \pm .048\%$ *SEM* vs. $M = 26.1\% \pm .038\%$ *SEM* ($P < .001$)] and fair [$M = 94.5\% \pm .016\%$ *SEM* vs. $M = 89.5\% \pm .028\%$ *SEM* ($P = .041$)] offers. We next examined the extent to which relationship closeness also modulates individuals' internal feeling states in response to unfairness. Analysis of participants' emotion ratings revealed greater forgiveness and benevolence in response to unfair offers from the partner than unfair offers from the unknown person (see Supplementary Table 1).

Control analysis

A potential concern is that individual differences in resource sharing could confound participants' acceptance rates in the UG. Greater resource sharing among couples may be reflected in greater acceptance of monetary offers made by the romantic partner. To rule out this interpretation, we repeated the behavioral analysis including resource sharing as a covariate. We replicated our main results, finding significant main effects of fairness ($P < 0.001$) and closeness ($P < 0.001$) as well as a significant fairness × closeness interaction ($P < 0.001$). All possible interactions with resource sharing (closeness × resource sharing, fairness × resource sharing, and closeness × fairness × resource sharing) were not significant (*P*'s were 0.574, 0.712, and 0.574, respectively). This control analysis indicates that the effects we observe in the UG task relate to differences in interpersonal closeness rather than resource sharing among couples.

Neuroimaging

We analyzed the neuroimaging data to reveal brain regions showing greater activation for unfair than fair offers from human proposers. Our results included the bilateral anterior insula (AI) extending into inferior frontal gyrus (IFG), dorsal anterior cingulate cortex (DACC), and right dorsolateral prefrontal cortex (DLPFC) (Figure 2(a,b); Table 1).

We then compared the neural correlates of unfairness separately for each proposer, and found that unfair offers from the partner and the unknown person activated an overlapping subset of brain areas, including right insular cortex/IFG, middle occipital gyrus, and DACC (Figure 3; Tables 2 and 3).

We next performed a contrast to reveal brain regions showing greater activation for unfair offers from one's partner than from an unknown person. This analysis revealed only one significant cluster in the MPFC (Figure 4(a,b)) [Montreal Neurological Institute (MNI) coordinates (*x, y, z*):–2, 66, 12]. The anterior MPFC has been implicated in meta-cognitive processes of reasoning about intentions (Amodio & Frith, 2006) and, hence, activation here might reflect increased efforts to understand the partner's motives for his/her unfair behavior (Campanhã et al., 2011; Meyer et al., 2013). This consideration motivated us to examine whether the region of MPFC identified would overlap with regions that have been preferentially associated with tasks relevant to mentalizing in the neuroimaging literature. As expected, the observed task-related change in MPFC activation overlapped to a large degree with the mentalizing network (Figure 4(c)), as determined by a formal

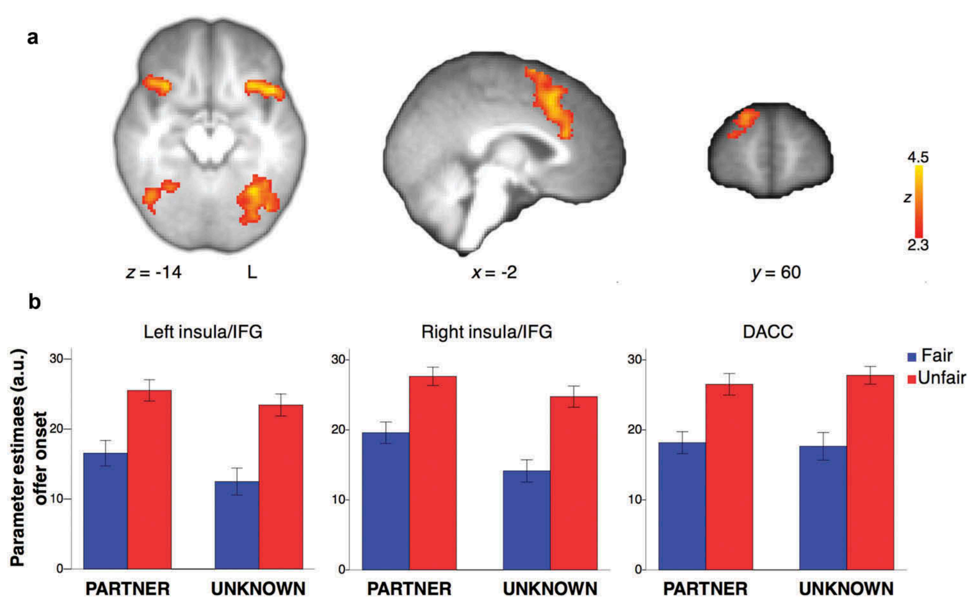

Figure 2. BOLD signal change related to unfairness. (a) Significant changes in BOLD signal for the contrast [unfair PARTNER + unfair UNKNOWN > fair PARTNER + fair UNKNOWN] were found in bilateral AI/IFG, DACC, right cerebellum/left middle occipital gyrus (extending into bilateral occipital lobe), and right DLPFC. For complete list of results, see Table 1. Results were thresholded at $z > 2.3$, $P < 0.05$, cluster corrected. L = left hemisphere. (b) Mean parameter estimates in arbitrary units (a.u.) within bilateral AI and DACC for fair and unfair offers from the partner and the unknown person. Error bars indicate one standard error of the mean, calculated within-subjects.

Table 1. Brain regions showing significantly greater activation for unfair than fair offers from a human partner (unfair PARTNER + unfair UNKNOWN > fair PARTNER + fair UNKNOWN).

Brain region	MNI coordinates *x*	*y*	*z*	Cluster size (voxels)	Peak *z*-value
L middle occipital gyrus (BA 19)	−38	−84	12	5203	5.39
R dorsal anterior cingulate cortex (DACC; BA 24/32)	14	12	60	2755	4.17
R insula/inferior frontal gyrus (BA 13/47)	34	20	−10	1542	4.54
R superior parietal lobule (BA 7)	22	−66	50	1344	4.55
R dorsolateral prefrontal cortex (DLPFC; BA 9/46)	34	50	16	1145	4.16
R cerebellum (BA 37)	48	−52	−34	1053	4.11
L insula/inferior frontal gyrus (BA 13/47)	−44	18	−12	889	4.68

MNI = Montreal Neurological Institute. $z > 2.3$, $P < 0.05$, cluster-corrected; L, left; R, right. BA = Brodmann's area. MNI coordinates refer to the largest peak voxel per cluster.

reverse-inference analysis using Neurosynth (Yarkoni, Poldrack, Nichols, Van Essen, & Wager, 2011; see Supplementary Results).

We had hypothesized that refraining from punishment of unfairness was related to the interaction between brain regions involved in mentalizing and regions that process cognitive conflict. Therefore, we predicted different functional connectivity between these regions in the context of unfair behavior from one's partner relative to unfair behavior from the unknown person. To this end, we performed a psychophysiological interaction (PPI; O'Reilly et al., 2012) analysis to identify brain regions exhibiting a change in functional connectivity with the MPFC depending on whether the partner or the unknown person made an unfair offer. This analysis revealed heightened functional connectivity between the MPFC and anterior portions of the DACC (Figure 5(a); Table 4).

We next tested to which degree MPFC-DACC functional connectivity was related to participants' acceptance of unfair partner offers (vs. unfair unknown offers). Therefore, we calculated the absolute difference between accepted unfair offers from the partner and the unknown person [Δ acceptance = (|% of accepted unfair $\text{offers}_{\text{PARTNER}}$ − % of accepted unfair $\text{offers}_{\text{UNKNOWN}}$|)]. We refer to this measure as the acceptance index. An acceptance index greater than zero indicates that an individual accepted unfair offers from his/her partner more frequently than from the unknown person and, hence, showed a higher tendency to "forgive" his/her partner (vs. the unknown person). We found that MPFC-DACC functional connectivity was inversely associated with participants' acceptance index (Figure 5(b)); individuals who showed lower levels of functional connectivity between the MPFC and DACC were better able to tolerate their partner's unfairness [$\beta = -0.58$, $t = -3.29$, $P = 0.003$ (two-tailed); 95% confidence interval (CI) for $\beta = -0.80/-0.22$].

Finally, we determined whether the task-related change in MPFC-DACC functional connectivity predicted individuals' forgiveness level in two independent

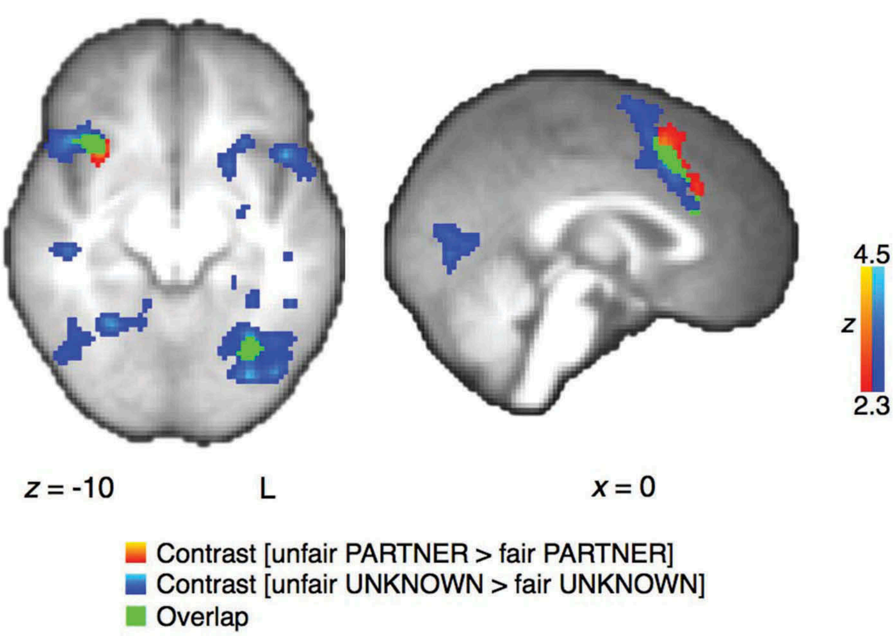

Figure 3. Overlap of BOLD signal change related to unfairness. Comparison of *z* statistic images of unfair offers from each proposer [unfair PARTNER > fair PARTNER in red-yellow; unfair UNKNOWN > fair UNKNOWN in blue-light blue]. Overlapping activations are displayed in green. The overlap included areas in right insular cortex/IFG, left middle occipital gyrus, and DACC. Results were thresholded at $z > 2.3$, $P < 0.05$, cluster corrected.

Table 2. Brain regions showing significantly greater activation for unfair than fair offers from a romantic partner (unfair PARTNER > fair PARTNER).

	MNI coordinates			Cluster size	
Brain region	*x*	*y*	*z*	(voxels)	Peak *z*-value
R dorsal anterior cingulate cortex (DACC; BA 24/32)	8	32	30	903	3.96
L middle occipital gyrus (BA 19)	−44	−78	8	872	3.64
R insula (BA 13)	32	18	−10	498	3.76

MNI = Montreal Neurological Institute. $z > 2.3$, $P < 0.05$, cluster-corrected; L, left; R, right. BA = Brodmann's area.

Table 3. Brain regions showing significantly greater activation for unfair than fair offers from an unknown person (unfair UNKNOWN > fair UNKNOWN).

	MNI coordinates			Cluster size	
Brain region	*x*	*y*	*z*	(voxels)	Peak *z*-value
L middle occipital gyrus (BA 19)	−36	−84	10	4305	5.11
R Cerebellum (BA 37)	50	−52	−34	2958	4.19
R dorsal anterior cingulate cortex (DACC; BA 24/32)	−6	10	46	2023	4.02
R insula/inferior frontal operculum (BA 13/38)	42	24	−14	1674	3.8
L temporal pole (BA 48)	−46	16	−12	1261	4.31
R middle occipital gyrus (BA 19)	40	−80	24	1182	3.92
L middle frontal gyrus (MFG; BA 46)	−38	44	32	688	3.98
L middle temporal gyrus (MTG; BA 21)	−54	−50	20	517	4.0

MNI = Montreal Neurological Institute. $z > 2.3$, $P < 0.05$, cluster-corrected; L, left; R, right. BA = Brodmann's area.

measures assessed outside the MRI scanner: one was a self-report questionnaire [Tendency to Forgive scale (TTF; Brown, 2003)] and one was a behavioral test [forgiveness IAT (Fatfouta et al., 2014)]. As Figure 5(c) illustrates, the task-related change in MPFC-DACC functional connectivity inversely predicted individuals' IAT scores [left panel; $\beta = -0.59$, $t = -3.34$, $P = 0.003$ (two-tailed); 95% CI for $\beta = -0.81/-0.23$], but not their TTF scores [right panel; $\beta = -0.01$, $t = -0.05$, $P = 0.961$ (two-tailed); 95% CI for $\beta = -0.42/0.40$]. In other words, those individuals who showed lower levels of functional connectivity between the MPFC and DACC showed stronger automatic associations between the self and forgiveness, which were not readily accessible in the self-report. This prediction was insensitive to any selection biases, as determined by a leave-one-subject out cross-validation (Esterman, Tamber-Rosenau, Chiu, & Yantis, 2010; see Supplementary Results).

Discussion

This study sought to determine how relationship closeness modulates behavioral and neural responses to unfairness, with the source of unfairness being one's own partner, compared to an unknown person. Our study yielded three main results: First, participants were more likely to accept unfair offers from their partner compared with the unknown person, and this effect remained significant after adjustment for resource sharing among couples. Second, when participants received unfair partner offers (vs. unfair unknown offers), a significant change in brain activation was observed in the MPFC; this region also demonstrated heightened functional connectivity with the DACC when the partner rather than the unknown person made an unfair offer. Third, lower levels of the task-related change in MPFC-DACC functional connectivity predicted participants' tendency to accept offers from one's partner following unfairness as well as behavioral but not self-reported expressions of forgiveness.

Our findings support the valuable-relationships theory (McCullough, 2008), with relationship closeness influencing economically relevant social decision-making: when a partner rather than an unknown person made an unfair offer, individuals were more likely to accept this offer. The observed effects of relationship closeness on participants' acceptance rates are unlikely to simply reflect differences in how resources were actually being shared with one's partner, as resource sharing did not affect participants' responses. Rather, this behavioral tendency to accept unfair partner offers may reflect a benevolent intention, as mirrored in participants' emotion ratings. In this respect, our findings fit with the notion that enforcing fairness norms through retaliation would be mitigated when interacting with a close ally, but prevail when interacting with a stranger (Campanhã et al., 2011). Consistent with this, related research has demonstrated that individuals accept more unfair offers when playing with an in-group member (and by extension, close ally) than with an out-group member (Diekhof, Wittmer, & Reimers, 2014).

When facing unfairness, participants activated a network of areas comprising bilateral AI/IFG, DACC, and DLPFC. These areas are consistent with those found in previous fMRI studies of the UG, and related tasks involving social decision-making (Gabay, Radua, Kempton, & Mehta, 2014; Sanfey et al., 2003). Extending these studies with regard to the relationship one has with the proposer, we found activation overlap in the right insular cortex/IFG and DACC for unfair offers from both proposers, that is, one's partner and the unknown person. This finding lends additional support for the notion that the insula/IFG and DACC play a critical role in detecting social norm violations (fairness norms, in particular) induced by inequity (Güroğlu, van den Bos, van

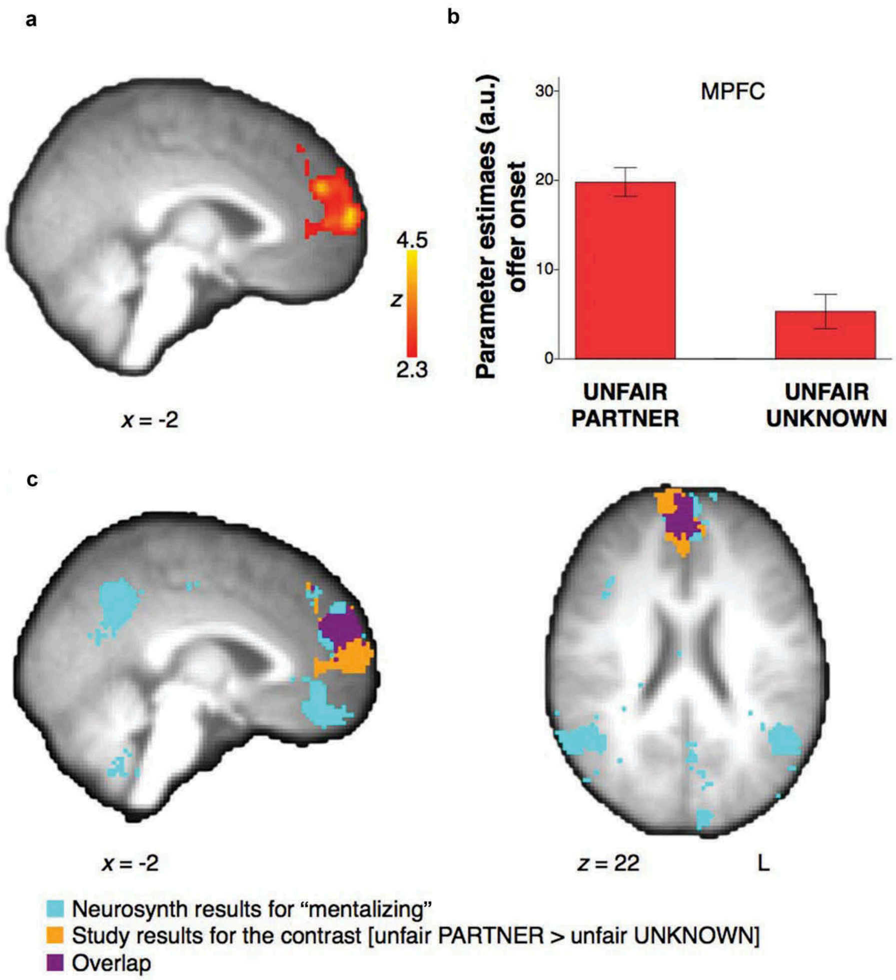

Figure 4. Brain activation associated with unfairness by the partner. (a) Significant changes in BOLD signal for the contrast [unfair PARTNER > unfair UNKNOWN] were found in MPFC. Results were thresholded at $z > 2.3$, $P < 0.05$, cluster corrected. (b) Mean parameter estimates in arbitrary units (a.u.) within MPFC for unfair offers from the partner and the unknown person. Error bars indicate one standard error of the mean, calculated within-subjects. (c) Comparison of Neurosynth meta-analysis search for the term "mentalizing" (light blue) to study results for the contrast [unfair PARTNER > unfair UNKNOWN] (orange, same as in A). Overlapping activations are displayed in purple (660 voxels, 1.98 cm^3). Neurosynth results were corrected for multiple comparisons using a whole-brain false discovery rate (FDR) threshold of $P < 0.05$.

Dijk, Rombouts, & Crone, 2011). The result that the insula activation also extended into IFG coheres with previous UG fMRI studies (Kirk, Downar, & Montague, 2011) and has been interpreted to reflect emotion regulation following unfairness (Güroğlu et al., 2011). In line with this view, such regulation might be related to the reappraisal of the proposers' intentions and/or behavior in a more or less negative way (Grecucci, Giorgetta, Bonini, & Sanfey, 2013). Increased AI/IFG activation thus may reflect enhanced efforts to cope with the unfairness at hand.

The fact that the anterior MPFC was the only region activated when participants received unfair offers from their partner, as compared to unfair offers from the unknown person, fits with the proposed role of the MPFC in mentalizing-related computations during social exchange (Hampton, Bossaerts, & O'Doherty, 2008), and also with the recent literature implicating MPFC in mentalizing about close others (Krienen et al., 2010). Activity in this region is also consistent with prior meta-analyses, which emphasize the importance of the MPFC (primarily Brodmann Area [BA] 10) for reasoning about belief (Denny, Kober, Wager, & Ochsner, 2012; Gilbert et al., 2006). Being treated unfairly by one's partner likely instigates a thought process, by which one might try to make sense of the reasons and motivations underlying his/her behavior as well as to anticipate his/her future behavior. The overlap between our fMRI results and the reverse-inference meta-analysis results confirmed this rationale. Concordance between the results of this study and fMRI studies of mentalizing indicates that processing unfair offers from one's

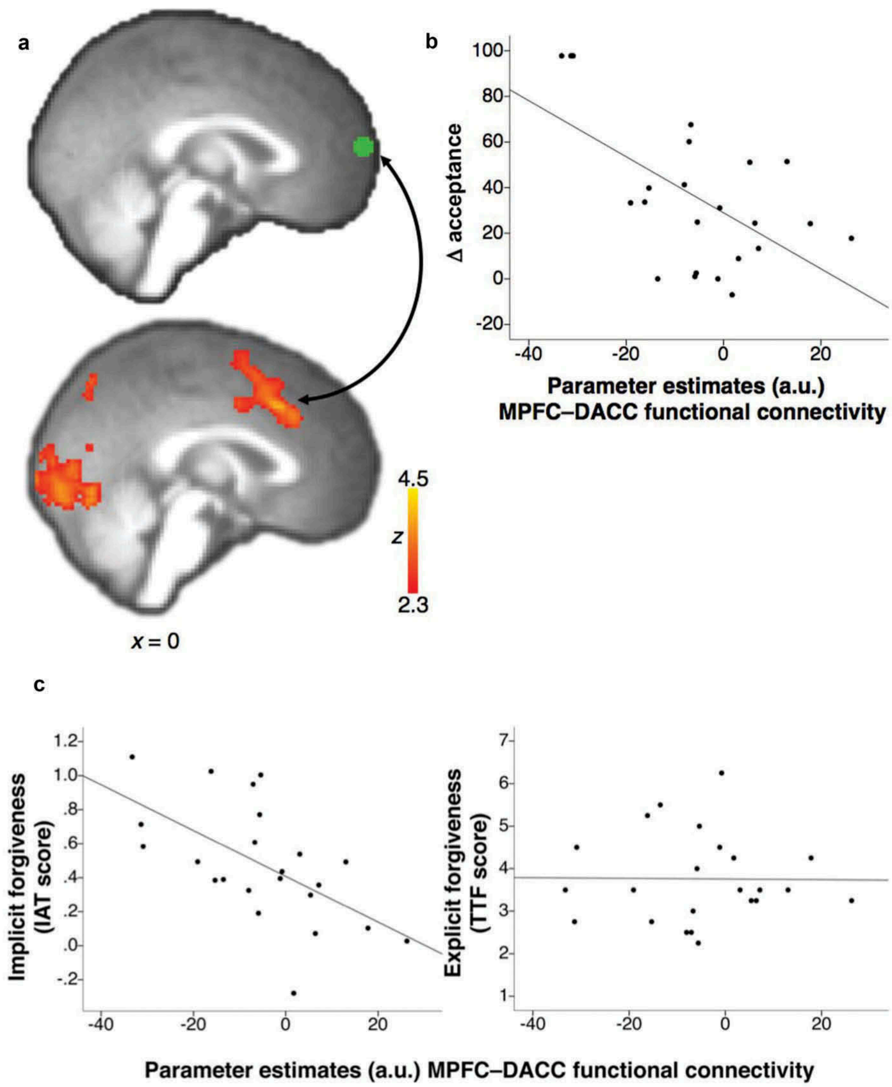

Figure 5. Reduced connectivity between MPFC and DACC predicts acceptance of unfairness. (a) PPI analysis identifying brain regions showing connectivity changes with the MPFC as a function of receiving unfair offers from the partner versus the unknown person. The upper panel depicts our functionally defined seed region (i.e., MPFC). The lower panel depicts brain regions whose connectivity with MPFC increases for unfair offers from the partner (vs. the unknown person). Results were thresholded at $z > 2.3$, $P < 0.05$, cluster corrected. (b) Regression line depicting the relationship between in MPFC-DACC functional connectivity and acceptance index (Δ) across participants. The x-axis represents the mean parameter estimates in arbitrary units (a.u.) for task-related change in MPFC-DACC functional connectivity. The y-axis represents the acceptance index, defined as the absolute difference between accepted unfair offers from the partner versus the unknown person [Δ acceptance = (|% of accepted unfair offers$_{\mathrm{PARTNER}}$—% of accepted unfair offers$_{\mathrm{UNKNOWN}}$|)]. (c) Regression lines depicting the relationship between MPFC-DACC functional connectivity and tolerance of unfairness across participants. The x-axis in both plots represents the mean parameter estimates in arbitrary units (a.u.) for task-related change in MPFC-DACC functional connectivity. The y-axis represents individual differences in forgiveness as measured by implicit (left panel; IAT score) and explicit (right panel; TTF score) measures.

partner versus an unknown person activates the same region as mental-state reasoning (i.e., MPFC).

Our results point to a functional relation between MPFC and DACC in situations in which participants decide whether to accept unfair partner behavior. Supporting our hypothesis, we found that the changes in brain activity in MPFC covaried significantly more with the DACC when one's partner behaved unfairly than when an unknown person behaved unfairly. Previous studies implicated the DACC in the evaluation of conflict (Botvinick, Nystrom, Fissell, Carter, & Cohen, 1999) and, more specifically, the conflict between cognitive ("accept") and emotional ("reject") motivations in the UG (Gabay et al., 2014; Sanfey et al., 2003). In the present study, unfair partner offers may present individuals with such an adverse situation: acting upon fairness norms might promote better future offers on the one hand, but also signal retaliatory cost impositions on

Table 4. Brain regions showing increased coupling with the MPFC as a function of receiving unfair offers from one's partner relative to the unknown person.

Brain region	MNI coordinates			Cluster size	
	x	*y*	*z*	(voxels)	Peak *z*-value
L lingual gyrus (BA 17/18)	−28	−84	−14	27026	5.59
L dorsal anterior cingulate cortex (DACC; BA 24/32)	0	22	38	1192	3.75
R precentral gyrus (BA 6)	44	6	42	1142	4.38

MNI = Montreal Neurological Institute. $z > 2.3$, $P < 0.05$, cluster-corrected; L, left; R, right. BA = Brodmann's area.

the other hand. In line with the regulatory role of prefrontal cortex in modulating ACC activation when competing response tendencies are at hand (Botvinick, Braver, Barch, Carter, & Cohen, 2001) and with the MPFC's connectivity with the DACC during decisional conflicts (Nakao et al., 2010), such conflict encoded in DACC may then be resolved through mentalizing-related MPFC activation.

Participants with reduced MPFC-DACC connectivity accepted unfair offers from one's partner more frequently than those from an unknown person. As discussed earlier, DACC activation appears to represent conflict (Botvinick et al., 1999) and MPFC likely represents mentalizing (Figure 5(c); Hampton et al., 2008; Krienen et al., 2010); therefore, we suggest that the less information about conflict that is online and readily available to the MPFC during a transgression, the more likely an individual is to adopt an accepting stance toward that transgression. Importantly, this pattern was also obtained with an independent behavioral measure outside the MRI scanner; reduced MPFC-DACC connectivity predicted individuals' automatic (implicit) mental associations regarding forgiveness. In line with related research on the neural basis of implicit social cognition (Stanley et al., 2008), only the implicit (IAT-measured) but not the explicit (self-reported) expression of forgiveness was predicted by brain activation. These results might be tentatively interpreted in terms of an enhanced tendency toward prosocial behavior in the context of close relationships rather than an enhancement of rational self-interested behavior. To achieve an optimal exchange in the context of close relationships, one may be better advised to forfeit even small gains as a way of preventing potential loss in the long run (e.g., fitness or other benefits such as social support).

It should be noted that there are possible limitations to our study. First, while previous studies similarly operationalized prosocial (i.e., "forgiving") reactions as refraining from punishment during social interactions (Will et al., 2015), caution is warranted in interpreting such a high-level cognitive concept. Our strategy for dealing with this issue was to employ multiple measures (i.e., emotion ratings and implicit and explicit forgiveness) rather than focusing on brain activation alone. Future work is thus needed to further validate our acceptance index and to refine the way it represents an attempt to promote conciliation after being transgressed against. Second, though we tried to control for potential confounds (e.g., resource sharing among couples), several other dynamics may act to modulate the decision to accept an unfair split. Future work would benefit from assessing additional variables such as responders' perceived intentionality of the unfair offer. Finally, our study used predetermined (i.e., constrained) offers to allow for experimental control. While this represents a standard practice in the UG literature (e.g., Sanfey et al., 2003), it might be equally informative to further investigate how romantic couples respond to unconstrained splits.

Overall, we found that feeling socially close to someone enhanced the tendency toward accepting unfairness, and this tendency was associated with a network comprised of the MPFC and DACC. Differences in connectivity between these regions were inversely related to a behavioral test of forgiveness assessed outside the scanner. In the future, the potential clinical utility of relationship closeness could be investigated in patients with aberrant social-exchange skills—in particular, psychopathological disorders following trauma that are accompanied by embitterment and revenge (Gäbler & Maercker, 2011). More generally, these findings have important implications for human decision-making because they demonstrate how relationship closeness modulates prefrontal brain function during strategic interactions, thereby advancing our knowledge of the context-dependent nature of social behavior.

Disclosure statement

No potential conflict of interest was reported by the authors.

ORCID

Ramzi Fatfouta http://orcid.org/0000-0001-6633-356X

References

Amodio, D. M., & Frith, C. D. (2006). Meeting of minds: The medial frontal cortex and social cognition. *Nature Reviews Neuroscience*, *7*(4), 268–277. doi:10.1038/nrn1884

Aron, A., Aron, E. N., & Smollan, D. (1992). Inclusion of other in the Self Scale and the structure of interpersonal closeness. *Journal of Personality and Social Psychology*, *63*(4), 596. doi:10.1037/0022-3514.63.4.596

Botvinick, M. M., Braver, T. S., Barch, D. M., Carter, C. S., & Cohen, J. D. (2001). Conflict monitoring and cognitive control. *Psychological Review*, *108*(3), 624–652. doi:10.1037/0033-295X.108.3.624

Botvinick, M. M., Nystrom, L. E., Fissell, K., Carter, C. S., & Cohen, J. D. (1999). Conflict monitoring versus selection-for-action in anterior cingulate cortex. *Nature*, *402*(6758), 179–181. doi:10.1038/46035

Brown, R. P. (2003). Measuring individual differences in the tendency to forgive: Construct validity and links with depression. *Personality & Social Psychology Bulletin*, *29*(6), 759–771. doi:10.1177/0146167203029006008

Camerer, C. F. (2003). *Behavioral game theory: Experiments in strategic interaction*. New York, NY: Russell Sage Foundation.

Campanhã, 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*(43), 15569–15574. doi:10.1523/JNEUROSCI.1253-11.2011

Crockett, M. J., Apergis-Schoute, A., Herrmann, B., Lieberman, M., Müller, U., Robbins, T. W., & Clark, L. (2013). Serotonin modulates striatal responses to fairness and retaliation in humans. *The Journal of Neuroscience*, *33*(8), 3505–3513. doi:10.1523/JNEUROSCI.2761-12.2013

De Waal, F. B., & Pokorny, J. J. (2005). Primate conflict resolution and its relation to human forgiveness. In E. Worthington (Ed.), *Handbook of forgiveness* (pp. 17–32). New York, NY: Routledge.

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*(8), 1742–1752. doi:10.1162/jocn_a_00233

Diekhof, E. K., Wittmer, S., & Reimers, L. (2014). Does competition really bring out the worst? Testosterone, social distance and inter-male competition shape parochial altruism in human males. *Plos One*, *9*(7), e98977–e98977. doi:10.1371/journal.pone.0098977

Esterman, M., Tamber-Rosenau, B. J., Chiu, Y.-C., & Yantis, S. (2010). Avoiding non-independence in fMRI data analysis: Leave one subject out. *Neuroimage*, *50*(2), 572–576. doi:10.1016/j.neuroimage.2009.10.092

Farrow, T. F., Hunter, M. D., Wilkinson, I. D., Gouneea, C., Fawbert, D., Smith, R., … Woodruff, P. W. R. (2005). Quantifiable change in functional brain response to empathic and forgivability judgments with resolution of posttraumatic stress disorder. *Psychiatry Research: Neuroimaging*, *140*(1), 45–53. doi:10.1016/j.pscychresns.2005.05.012

Farrow, T. F., Zheng, Y., Wilkinson, I. D., Spence, S. A., Deakin, J. F. W., Tarrier, N., … Woodruff, P. W. R. (2001). Investigating the functional anatomy of empathy and forgiveness. *NeuroReport: For Rapid Communication of Neuroscience Research*, *12*(11), 2433–2438. doi:10.1097/00001756-200108080-00029

Fatfouta, R., Schröder-Abé, M., & Merkl, A. (2014). Forgiving, fast and slow: Validity of the implicit association test for predicting differential response latencies in a transgression-recall paradigm. *Frontiers in Psychology*, *5*, 728. doi:10.3389/fpsyg.2014.00728

Fatfouta, R., Schulreich, S., Meshi, D., & Heekeren, H. (2015). So close to a deal: Spatial-distance cues influence economic decision-making in a social context. *PLoS One*, *10*(8), e0135968. doi:10.1371/journal.pone.0135968

Gabay, A. S., Radua, J., Kempton, M. J., & Mehta, M. A. (2014). The ultimatum game and the brain: A meta-analysis of neuroimaging studies. *Neuroscience & Biobehavioral Reviews*, *47*, 549–558. doi:10.1016/j.neubiorev.2014.10.014

Gäbler, I., & Maercker, A. (2011). Revenge after trauma: Theoretical outline. In M. Linden & A. Maercker (Eds.), *Embitterment. Societal, psychological, and clinical perspectives* (pp. 42–69). Wien: Springer.

Gilbert, S. J., Spengler, S., Simons, J. S., Steele, J. D., Lawrie, S. M., Frith, C. D., & Burgess, P. W. (2006). Functional specialization within rostral prefrontal cortex (Area 10): A Meta-analysis. *Journal of Cognitive Neuroscience*, *18*(6), 932–948. doi:10.1162/jocn.2006.18.6.932

Grecucci, A., Giorgetta, C., Bonini, N., & Sanfey, A. G. (2013). Reappraising social emotions: The role of inferior frontal gyrus, temporo-parietal junction and insula in interpersonal emotion regulation. *Frontiers in Human Neuroscience*, *7*. doi:10.3389/fnhum.2013.00523

Greenwald, A. G., McGhee, D. E., & Schwartz, J. L. (1998). Measuring individual differences in implicit cognition: The implicit association test. *Journal of Personality and Social Psychology*, *74*(6), 1464–1480. doi:10.1037/0022-3514.74.6.1464

Greenwald, A. G., Nosek, B. A., & Banaji, M. R. (2003). Understanding and using the implicit association test: I. An improved scoring algorithm. *Journal of Personality and Social Psychology*, *85*(2), 197–216. doi:10.1037/0022-3514.85.2.197

Greve, D. N., & Fischl, B. (2009). Accurate and robust brain image alignment using boundary-based registration. *Neuroimage*, *48*(1), 63–72. doi:10.1016/j.neuroimage.2009.06.060

Güroğlu, B., van den Bos, W., van Dijk, E., Rombouts, S. A. R. B., & Crone, E. A. (2011). Dissociable brain networks involved in development of fairness considerations: Understanding intentionality behind unfairness. *NeuroImage*, *57*(2), 634–641. doi:10.1016/j.neuroimage.2011.04.032

Güth, W., Schmittberger, R., & Schwarze, B. (1982). An experimental analysis of ultimatum bargaining. *Journal of Economic Behavior & Organization*, *3*(4), 367–388. doi:10.1016/0167-2681(82)90011-7

Hampton, A. N., Bossaerts, P., & O'Doherty, J. P. (2008). Neural correlates of mentalizing-related computations during strategic interactions in humans. *PNAS Proceedings of the National Academy of Sciences of the United States of America*, *105*(18), 6741–6746. doi:10.1073/pnas.0711099105

Hayashi, A., Abe, N., Ueno, A., Shigemune, Y., Mori, E., Tashiro, M., & Fujii, T. (2010). Neural correlates of forgiveness for moral transgressions involving deception. *Brain Research*, *1332*, 90–99. doi:10.1016/j.brainres.2010.03.045

Jenkinson, M., Bannister, P., Brady, M., & Smith, S. (2002). Improved optimization for the robust and accurate linear registration and motion correction of brain images. *Neuroimage*, *17*(2), 825–841. doi:10.1006/nimg.2002.1132

Jenkinson, M., & Smith, S. (2001). A global optimisation method for robust affine registration of brain images. *Medical Image Analysis*, *5*(2), 143–156. doi:10.1016/S1361-8415(01)00036-6

Karremans, J. C., & Aarts, H. (2007). The role of automaticity in determining the inclination to forgive close others. *Journal of Experimental Social Psychology*, *43*(6), 902–917. doi:10.1016/j.jesp.2006.10.012

Karremans, J. C., Regalia, C., Paleari, F. G., Fincham, F. D., Cui, M., Takada, N., … Uskul, A. K. (2011). Maintaining harmony across the globe: The cross-cultural association between closeness and interpersonal forgiveness. *Social Psychological and Personality Science*, *2*(5), 443–451. doi:10.1177/1948550610396957

Kirk, U., Downar, J., & Montague, P. R. (2011). Interoception drives increased rational decision-making in meditators playing the ultimatum game. *Frontiers in Neuroscience*, *5*. doi:10.3389/fnins.2011.00049

Koenigs, M., & Tranel, D. (2007). Irrational economic decision-making after ventromedial prefrontal damage: Evidence from the Ultimatum Game. *The Journal of Neuroscience*, *27*(4), 951–956. doi:10.1523/JNEUROSCI.4606-06.2007

Koski, S. E., Koops, K., & Sterck, E. H. M. (2007). Reconciliation, relationship quality, and postconflict anxiety: Testing the integrated hypothesis in captive chimpanzees. *American Journal of Primatology*, *69*(2), 158–172. doi:10.1002/ajp.20338

Krienen, F. M., Tu, P.-C., & Buckner, R. L. (2010). Clan mentality: Evidence that the medial prefrontal cortex responds to close others. *The Journal of Neuroscience*, *30*(41), 13906–13915. doi:10.1523/JNEUROSCI.2180-10.2010

Li, J.-C. A., & Chen, Y.-C. (2012). Personality, Affects, and Forgiving Behavior in Games. *SSRN Electronic Journal*. doi:10.2139/ssrn.2120811

McCullough, M. E. (2008). *Beyond revenge: The evolution of the forgiveness instinct*. San Francisco, CA: Jossey-Bass.

McCullough, M. E., Kurzban, R., & Tabak, B. A. (2013). Cognitive systems for revenge and forgiveness. *The Behavioral and Brain Sciences*, *36*(1), 1–15. doi:10.1017/s0140525x11002160

McCullough, M. E., Worthington, E., & Rachal, K. C. (1997). Interpersonal forgiving in close relationships. *Journal of Personality and Social Psychology*, *73*(2), 321–336. doi:10.1037/0022-3514.73.2.321

Mehta, P. H., & Beer, J. (2010). Neural Mechanisms of the Testosterone-Aggression Relation: The Role of Orbitofrontal Cortex. *Journal of Cognitive Neuroscience*, *22*(10), 2357–2368. doi:10.1162/jocn.2009.21389

Meyer, M. L., Masten, C. L., Ma, Y., Wang, C., Shi, Z., Eisenberger, N. I., & Han, S. H. (2013). Empathy for the social suffering of friends and strangers recruits distinct patterns of brain activation. *Social Cognitive and Affective Neuroscience*, *8*(4), 446–454. doi:10.1093/scan/nss019

Murray, S. L., & Holmes, J. G. (1997). A leap of faith? Positive illusions in romantic relationships. *Personality and Social Psychology Bulletin*, *23*(6), 586–604. doi:10.1177/0146167297236003

Nakao, T., Osumi, T., Ohira, H., Kasuya, Y., Shinoda, J., Yamada, J., & Northoff, G. (2010). Medial prefrontal cortex-dorsal anterior cingulate cortex connectivity during behavior selection without an objective correct answer. *Neuroscience Letters*, *482*(3), 220–224. doi:10.1016/j.neulet.2010.07.041

O'Reilly, J. X., Woolrich, M. W., Behrens, T. E., Smith, S. M., & Johansen-Berg, H. (2012). Tools of the trade: Psychophysiological interactions and functional connectivity. *Social Cognitive and Affective Neuroscience*, *7*(5), 604–609. doi:10.1093/scan/nss055

Ricciardi, E., Rota, G., Sanil, L., Gentili, C., Gaglianese, A., Guazzelli, M., & Pietrini, P. (2013). How the brain heals emotional wounds: The functional neuroanatomy of forgiveness. *Frontiers in Human Neuroscience*, *7*. doi:10.3389/fnhum.2013.00839

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*(5626), 1755–1758. doi:10.1126/science.1082976

Smith, S. M. (2002). Fast robust automated brain extraction. *Human Brain Mapping*, *17*(3), 143–155. doi:10.1002/hbm.10062

Smith, S. M., Jenkinson, M., Woolrich, M. W., Beckmann, C. F., Behrens, T. E., Johansen-Berg, H., … Matthews, P. M. (2004). Advances in functional and structural MR image analysis and implementation as FSL. *Neuroimage*, *23*(Suppl 1), S208–S219. doi:10.1016/j.neuroimage.2004.07.051

Stanley, D., Phelps, E., & Banaji, M. (2008). The neural basis of implicit attitudes. *Current Directions in Psychological Science*, *17*(2), 164–170. doi:10.1111/j.1467-8721.2008.00568.x

Strang, S., Utikal, V., Fischbacher, U., Weber, B., & Falk, A. (2014). Neural correlates of receiving an apology and active forgiveness: An fMRI study. *PLoS ONE*, *9*(2), 1–7. doi:10.1371/journal.pone.0087654

Will, G.-J., Crone, E. A., & Güroğlu, B. (2015). Acting on social exclusion: Neural correlates of punishment and forgiveness of excluders. *Social Cognitive and Affective Neuroscience*, *10*(2), 209–218. doi:10.1093/scan/nsu045

Worthington, E. L. (2006). *Forgiveness and reconciliation: Theory and application*. New York, NY: Routledge/Taylor & Francis Group.

Yarkoni, T., Poldrack, R. A., Nichols, T. E., Van Essen, D. C., & Wager, T. D. (2011). Large-scale automated synthesis of human functional neuroimaging data. *Nature Methods*, *8*(8), 665–670. doi:10.1038/nmeth.1635

Young, L., & Saxe, R. (2009). Innocent intentions: A correlation between forgiveness for accidental harm and neural activity. *Neuropsychologia*, *47*(10), 2065–2072. doi:10.1016/j.neuropsychologia.2009.03.020