

Responses of medial and ventrolateral prefrontal cortex to interpersonal conflict for resources

Leonie Koban,^{1,2} Swann Pichon,^{1,2} and Patrik Vuilleumier^{1,2}

¹Neurology and Imaging of Cognition, Neuroscience Center, University Medical Center, 1211 Geneva, Switzerland and ²Swiss Center for Affective Sciences, University of Geneva, 1205 Geneva, Switzerland

Little is known about brain mechanisms recruited during the monitoring and appraisal of social conflicts—for instance, when individuals compete with each other for the same resources. We designed a novel experimental task inducing resource conflicts between two individuals. In an event-related functional magnetic resonance imaging (fMRI) design, participants played with another human participant or against a computer, who across trials chose either different (no-conflict) or the same tokens (conflict trials) in order to obtain monetary gains. In conflict trials, the participants could decide whether they would share the token, and the resulting gain, with the other person or instead keep all points for themselves. Behaviorally, participants shared much more often when playing with a human partner than with a computer. fMRI results demonstrated that the dorsal mediofrontal cortex was selectively activated during human conflicts. This region might play a key role in detecting situations in which self- and social interest are incompatible and require behavioral adjustment. In addition, we found a conflict-related response in the right ventrolateral prefrontal cortex that correlated with measures of social relationship and individual sharing behavior. Taken together, these findings reveal a key role of these prefrontal areas for the appraisal and resolution of interpersonal resource conflicts.

Keywords: conflict; social cognition; dorsal mediofrontal cortex; ventrolateral prefrontal cortex

INTRODUCTION

Despite the frequency and sometimes important consequences of social conflicts in human life, the brain mechanisms that underlie the detection and resolution of conflicts with other individuals have rarely been explored. Social conflicts can result from an incompatibility of expectations, motivation, goals or values between two or more individuals or groups. In many situations, social conflicts reflect a competition for common and limited resources, goods or territories. As such, interpersonal (i.e. social) conflicts share similar properties with intrapersonal (i.e. cognitive) conflicts elicited by competition for cognitive resources, a situation more frequently studied in neuroscience (Botvinick *et al.*, 2004; Ridderinkhof *et al.*, 2004; Nee *et al.*, 2007) and typically related to an interference between concurrent processing pathways or response options (Botvinick *et al.*, 2001; Verguts and Notebaert, 2008). Moreover, whereas intrapersonal cognitive conflicts are important triggers for adjustments in cognitive and behavioral control (Botvinick *et al.*, 2001; Ridderinkhof *et al.*, 2004), the detection of and adjustment to interpersonal conflict signals are also likely to be critical for appropriate behavior in social contexts. In addition, both intra- and interpersonal conflicts are associated with negative affect (Huhman, 2006; Shackman *et al.*, 2011; Dreisbach and Fischer, 2012). However, despite these apparent similarities between cognitive (intrapersonal) and social (interpersonal) conflicts, it is unknown whether these two types of events recruit similar brain mechanisms.

Studies on the neurophysiological correlates of cognitive (intrapersonal) conflict have generally used interference paradigms where conflict is induced by an incongruence between different stimulus dimensions (e.g. Stroop task and flanker task) or between stimulus properties and behavioral response (e.g. Simon task), leading to greater recruitment of executive control systems (Botvinick *et al.*, 2001; Kerns

et al., 2004; Carter and van Veen, 2007). Brain imaging studies have shown that these interference effects lead to increased activity in the anterior cingulate cortex (ACC) or, more broadly, the dorsal mediofrontal cortex (dmFC), alongside with other regions in dorsolateral prefrontal cortex and anterior insula (AI) (Nee *et al.*, 2007). It is assumed that the ACC/dmFC subserves the detection of conflict and triggers subsequent adjustments in behavior, implemented by lateral prefrontal areas, thereby improving cognitive control and reducing interference in subsequent trials (Botvinick, 2007; but see also Holroyd and Coles, 2002; Nieuwenhuis *et al.*, 2002; Ridderinkhof *et al.*, 2004; Rushworth *et al.*, 2007).

Although interpersonal resource conflicts have not directly been studied, a few recent studies suggest that a similar network of brain regions might be involved during the monitoring of actions made by other people. Both the dmFC and AI activate not only to action errors caused by oneself but also to the observation of others' errors (Miltner *et al.*, 2004; van Schie *et al.*, 2004; de Bruijn *et al.*, 2009; Koban *et al.*, 2012b; Koban *et al.*, 2013). Likewise, dmFC responds to the observation of social threat (Pichon *et al.*, 2012), to incongruent social cues (Zaki *et al.*, 2010; Ruz and Tudela, 2011), social exclusion (Eisenberger *et al.*, 2003) and disagreement between subjective judgments relative to normative group opinions (see Campbell-Meiklejohn *et al.*, 2010; Klucharev *et al.*, 2009; Zaki *et al.*, 2011). Furthermore, in studies of social decision making in economic paradigms, such as the Ultimatum Game, the Dictator Game or the Prisoners' Dilemma (see reviews by Behrens *et al.*, 2009; Rilling *et al.*, 2008), the AI and dmFC regions were often found to activate in situations when one's own behavior is in conflict with others' expectations and social norms (Chang *et al.*, 2011; Chang and Sanfey, 2013). Unfair economic proposals by others are also associated with activation in these brain regions and increased physiological arousal (Sanfey *et al.*, 2003; Corradi-Dell'Acqua *et al.*, 2013). However, whether the dmFC and AI are implicated in the detection of interpersonal conflicts during competition for resources has not yet been addressed.

Importantly, the processing of social conflict situations depends partly on the interpersonal relationship with the other party. Conflict constitutes a potential risk for the social relationship, which requires

Received 7 October 2012; Accepted 12 February 2013

Advance Access publication 4 March 2013

We thank Christoph Hofstetter and Sebastian Rieger for help with data collection, and Corrado Corradi-Dell'Acqua and Benoit Bediou for helpful discussions. Research was supported by the National Center of Competence in Research (NCCR) Affective sciences financed by the Swiss National Science Foundation (No. 51NF40-104897) and hosted by the University of Geneva, and by a research award from the Evens Foundation.

Correspondence should be addressed to: Leonie Koban, NCCR Affective Sciences, University of Geneva, 7 rue des Battoirs, CH-1205 Geneva, Switzerland. E-mail: leonie.koban@gmail.com

efficient emotion regulation as well as conflict resolution between self-centered and other-regarding motives. Interestingly, one recent study showed that emotion regulation ability during conflict with one's romantic partner was predicted by the magnitude of activation in right ventrolateral prefrontal cortex (vlPFC) during observation of his/her emotional expressions (Hooker *et al.*, 2010). Activity in the latter region was also correlated with the acceptance rate of unfair offers in an Ultimatum Game (Tabibnia *et al.*, 2008). Taken together, these studies suggest that vlPFC may be important for regulating affective and behavioral responses to social conflict situations.

The present fMRI study aimed at investigating whether brain regions involved in cognitive conflict are also involved in the detection of interpersonal conflict between two individuals competing over limited resources, and whether such responses are influenced by inter-individual differences in social relationship factors (i.e. interpersonal closeness). We designed a novel task in which a participant in the scanner played with a confederate (social condition) or against a computer (non-social condition). In each trial, both players chose between two tokens at the same time before being presented with their respective choices. Critically, these choices could either be conflicting (if both players chose the same token) or non-conflicting (in case they did not). Following a conflict situation, the participant in the scanner was given the possibility to either keep the gain associated with the token or to share it with his/her co-player. Unbeknown to the participant in the scanner, choices of the second player were actually always generated by a computer, both in the social and non-social conditions. Therefore, these conditions were identical in all points except for the participant's consideration of the agent he was playing with. Furthermore, in order to investigate the influence of interpersonal relationship on conflict processing, we measured perceived interpersonal closeness with the 'Inclusion of Other into Self' scale (IOS; Aron *et al.*, 1992), and used a personality judgment task (with trait adjectives, Kelley *et al.*, 2002; Wagner *et al.*, 2011) to infer the degree of subjective similarity in self- and other-related attributes. Based on the hypothesis that cooperativeness is related to closeness and perceived similarity (Koban *et al.*, 2010), we expected more frequent 'share' decisions when the participant perceived the co-player as close and similar to him. In line with previous studies on cognitive and affective conflicts, we expected that interpersonal conflict would activate the dmPFC, together with other regions frequently recruited for errors and conflict monitoring, such as the AI and lateral prefrontal cortex.

METHODS

Participants

Twenty-two healthy participants (mean age 23.5 years, 10 females, 2 left-handed) were recruited at the University of Geneva. The second volunteer was played by a sex-matched confederate unfamiliar to the real subject. Participants met the confederate briefly in the beginning of the experiment (~5 min). The confederate, however, was not taking part in the task and a computer emulated his game instead. Four participants (three male) who, during the debriefing, expressed doubts about the identity of the confederate or about the computer condition (see below), were excluded from the sample as well as one male participant with abnormalities in brain structure that impeded preprocessing of the functional imaging data, resulting in a final sample size of 17 naive participants. All volunteers gave written informed consent and were paid for their time and participation. Further, they received an additional monetary bonus that depended on the points earned in a random subset of 10 experimental trials.

Stimuli and procedure

The participant and the confederate were briefed together. They were told that they would take part in a study on social decision making,

during which they had to interact via a computer network, one of them being in the magnetic resonance imaging (MRI) scanner and the second in another testing room. The participant in the scanner performed two sessions of a novel task that was designed to elicit interpersonal conflict situations. He/she was told that he would play some blocks of the task against the other volunteer, and some blocks against a pre-programmed script run by the computer. In fact, the choices and reactions of the co-player were always generated by the experimental script (controlled by E-prime software). As confirmed during careful post-experimental debriefing, none of the 17 naive participants reported being aware of this manipulation.

During fMRI, each trial (Figure 1) started with a fixation cross (2000–4000 ms, pseudo-random duration), followed by a central cue (1000 ms) that indicated the name of the current opponent (confederate or computer) as well as a target color (red, green or blue), which was selected randomly for each trial and participant. Participants were told that both players were assigned their own target color, randomly selected on each trial and associated with a varying amount of points, with each player being unaware of the other's color at the beginning of the trial. Thus, this initial choice was made independently of the second player (Figure 1). Next, two equiluminant but differently colored patterns were shown on the left and right side of the screen, and players had to select within 3 s their respective target color. Participants were told that their choices would yield points that would subsequently be transformed into bonus money after the session. They were instructed to collect as many points as possible, irrespective of choices made by the other player. After another jitter of 1000–2000 ms following the participant's response (indicated by a change in the corresponding arrow's color), the choice of the co-player was presented for 3000 ms, similarly indicated by an arrow (with a different color). The choices of the two players could therefore be either incompatible (same colors—conflict trial) or compatible (different colors—no-conflict trial) with each other. When the co-player chose a different color (no-conflict), each player won 100 points. In contrast, when their choices were in conflict, the participant *in the scanner* had to decide whether to share the points or to keep all points for him/herself. Importantly, participants were completely free how to decide (e.g. 'share' or 'keep') in case of conflict, with no possibility of the other player to punish or reject these decisions. The two options ('keep' or 'share') were presented on the left and right side of the screen (random position across trials), respectively, for a maximum of 3 s (Figure 1). Three seconds after the decision, the final outcome was presented centrally on the screen: 100 points following 'keep' decisions, or 50 points following 'share' decisions. Participants were told that the opponent would get no points in trials where they kept their chosen color and 100 points when they shared it. Please note that opponents had no opportunity to share or keep, leading to an asymmetric relationship between the two players following conflict trials, similar to a Dictator Game. To maintain motivation and keep the outcome phase engaging, the gains were unpredictably reduced to 5 or 10 points in a small proportion of trials (10%, independent of actual decisions); these trials were not further analyzed.

Participants started with a block of 48 trials against the confederate (HUMAN condition) and then 24 trials against the computer (COMP) in the first session, while the order was reversed in the second session (24 COMP trials followed by 48 HUMAN trials). Both conditions contained 50% conflict and 50% no-conflict trials, presented in a randomized order. Each of the two functional imaging sessions took 15 min.

Behavioral measures

Decision making in the conflict task was measured as the percentage of share *vs* keep decisions following conflict situations, separately for human and computer blocks.

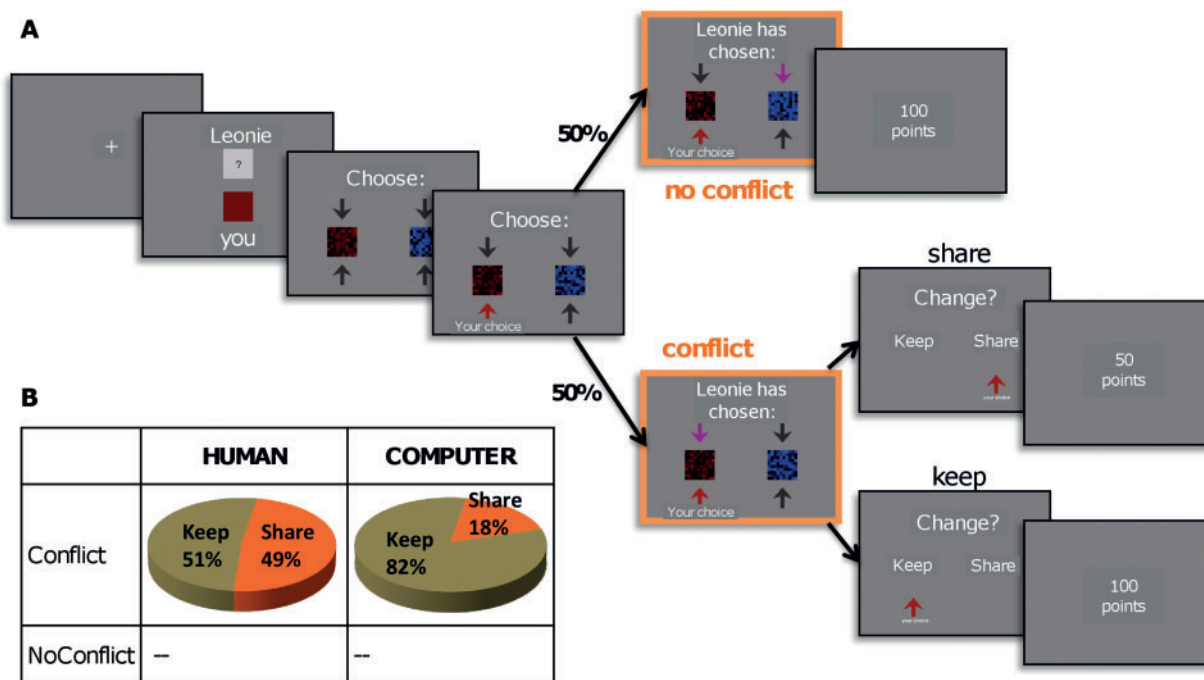


Fig. 1 Experimental design and behavioral results. (A) Each trial started with a cue indicating which color would yield points in this given trial. At this stage, the participant in the scanner (red token in the present example) does not know the other person’s color. Then, after a color token is selected by each player independently, the two respective choices are presented to the participant. This could lead to conflict when both players have chosen the same token (50% of trials) or to no-conflict when the other player has made a different and hence compatible choice (50% of trials). In no-conflict trials, each subject wins their respective token and points. In conflict trials, the participant in the scanner has to decide whether he/she wants to keep the token or share the points with the other player. Note that this lead to an asymmetric relationship between the two participants, as the opponent outside the scanner had no opportunity to refuse offers or to punish the participant in the scanner. (B) The 2 × 2 design included the factor CONDITION (human vs computer opponent) and the factor CONFLICT (conflict vs no-conflict). Participants shared significantly more often in the human compared with the computer condition.

In order to assess how participants perceived their co-player, they performed an additional mentalizing task probing the access to self and other representations (Wagner *et al.*, 2011). Using a Likert-like scale from 1 to 4, they had to rate the extent to which different personality trait adjectives described themselves (SELF condition) or the co-player (OTHER condition). A third (control) condition required indicating the number of syllables of a given adjective. We calculated average reaction times for each of the three conditions. We reasoned that the more different the other player is perceived from the self (i.e. the more adjustments is necessary from a direct self-projection, see Tamir and Mitchell, 2011), the larger the difference in reaction time (RT) would be between self and other judgments ($\Delta RT = RT_{Other} - RT_{Self}$).

Perceived relationship closeness was also measured after the MRI sessions with an adapted French version of the Inclusion of Other into Self Scale (Aron *et al.*, 1992). Affective responses to conflict and no-conflict trials were assessed with an emotion rating questionnaire (see Supplementary Material).

MRI image acquisition and analysis

Image acquisition and preprocessing followed standard practice and are described in the Supplementary Methods. We performed standard first-level analysis using the general linear model as implemented in SPM8. Fourteen regressors (convolved with a canonical hemodynamic response function) modeled the onsets of the different experimental events: the onsets of target-stimuli for the human and computer conditions, the four different conflict conditions (Human-Conflict, Human-No Conflict, Computer-Conflict and Computer-No Conflict), the different decisions to share or insist on the choice (Human-Keep, Human-Share and Computer-Keep), as well as the different outcomes. We did not include ‘Computer-Share’ events as separate conditions (at decision and outcome stage), as they were too

few for most of the participants (mean = 4.1 trials, s.d. = 4.3 trials). Time derivatives of the 14 regressors of interest were added to the model in order to correct for slice time differences. Six regressors of no interest corrected for movement artifacts. The model included a high-pass filter to reduce low-frequency noise (cutoff 128 s).

The statistical estimation of model parameters used restricted maximum likelihood and an autoregressive AR(1) model to account for temporal autocorrelation. On the level of individual participants, contrast images were calculated for the four different conflict conditions (Human-Conflict, Human-No Conflict, Comp-Conflict and Comp-No Conflict). At the group level, the resulting contrast images were submitted to a second-level random effects factorial model. Contrasts for the decision phase (e.g. share vs keep decisions) were not further investigated, as the resulting activation maps did not yield any significant effects ($P < 0.001$). To test for a modulation of conflict-specific activations by personal and relationship variables, we calculated additional conflict-specific contrasts (Human-Conflict > Human-No Conflict and Comp-Conflict > Comp-No Conflict), using second-level one-sided *t*-tests which included the three covariates of interest (IOS, ΔRT and percentage of share decisions/%share). Group-level *t* maps were thresholded at $P < 0.001$ (uncorrected) with an extent threshold on $k > 50$ voxel (corresponding to a volume of 400 mm³), in line with previous fMRI studies on social decision making (e.g. Tabibnia *et al.*, 2008).

RESULTS

Behavioral results

During the fMRI task, behavior during conflict trials was strongly modulated as a function of the two experimental conditions. Participants shared significantly more often (and thus kept less) when they believed that they were playing with another human

participant (48.5% share decisions) when compared with a computer (17.5% share, $t(16) = 5.2$, $P < 0.001$). Moreover, the interindividual variation in the amount of sharing was very high, especially when playing with another human, ranging from 2 to 96%, pointing to the interest of taking into account this interindividual variance for our subsequent covariance analysis. RTs did not differ between types of decisions (human-keep: 917 ms, human-share: 930 ms, computer-keep: 897 ms, $F(2,28) = 0.33$, $P = 0.73$), ruling out the possibility that modulation of brain activations (see below) in the different conflict conditions could be driven by RT differences (Grinband et al., 2011).

For the self-other judgment task, RTs were faster for SELF ratings (1595 ms) than OTHER ratings (1689 ms), resulting in an average ΔRT of 94 ms ($t(16) = 2.5$, $P = 0.024$). Against our expectations, neither the ΔRT ($r = -0.08$, n.s.) nor the percentage of sharing ($r = -0.19$, n.s.) was correlated with the subjective interpersonal closeness score as measured by IOS. However, ΔRT was negatively correlated with the decision making (% share) in the human conflict condition ($r = -0.72$, $P < 0.001$), indicating that a smaller ΔRT between self and other was linked to a higher percentage of 'share' decisions (and fewer 'keep' decisions).

Functional imaging results

We first examined the main effect of (Conflict > No Conflict). This showed prominent activations in bilateral AI, ACC and supplementary motor area, as well as putamen and right cerebellum (Table 1). Activations for (Conflict > No Conflict) are depicted separately for the HUMAN and the COMPUTER condition in Figure 2, showing partly overlapping activations. These effects were generally larger and extended more dorsally in the medial frontal cortex for the human condition. A conjunction analysis [Human (Conflict > No Conflict)] \cap [Computer (Conflict > No Conflict)] formally confirmed common activations in bilateral AI for both HUMAN and COMPUTER conflicts (see Table 3 and Figure 3B and C). Common effects in ACC did not reach significance.

Next, we directly tested for the CONDITION \times CONFLICT interaction effects [Human (Conflict > No Conflict)] > [Computer (Conflict > No Conflict)], in order to pull apart activations specific to the human conflict situations. This contrast yielded a single cluster in the dMFC, centered on the medial superior frontal gyrus and pre-SMA (Table 2 and Figure 3A). Pairwise comparisons on the extracted average beta-values across the whole cluster confirmed that activity for Conflict vs No Conflict was different only for the HUMAN

Table 1 Main effect CONFLICT [Conflict > No Conflict]

Regions	Left/ right	Peak voxel MNI coordinates (mm)			Cluster size (voxels)	Peak <i>t</i> -value
		x	y	z		
Anterior insula, inferior frontal gyrus	L	-36	16	-2	631 ^a	5.91
		-36	22	-8		5.80
		-48	26	-4		3.81
Anterior insula, inferior frontal gyrus	R	32	26	-6	774 ^a	5.88
		44	20	-2		5.50
Anterior cingulate cortex	R	10	28	24	649 ^a	5.37
		10	34	14		4.52
		6	22	32		4.30
Pallidum, nucleus caudatus	R	16	2	-4	260 ^a	5.09
		14	6	8		4.77
		12	14	10		4.16
Thalamus	L	-10	-20	8	188 ^a	4.98
Cerebellum	R	34	-52	-34	158 ^a	4.68
		40	-60	-30		4.16
		46	-52	-32		3.35
Putamen	L	-14	4	10	128	4.19
		-18	8	0		3.79
Precuneus	R	8	-66	42	51	4.04
Angular gyrus	R	54	-54	36	53	3.99
		58	-48	28		3.52

All activations are reported at a threshold of $P < 0.001$ uncorrected, extent threshold $k > 50$.
^aClusters surviving a FWE-corrected threshold of $P < 0.05$.

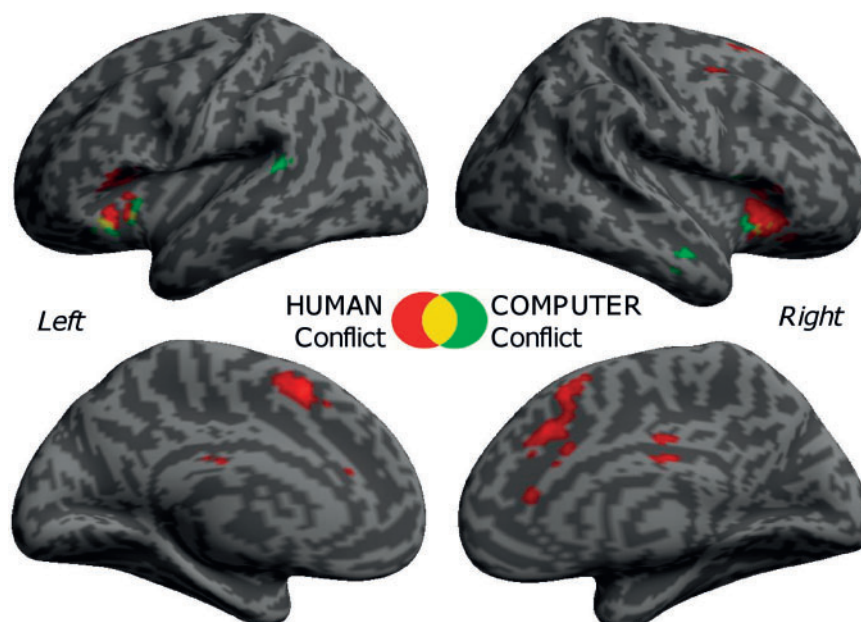


Fig. 2 Brain responses to conflict trials separately for the human and computer conditions. The statistical maps show increased BOLD activity during conflict with another human choice in red (Human [Conflict > No Conflict]) and conflict with a computer in green (Computer [Conflict > No Conflict]). Overlapping activations are displayed in yellow. Thresholds of activation maps are set at $P < 0.001$ with a minimum cluster extent of 50 voxels. Maps were rendered on partially inflated lateral and medial views of the two cerebral hemispheres.

(Bonferroni-corrected $P=0.001$), but not for the COMPUTER condition ($P=1.0$).

To determine the influence of interpersonal relationship on brain responses to social conflicts, we tested for any modulation of these effects by social appraisal factors derived from our behavioral measures and questionnaires. Individual data concerning subjective interpersonal closeness (IOS scale), self-other mentalizing differences (ΔRT) and percent of share decisions on conflict trials were included as covariates in a multiple regression model using the contrast [Human (Conflict > No Conflict)] in the second-level group analysis (Figure 4).

Interpersonal closeness as measured with IOS correlated positively with increases in dorsal cingulate cortex and right vLPFC (inferior-anterior part of the middle frontal gyrus, see Table 4 for full results). Thus, the closer the relationship between the two players, the stronger the response in those regions during conflict (*vs* no-conflict trials). No negative correlations were found.

On the other hand, the ΔRT reflecting self-other judgment differences was negatively related to conflict-activations in the same areas of the right vLPFC, plus the superior frontal gyrus (Table 4), indicating that smaller reaction time differences between self- and other-judgments (likely to reflect greater subjective similarity) were

associated with stronger conflict-related activity in these two areas. We found no positive correlations for this contrast. Similarly, the percentage of share decisions again highlighted one significant cluster in right vLPFC, correlating negatively with the percentage of share decisions on human conflict trials (Table 4).

Remarkably, all these covariates modulated conflict-related activation in a partly overlapping area of the right vLPFC, more specifically, in the ventral portion of the right middle frontal gyrus (Figure 4). Therefore, we performed an intermediate conjunction analyses (Friston *et al.*, 2005) across these three contrasts, which yielded one cluster in the right vLPFC (peak at $x=38, y=46, z=-4; t=4.29; k=169$), and thus confirmed that this area was commonly modulated by IOS, ΔRT and percentage of sharing. Figure 4B provides a four-dimensional scatter plot of activation parameters (extracted betas averaged across voxels) from this conjunction cluster, illustrating the joint contribution of these three different behavioral measures. Importantly, none of the three covariates was significantly correlated with vLPFC activity in the computer conflict conditions, indicating that these effects were unique to the human conflict condition.

Table 2 Interaction effect CONFLICT \times CONDITION: [Human (Conflict > No Conflict)] > [Computer (Conflict > No Conflict)]

Regions	Left/ right	Peak voxel MNI coordinates (mm)			Cluster size (voxels)	Peak t-value
		x	y	z		
		Dorsal mediofrontal cortex	R/L	-10		
		-2	12	54		3.70
		6	16	52		3.66

Table 3 Conjunction analysis [HUMAN (Conflict > No Conflict)] AND [COMPUTER (Conflict > No Conflict)]

Regions	Left/ right	Peak voxel MNI coordinates (mm)			Cluster size (voxels)	Peak t-value
		x	y	z		
Anterior insula	L	-36	22	-8	147	4.47
		-38	16	-2		
Anterior insula	R	48	18	0	111	4.03
		34	18	-10		
		30	24	-6		

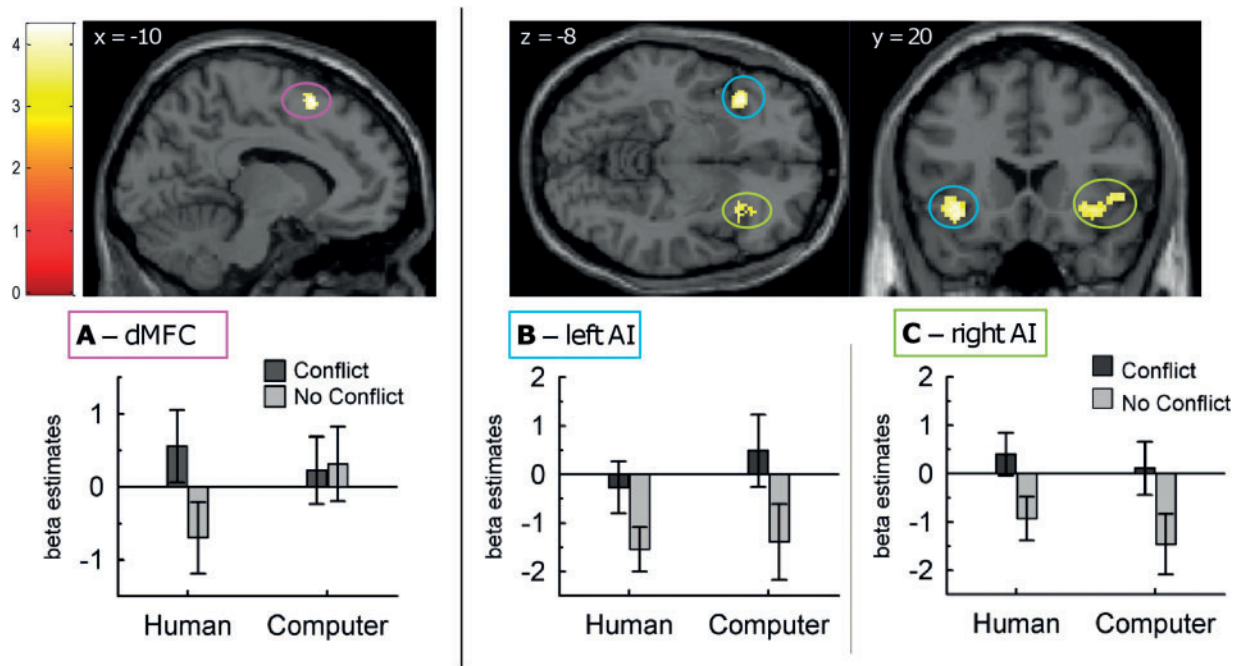


Fig. 3 Dissociation between dMFC and AI for interpersonal conflict monitoring. (A) Interaction effect in the dMFC [Human Conflict > No Conflict] > [Computer Conflict > No Conflict], thresholded at $P < 0.001$, extent threshold of $k > 50$ voxels. This cluster in dMFC was differentially activated for human resource conflict vs no-conflict, but not for computer conflicts. (B and C) The conjunction analysis [Human Conflict > No Conflict] \cap [Computer Conflict > No Conflict] demonstrated that the left (B) and right AI (C) were activated for conflict compared with no-conflict trials in both the human and the computer conditions. Vertical bars denote standard errors.

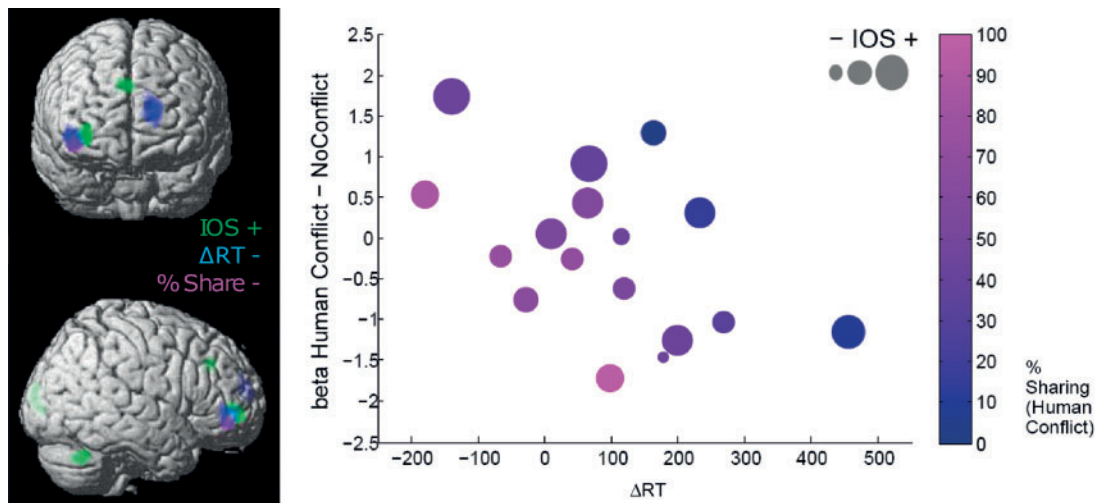


Fig. 4 Whole-brain covariate analysis. Interpersonal closeness, mentalizing RT difference and percentage of ‘share’ vs ‘keep’ decisions were found to correlate with the contrast [Human Conflict > No Conflict] in an overlapping cluster in the right vIPFC. To illustrate this multivariate correlation, the relationships between all four measures are plotted in a four-dimensional scatterplot (one point describes one subject). vIPFC beta estimates and ΔRT are plotted on the y - and x -axes, respectively, whereas the point color illustrates conflict behavior (% share decision), and the point size indicates interpersonal closeness (IOS questionnaire).

Table 4 Regression of the contrast [HUMAN (Conflict vs No Conflict)] with behavioral measures: interpersonal closeness (IOS), self-other mentalizing distance (ΔRT), and percentage of sharing in the human condition

Measure and Regions	Left/ right	Peak voxel MNI coordinates (mm)			Cluster size (voxels)	Peak t -value
		x	y	z		
IOS—positive						
Cingulate cortex	R/L	8	32	36	77	6.40
Cerebellum	R	36	-60	-34	70	5.85
Ventrolateral prefrontal cortex	R	32	50	0	113	5.71
Cerebellum	L	-34	-70	-32	63	5.64
Cuneus	L	-8	-94	8	102	5.37
		-6	-96	18		5.25
		-6	-92	0		4.45
ΔRT—negative						
Superior frontal gyrus	L	-18	56	20	194	7.47
		-16	62	12		6.25
		-10	56	16		5.51
Ventrolateral prefrontal cortex	R	44	48	-2	185	6.30
		40	42	-6		5.58
		34	44	0		4.51
Percentage of sharing—negative						
Ventrolateral prefrontal cortex	R	40	44	-10	63	5.01

Reported activations are $P < 0.001$ uncorrected (threshold of $k > 50$ voxels per cluster).

DISCUSSION

We designed a new experimental paradigm to investigate the brain mechanisms recruited during the perception and resolution of social conflict involving resource competition—a common and fundamental source of contention between individuals or groups in real life. Participants made choices that could be compatible or incompatible with those of their partner and had then to decide whether to share or keep their potential gains. Behavioral results showed that people shared more when they assumed to be playing with another human, rather than with a computer. We also observed that the propensity to share was associated with a smaller RT difference when making personality judgments about oneself vs the other player (an implicit measure of perceived self-other similarity) during an independent mentalizing

task. These behavioral findings provide evidence that participants differentially weighted self-centered and other-regarding motives in response to resource conflict during the virtual human interactions in our paradigm.

Brain imaging results obtained during this paradigm revealed a functional dissociation in the neural network typically recruited during conflict monitoring in cognitive tasks. On the one hand, we found an activation of bilateral AI that was common to both human and computer conflicts, when compared with no-conflict trials. On the other hand, the activation of one area within pre-SMA/dMFC was specific to human (compared with non-human) conflict situations. Furthermore, a region in the right vIPFC selectively responded to human conflict vs no-conflict as a function of the quality of social relationship with the co-player and the individual propensity to resource sharing.

So far, most of the neuroscience research on conflict monitoring has focused on cognitive/intrapersonal levels of response or processing competition, with strong evidence that a dedicated network encompassing dMFC, dlPFC and insula is crucially implicated in the monitoring of cognitive and affective conflicts (Botvinick *et al.*, 2001; Ochsner *et al.*, 2009; Chiew and Braver, 2011). Our current findings extend this literature by showing that the dMFC is also implicated in the detection of interpersonal conflicts due to competition for common gains, which require a resolution between self-interest and social motives. Although this study is the first to investigate the monitoring of social resource conflicts, our results are consistent with recent work on other social conflict situations that reported dMFC activation in response to incongruent social cues (Zaki *et al.*, 2010; Ruz and Tudela, 2011), and to conflict between one’s own and group opinion (Sanfey *et al.*, 2003; Klucharev *et al.*, 2009; Campbell-Meiklejohn *et al.*, 2010; Chang and Sanfey, 2013). It has been suggested that the dMFC may respond to conflicts in a domain-general manner for various kinds of interference during cognitive and affective processing (Ochsner *et al.*, 2009; Koban *et al.*, 2012a). Accordingly, a much stronger and extensive activation was elicited in dMFC, in particular more dorsally in the superior medial frontal gyrus (Figures 2 and 3), during interpersonal conflicts with the human co-player. Only during the human conflict trials were our participants confronted with a true conflict between different intra- and interpersonal motivations,

namely self-centered monetary incentive *vs* other-regarding social sharing. Interestingly, nearby regions in medial frontal gyrus (dorsal to ACC proper) have previously been related to mentalizing and altruistic behavior (Waytz *et al.*, 2012). The selective increases of dmFC in the human conflict condition might reflect a key role of this region in signaling the need for behavioral adjustment with regard to the goals of the other co-player and social norms. It remains to be clarified whether dmFC comprises different subareas, which are domain-specific for cognitive, affective or social conflict situations, or for other events requiring adjustments in behavioral and cognitive control (Shackman *et al.*, 2011; Koban *et al.*, 2012a).

The AI showed a different response pattern than dmFC. In this region, we found increased activation in conflict *vs* no-conflict trials, across both the human and computer conditions. Next to dmFC and dlPFC, the AI has consistently been associated to cognitive control during interference and conflict processing (Nee *et al.*, 2007; Higo *et al.*, 2011), as well as to various other affective functions (Kober *et al.*, 2008). The role of AI in conflict monitoring is still poorly understood but has been linked to emotional experience or awareness associated with action control and errors (Ullsperger *et al.*, 2010; Koban *et al.*, 2013). In the context of social decision-making paradigms, such as the ultimatum game, activity in the AI has been related to perceived unfairness (Sanfey *et al.*, 2003) and violation of social norms and expectations (Chang *et al.*, 2011; Koban *et al.*, 2013). AI has also been linked to empathy (Singer *et al.*, 2004) and feeling state prediction (Singer *et al.*, 2009). However, in the present task, the AI responses were not specific to social conflict monitoring, but appeared to reflect more basic processes similarly recruited in the computer condition. These effects could perhaps be related to negative affect or arousal signals evoked by the detection of any conflict with own choices. Differences in the trial structure between conflict and non-conflict trials might also explain this general pattern of activity, for instance due to the need of preparing a further motor response or the maintenance of task-related activity (Sridharan *et al.*, 2008) in the presence of conflict cues, but not in the absence of conflict (where outcome feedback was immediately presented).

Finally, we observed a selective modulation of the right vlPFC to social conflict by several distinct measures of interpersonal relationship, including self-other similarity and mentalizing processes, as well as by the actual social decision making of individual participants. The neural response to conflict in this region was enhanced by interpersonal closeness (as determined on the IOS scale), and by faster judgments of personal attributes for the co-player [as indicated by smaller differences in mentalizing reaction times for SELF- *vs* OTHER (see Tamir and Mitchell, 2011)]. Additionally, activation to conflict *vs* no-conflict situations in the right vlPFC was also correlated with the percentage of keep *vs* share decisions. An activation of this region has frequently been reported for emotional processing, especially in tasks involving affect labeling and emotion regulation by cognitive reappraisal (Ochsner and Gross, 2005; Lieberman *et al.*, 2007; Kober *et al.*, 2008; Wager *et al.*, 2008; Kanske *et al.*, 2011). Recent studies also demonstrated that the vlPFC is involved in the implicit regulation of emotion and behavior in the context of social relationships (Tabibnia *et al.*, 2008; Hooker *et al.*, 2010; Meyer *et al.*, 2011). In particular, Meyer *et al.* (2011) suggested that this region may be crucial for the incidental regulation of potential threats to a relationship. Interestingly, the vlPFC peak activation reported in their study is only 5.7 mm off the peak voxel found here. One plausible interpretation for our results is that the right vlPFC could mediate the regulation and resolution of social conflict situations, as a function of the relationship with the other individual and the enforcement of self-centered motives (keeping *vs* sharing resources). Thus, participants who felt closer and mentalized about the other in a more self-based

way (Tamir and Mitchell, 2011), but nevertheless also made a high amount of 'keep' decisions, may have recruited the right vlPFC to a greater extent in order to reduce conflict or cognitive dissonance (Festinger, 1957; van Veen *et al.*, 2009) between self-centered monetary interests and social motives, allowing them in turn to make a higher amount of 'keep' decisions. Whether this activation reflects greater ability to integrate self-centered and other-regarding behavior in social context, or to disregard the goals of others when they are subjectively perceived as close to one's own goals, will require further investigations using variants of similar paradigms with more specific manipulation of social factors and moral norms.

It is noteworthy that we did not observe increased activation of the dorsolateral PFC in response to conflict situations, as this region has been frequently implicated in both cognitive control and in social or moral decision making (e.g. Greene *et al.*, 2004; Knoch *et al.*, 2009; Baumgartner *et al.*, 2011). This could be due to the character of our decision-making task. Participants in our task were free to decide according to their social preferences without any possibility for their opponent to reject this decision. Therefore, less cognitive control than in other social decision-making situations might be needed. To our knowledge, only two recent studies have investigated brain activation of *allocators* in the Dictator Game, which resembles the decision phase in our experiment. Weiland *et al.* (2012) compared brain activations correlated to fair behavior in the Ultimatum and the Dictator Game and found more lateral prefrontal and parietal activity during the Ultimatum game, which could be related to the higher requirement for strategical decision making (Weiland *et al.*, 2012). A second recent study (Gunther Moor *et al.*, 2012) investigated brain activity during a Dictator Game, which participants played against individuals who had previously included or rejected them in a ball-tossing game (i.e. Cyberball game, Eisenberger *et al.*, 2003). Interestingly, Gunther Moor *et al.* reported increased activity of the left TPJ, the right STS and the bilateral vlPFC when participants made offers to players that previously excluded them. These results suggest that activity in these regions may be mediating social punishment behavior (Rilling *et al.*, 2008; Gunther Moor *et al.*, 2012) and are also in line with the idea that vlPFC may be important in the regulation of opposing social goals.

We note that our design has some limitations that could be improved in future studies. Firstly, as noted above, a possible confound between conflict and no-conflict trials is the difference in trial structure. In order to create a semi-realistic task, no-conflict situations included no further decision stage (share/keep), as opposed to conflict trials. Thus, no-conflict situations were associated with a shorter temporal delay of reward outcome (won points), and possibly also with higher reward expectation. However, in order to control for this possible confound, we included a non-social computer condition in our design. In addition, we cannot definitely distinguish whether the greater activation (main effect) for conflict *vs* no-conflict trials in AI, thalamus and cerebellum was driven by the conflict detection itself, or by the need for a new decision and subsequent motor response. For example, it is possible that activity in AI, thalamus and cerebellum might relate to a general increase in arousal, non-specific effort or expected delay of reward. However, importantly, the interaction effect in dmFC was specifically driven by conflicts in the human condition and, therefore, unlikely to be explained by these factors. Secondly, our homogenous population and relatively modest sample size do not allow for strong claims regarding individual differences concerning the modulation of right vlPFC by personality factors and social relationships. Nevertheless, taken together with other studies, our findings point to an important role of this region in affective or regulative aspects of interpersonal behavior, particularly during conflict processing and resolution (Tabibnia *et al.*, 2008; Hooker *et al.*, 2010; Meyer *et al.*, 2011).

More generally, our study may open new perspectives to understand brain mechanisms responsible for the appraisal and response to conflict between choices, goals, or values of different individuals or even groups. These processes are crucial components of behavior and moral norms in human societies. Future studies are needed to clarify how responses in dmFC, AI and vMPFC are modulated during social conflict detection and resolution, and how they interact with social and individual factors related to mentalizing abilities, interpersonal relationship as well as various personality or even genetic characteristics.

SUPPLEMENTARY DATA

Supplementary data are available at SCAN online.

REFERENCES

- 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), 596–612.
- Baumgartner, T., Knoch, D., Hotz, P., Eisenegger, C., Fehr, E. (2011). Dorsolateral and ventromedial prefrontal cortex orchestrate normative choice. *Nature Neuroscience*, 14(11), 1468–74.
- Behrens, T.E., Hunt, L.T., Rushworth, M.F. (2009). The computation of social behavior. *Science*, 324(5931), 1160–4.
- 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), 624–52.
- Botvinick, M.M., Cohen, J.D., Carter, C.S. (2004). Conflict monitoring and anterior cingulate cortex: an update. *Trends in Cognitive Sciences*, 8(12), 539–46.
- Botvinick, M.M. (2007). Conflict monitoring and decision making: reconciling two perspectives on anterior cingulate function. *Cognitive, Affective & Behavioral Neuroscience*, 7(4), 356–66.
- Campbell-Meiklejohn, D.K., Bach, D.R., Roepstorff, A., Dolan, R.J., Frith, C.D. (2010). How the opinion of others affects our valuation of objects. *Current Biology*, 20(13), 1165–70.
- Carter, C.S., van Veen, V. (2007). Anterior cingulate cortex and conflict detection: an update of theory and data. *Cognitive, Affective & Behavioral Neuroscience*, 7(4), 367–79.
- 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.
- Chang, L.J., Smith, A., Dufwenberg, M., Sanfey, A.G. (2011). Triangulating the neural, psychological, and economic bases of guilt aversion. *Neuron*, 70(3), 560–72.
- Chiew, K.S., Braver, T.S. (2011). Neural circuitry of emotional and cognitive conflict revealed through facial expressions. *PLoS One*, 6(3), e17635.
- Corradi-Dell'Acqua, C., Civali, 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(4), 424–31.
- de Bruijn, E.R., de Lange, F.P., von Cramon, D.Y., Ullsperger, M. (2009). When errors are rewarding. *The Journal of Neuroscience*, 29(39), 12183–6.
- Dreisbach, G., Fischer, R. (2012). Conflicts as aversive signals. *Brain and Cognition*, 78(2), 94–8.
- Eisenberger, N.I., Lieberman, M.D., Williams, K.D. (2003). Does rejection hurt? An FMRI study of social exclusion. *Science*, 302(5643), 290–2.
- Festinger, L. (1957). *A Theory of Cognitive Dissonance*. Stanford, CA: Stanford Univ Press.
- Friston, K.J., Penny, W.D., Glaser, D.E. (2005). Conjunction revisited. *NeuroImage*, 25(3), 661–7.
- Greene, J.D., Nystrom, L.E., Engell, A.D., Darley, J.M., Cohen, J.D. (2004). The neural bases of cognitive conflict and control in moral judgment. *Neuron*, 44(2), 389–400.
- Grinband, J., Savitskaya, J., Wager, T.D., Teichert, T., Ferrera, V.P., Hirsch, J. (2011). The dorsal medial frontal cortex is sensitive to time on task, not response conflict or error likelihood. *NeuroImage*, 57(2), 303–11.
- Gunther, M.G., Bregtje, B., Op de Macks, Z.A., Rombouts, S.A., Van der Molen, M.W., Crone, E.A. (2012). Social exclusion and punishment of excluders: neural correlates and developmental trajectories. *NeuroImage*, 59(1), 708–17.
- Higo, T., Mars, R.B., Boorman, E.D., Buch, E.R., Rushworth, M.F. (2011). Distributed and causal influence of frontal operculum in task control. *Proceedings of the National Academy of Sciences of the United States of America*, 108(10), 4230–5.
- Holroyd, C.B., Coles, M.G. (2002). The neural basis of human error processing: reinforcement learning, dopamine, and the error-related negativity. *Psychological Review*, 109(4), 679–709.
- Hooker, C.I., Gyurak, A., Verosky, S.C., Miyakawa, A., Ayduk, O. (2010). Neural activity to a partner's facial expression predicts self-regulation after conflict. *Biological Psychiatry*, 67(5), 406–13.
- Huhman, K.L. (2006). Social conflict models: can they inform us about human psychopathology? *Hormones and Behavior*, 50(4), 640–6.
- Kanske, P., Heissler, J., Schönfelder, S., Bongers, A., Wessa, M. (2011). How to regulate emotion? Neural networks for reappraisal and distraction. *Cerebral Cortex*, 21(6), 1379–88.
- Kelley, W.M., Macrae, C.N., Wyland, C.L., Caglar, S., Inati, S., Heatherton, T.F. (2002). Finding the self? An event-related fMRI study. *Journal of Cognitive Neuroscience*, 14(5), 785–94.
- Kerns, J.G., Cohen, J.D., MacDonald, A.W.3rd, Cho, R.Y., Stenger, V.A., Carter, C.S. (2004). Anterior cingulate conflict monitoring and adjustments in control. *Science*, 303(5660), 1023–6.
- Klucharev, V., Hytönen, K., Rijpkema, M., Smidts, A., Fernández, G. (2009). Reinforcement learning signal predicts social conformity. *Neuron*, 61(1), 140–51.
- Knoch, D., Schneider, F., Schunk, D., Hohmann, M., Fehr, E. (2009). Disrupting the prefrontal cortex diminishes the human ability to build a good reputation. *Proceedings of the National Academy of Sciences of the United States of America*, 106(49), 20895–9.
- Koban, L., Brass, M., Lynn, M.T., Pourtois, G. (2012a). Effects of placebo analgesia on brain correlates of error monitoring. *PLoS One*, 7(11), e49784.
- Koban, L., Corradi-Dell'Acqua, C., Vuilleumier, P. (2013). Integration of error agency and representation of others' pain in the Anterior Insula. *Journal of Cognitive Neuroscience*, 25(2), 258–72.
- Koban, L., Pourtois, G., Bediou, B., Vuilleumier, P. (2012b). Effects of social context and predictive relevance on action outcome monitoring. *Cognitive, Affective & Behavioral Neuroscience*, 12(3), 460–78.
- Koban, L., Pourtois, G., Vocat, R., Vuilleumier, P. (2010). When your errors make me lose or win: event-related potentials to observed errors of cooperators and competitors. *Social Neuroscience*, 5(4), 360–74.
- Kober, H., Barrett, L.F., Joseph, J., Bliss-Moreau, E., Lindquist, K., Wager, T.D. (2008). Functional grouping and cortical-subcortical interactions in emotion: a meta-analysis of neuroimaging studies. *NeuroImage*, 42(2), 998–1031.
- Lieberman, M.D., Eisenberger, N.I., Crockett, M.J., Tom, S.M., Pfeifer, J.H., Way, B.M. (2007). Putting feelings into words: affect labeling disrupts amygdala activity in response to affective stimuli. *Psychological Science*, 18(5), 421–8.
- Meyer, M.L., Berkman, E.T., Karremans, J.C., Lieberman, M.D. (2011). Incidental regulation of attraction: the neural basis of the derogation of attractive alternatives in romantic relationships. *Cognition & Emotion*, 25(3), 490–505.
- Miltner, W.H.R., Brauer, J., Hecht, H., Trippe, R., Coles, M.G.H. (2004). Parallel brain activity for self-generated and observed errors. In: Ullsperger, M., Falkenstein, M., editors. *Errors, Conflicts, and the Brain. Current Opinions on Performance Monitoring*. Leipzig: MPI of Cognitive Neuroscience, pp. 124–9.
- Nee, D.E., Wager, T.D., Jonides, J. (2007). Interference resolution: insights from a meta-analysis of neuroimaging tasks. *Cognitive, Affective & Behavioral Neuroscience*, 7(1), 1–17.
- Nieuwenhuis, S., Ridderinkhof, K.R., Talsma, D., Coles, M.G., Holroyd, C.B., Kok, A., van der Molen, M.W. (2002). A computational account of altered error processing in older age: dopamine and the error-related negativity. *Cognitive, Affective & Behavioral Neuroscience*, 2(1), 19–36.
- Ochsner, K.N., Gross, J.J. (2005). The cognitive control of emotion. *Trends in Cognitive Sciences*, 9(5), 242–9.
- Ochsner, K.N., Hughes, B., Robertson, E.R., Cooper, J.C., Gabrieli, J.D. (2009). Neural systems supporting the control of affective and cognitive conflicts. *Journal of Cognitive Neuroscience*, 21, 1842–55.
- Pichon, S., de Gelder, B., Grèzes, J. (2012). Threat prompts defensive brain responses independently of attentional control. *Cerebral Cortex*, 22(2), 274–85.
- Ridderinkhof, K.R., Ullsperger, M., Crone, E.A., Nieuwenhuis, S. (2004). The role of the medial frontal cortex in cognitive control. *Science*, 306(5695), 443–7.
- Rilling, J.K., King-Casas, B., Sanfey, A.G. (2008). The neurobiology of social decision-making. *Current Opinion in Neurobiology*, 18(2), 159–65.
- Rushworth, M.F.S., Behrens, T.E.J., Rudebeck, P.H., Walton, M.E. (2007). Contrasting roles for cingulate and orbitofrontal cortex in decisions and social behaviour. *Trends in Cognitive Sciences*, 11(4), 168–76.
- Ruz, M., Tudela, P. (2011). Emotional conflict in interpersonal interactions. *NeuroImage*, 54(2), 1685–91.
- 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), 1755–8.
- Shackman, A.J., Salomons, T.V., Slagter, H.A., Fox, A.S., Winter, J.J., Davidson, R.J. (2011). The integration of negative affect, pain and cognitive control in the cingulate cortex. *Nature Reviews Neuroscience*, 12(3), 154–67.
- Singer, T., Critchley, H.D., Preuschoff, K. (2009). A common role of insula in feelings, empathy and uncertainty. *Trends in Cognitive Sciences*, 13(8), 334–40.
- Singer, T., Seymour, B., O'Doherty, J., Kaube, H., Dolan, R.J., Frith, C.D. (2004). Empathy for pain involves the affective but not sensory components of pain. *Science*, 303(5661), 1157–62.
- Sridharan, D., Levitin, D.J., Menon, V. (2008). A critical role for the right fronto-insular cortex in switching between central-executive and default-mode networks. *Proceedings of the National Academy of Sciences of the United States of America*, 105(34), 12569–74.

- 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(4), 339–47.
- Tamir, D.I., Mitchell, J.P. (2011). Neural correlates of anchoring-and-adjustment during mentalizing. *Proceedings of the National Academy of Sciences of the United States of America*, 107(24), 10827–32.
- Ullsperger, M., Harsay, H.A., Wessel, J.R., Ridderinkhof, K.R. (2010). Conscious perception of errors and its relation to the anterior insula. *Brain Structure & Function*, 214(5-6), 629–43.
- van Schie, H.T., Mars, R.B., Coles, M.G., Bekkering, H. (2004). Modulation of activity in medial frontal and motor cortices during error observation. *Nature Neuroscience*, 7(5), 549–54.
- van Veen, V., Krug, M.K., Schooler, J.W., Carter, C.S. (2009). Neural activity predicts attitude change in cognitive dissonance. *Nature Neuroscience*, 12(11), 1469–74.
- Verguts, T., Notebaert, W. (2008). Hebbian learning of cognitive control: dealing with specific and nonspecific adaptation. *Psychological Review*, 115(2), 518–25.
- Wager, T.D., Davidson, M.L., Hughes, B.L., Lindquist, M.A., Ochsner, K.N. (2008). Prefrontal-subcortical pathways mediating successful emotion regulation. *Neuron*, 59(6), 1037–50.
- Wagner, U., N'Diaye, K., Ethofer, T., Vuilleumier, P. (2011). Guilt-specific processing in the prefrontal cortex. *Cerebral Cortex*, 21(11), 2461–70.
- Weiland, S., Hewig, J., Hecht, H., Mussel, P., Miltner, W.H. (2012). Neural correlates of fair behavior in interpersonal bargaining. *Social Neuroscience*, 7, 537–51.
- Waytz, A., Zaki, J., Mitchell, J.P. (2012). Response of Dorsomedial Prefrontal Cortex Predicts Altruistic Behavior. *The Journal of Neuroscience*, 32, 7646–50.
- Zaki, J., Hennigan, K., Weber, J., Ochsner, K.N. (2010). Social cognitive conflict resolution: contributions of domain-general and domain-specific neural systems. *The Journal of Neuroscience*, 30(25), 8481–8.
- Zaki, J., Schirmer, J., Mitchell, J.P. (2011). Social influence modulates the neural computation of value. *Psychological Science*, 22(7), 894–900.