



# Neural correlates of dynamically evolving interpersonal ties predict prosocial behavior

Johannes J. Fahrenfort<sup>1,2\*</sup>, Frans van Winden<sup>2,3</sup>, Benjamin Pelloux<sup>2,3,4</sup>, Mirre Stallen<sup>5,6</sup> and K. Richard Ridderinkhof<sup>2,7</sup>

<sup>1</sup> Brain and Cognition, Department of Psychology, University of Amsterdam, Amsterdam, Netherlands

<sup>2</sup> Cognitive Science Center Amsterdam, University of Amsterdam, Amsterdam, Netherlands

<sup>3</sup> Amsterdam School of Economics, Center for Research in Experimental Economics and Political Decision-Making, University of Amsterdam, Amsterdam, Netherlands

<sup>4</sup> Groupe d'Analyse et de Théorie Economique, CNRS and University of Lumiere Lyon 2, Lyon, France

<sup>5</sup> Rotterdam School of Management, Erasmus University, Rotterdam, Netherlands

<sup>6</sup> Donders Institute for Brain, Cognition and Behavior, Radboud University, Nijmegen, Netherlands

<sup>7</sup> Department of Psychology, Amsterdam Center for the Study of Adaptive Control in Brain and Behavior, University of Amsterdam, Amsterdam, Netherlands

## Edited by:

Itzhak Aharon, The Interdisciplinary Center, Israel

## Reviewed by:

Sabine Windmann, Goethe University Frankfurt, Germany

Xinying Cai, Washington University in St Louis, USA

## \*Correspondence:

Johannes J. Fahrenfort, Brain and Cognition, Department of Psychology, University of Amsterdam, Room 3.19, Weesperplein 4, 1018XA Amsterdam, Netherlands.

e-mail: fahrendort@yahoo.com

There is a growing interest for the determinants of human choice behavior in social settings. Upon initial contact, investment choices in social settings can be inherently risky, as the degree to which the other person will reciprocate is unknown. Nevertheless, people have been shown to exhibit prosocial behavior even in one-shot laboratory settings where all interaction has been taken away. A logical step has been to link such behavior to trait empathy-related neurobiological networks. However, as a social interaction unfolds, the degree of uncertainty with respect to the expected payoff of choice behavior may change as a function of the interaction. Here we attempt to capture this factor. We show that the interpersonal tie one develops with another person during interaction – rather than trait empathy – motivates investment in a public good that is shared with an anonymous interaction partner. We examined how individual differences in trait empathy and interpersonal ties modulate neural responses to imposed monetary sharing. After, but not before interaction in a public good game, sharing prompted activation of neural systems associated with reward (striatum), empathy (anterior insular cortex and anterior cingulate cortex) as well as altruism, and social significance [posterior superior temporal sulcus (pSTS)]. Although these activations could be linked to both empathy and interpersonal ties, only tie-related pSTS activation predicted prosocial behavior during subsequent interaction, suggesting a neural substrate for keeping track of social relevance.

**Keywords:** interpersonal ties, social ties, empathy, social decision-making, public good game, pSTS, ACC, insula

## INTRODUCTION

In experience-based decision-making, rather than being provided with a description of the probability distribution of outcomes associated with certain choices, subjects have incomplete information about choice outcomes and their respective probabilities. Information regarding choice–outcome is acquired and integrated through continuous or pre-choice sampling with feedback. Evidence seems to suggest that people have different choice–preferences in decisions-from-experience paradigms than during description-based choice paradigms, although the origins of this putative preference gap are still under investigation (Rakow and Newell, 2010). Given the naturalistic nature of experience-based decision-making, disentangling its underlying cognitive and neural mechanisms may be crucial to understanding human choice behavior. Economic trade for example, is an omnipresent real-world example of experience-based decision-making. The probability distribution of outcomes during economic interactions can often only be gauged through the interaction itself. Interestingly, intermediate outcomes of the interaction may impact the probability distribution of future outcomes, lessening the degree

of uncertainty along the way. Hence, not only is the subject gauging the probability distribution of outcomes, he or she may actually modulate the probability distribution of outcomes as the interaction unfolds. It has been suggested that interaction partners keep track of this dynamic process internally by means of the interpersonal tie one develops with one's interaction partner (van Dijk et al., 2002; van Winden et al., 2008). Another suggested influence on human choice behavior in social settings has been the presence of empathy (e.g., Hein et al., 2010). Here, we attempt to establish how well a trait variable like empathy and a dynamic state variable such as the interpersonal tie are able to predict prosocial choices during interactive decision-making.

The willingness to share resources without guarantee of return-benefit is crucial to achieve cooperation. This is illustrated by the Public Good family of economic games, where free riding typically leads to a breakdown of cooperation (Isaac and Walker, 1988). Experimental evidence suggests that the willingness to invest in a public good (including the voluntary punishment of free riders) is motivated by social factors beyond immediate personal gain (Fehr and Gächter, 2002; Rilling et al., 2002). These factors

are often captured using the term altruism, which is observed in humans as well as various other mammals (Fehr and Fischbacher, 2003; de Waal, 2008). Many have linked the proximate cause of altruism to empathy, referring to the capacity to experience and understand the affective state of another person (Batson, 1991; Eisenberg, 2000; de Waal, 2008; Singer and Lamm, 2009). Recently, neural evidence has been presented in support of this view, linking empathy-related activation in the insula (Hein et al., 2010; Masten et al., 2011) and medial prefrontal cortex (Mathur et al., 2010; Masten et al., 2011) to prosocial behavior. Empathy promotes social cohesion (Farrow et al., 2001), scales with feelings of group membership (Sturmer et al., 2006) and as such may be a key motive for the willingness to share.

Individual differences in the willingness to share are often described in terms of social value orientations (Van Lange, 1999), which can be estimated by having subjects make choices between different monetary distributions regarding themselves and an unspecified other (Liebrand, 1984; Haruno and Frith, 2010). Social value orientation, like empathy, is considered to be a trait-like, stable disposition (Van Lange, 1999). However, recent evidence suggests that social value orientation toward an interaction partner can change, depending on the success or failure of an intermediate interaction (van Dijk et al., 2002). A complementary influence on prosocial behavior might therefore be the development of interpersonal ties, or feelings of social connectedness with an interaction partner. Such ties are different from empathy in that they consist of the extent to which someone cares about another person (van Dijk et al., 2002), rather than one's ability to share someone's emotions (Eisenberg, 2000; de Waal, 2008; Singer and Lamm, 2009). The degree to which one is able to experience empathy seems to be a personality characteristic (Mehrabian, 1997; Singer et al., 2004b), while interpersonal ties are dynamic, resulting from interaction itself (van Dijk et al., 2002; van Winden et al., 2008). Positively developing ties may instill prosocial behavior, while negatively developing ties may be associated with antisocial behavior, such as a reduced willingness to invest further in the other's well-being, or even vengeful behavior, such as punishing the other at a cost to oneself (Fehr and Gächter, 2002).

Compared to empathy, the cognitive, affective, and neural mechanisms of interpersonal ties have remained relatively exempt of empirical scrutiny. Various studies have focused on normative aspects of social interactions, such as trust (King-Casas et al., 2005) and fairness (Singer et al., 2006). Other studies have focused on the neural basis of cooperation itself. For example, left orbitofrontal cortex (OFC) and medial parts of the frontal cortex have been associated with cooperative and competitive behavior respectively (Decety et al., 2004). Cooperative behavior in an iterated prisoner's dilemma game has also been found to elicit increased activity in reward-related areas such as the striatum (Moll et al., 2006; Haber and Knutson, 2010), as well as the ventromedial prefrontal cortex (VMPFC) and the OFC (Rilling et al., 2002). Increased (decreased) VMPFC activation has further been found when inferring generous (selfish) play in a public good game played by others (Cooper et al., 2010). Although these regions may be involved in the formation of interpersonal ties, it is unknown whether they are also involved in keeping track of such ties over longer periods of time. Studies on social cognition point to the relevance of the posterior

superior temporal sulcus (pSTS) in representing the social significance of other agents. For example, gray matter density abnormalities in the STS have been associated with autism in children (Boddaert et al., 2004), stimuli that have acquired relevance during meaningful social interaction (e.g., faces of cooperative game partners) have been shown to activate the pSTS (Singer et al., 2004a); the (p)STS is important in keeping track of one's influence on other agent's intentions and strategies (Hampton et al., 2008; Haruno and Kawato, 2009); and activation of the pSTS during the perception of agency correlates with individuals' self-reported altruism (Tankersley et al., 2007).

To determine the influence of empathy and interpersonal ties in determining prosocial behavior, we first established the presence of neural responses related to empathy and interpersonal ties during computer controlled (imposed) sharing. In a second step, we determined how these neural responses related to previous cooperation success and whether they predicted subsequent prosocial behavior. To establish neural empathy and tie-related responses during computer-imposed sharing, we administered a distributional outcome test (DOT; Liebrand, 1984; Haruno and Frith, 2010), with monetary outcomes benefiting the other at the expense of oneself and vice versa (see **Table 1**), both before and after anonymous interaction in a public good game. In the DOT, subjects were repeatedly confronted with changing payoff distributions, of which the computer would choose one on each trial. We used functional magnetic resonance imaging (fMRI) to examine individual differences in the neural responses to DOT outcomes. Subtracting fMRI responses during the pre-interaction DOT from the responses during the post-interaction DOT, allowed us to isolate neural responses to sharing that could be related to the intermediate public good game interaction (see Materials and Methods). Choices between distributional outcomes were made by the computer, kept equal across participants, and kept equal between pre- and post-interaction DOTs, to ensure that neural responses could be compared across time and across participants.

## MATERIALS AND METHODS

### SUBJECTS

We collected data from 34 interaction pairs. Due to our complex experimental setup, five subjects had to be excluded from analysis because of computer or human error during data collection and transfer. Subjects were students from the University of Amsterdam or international exchange students enrolled in courses at the University of Amsterdam. All subjects gave written informed consent prior to the onset of the experiment. The 29 subjects in the scanner had an average age of 22.6 (SD  $\pm$  2.7), 14 were female. Their 29 interaction partners had an average age of 23.3 (SD  $\pm$  4.5), 11 were female. In 14 interaction pairs both partners had the same sex. In five pairs both partners were female. Chi-square and *t*-tests confirmed that the scanning group did not differ significantly from the group outside the scanner in age, sex, or any of the behavioral measures.

### EXPERIMENTAL PROCEDURE

Subjects were tested in pairs. One subject was positioned in the scanner while the other was seated in a separate room behind a laptop. The procedure was fully anonymous, such that subjects would

**Table 1 | Alternatives during the scanned pre- and post-interaction distributional outcome test (DOT).**

Trial	Alternative A		Alternative B		Other gains at expense of self	Self gains at expense of other	Equality in one of the alternatives
	Self	Other	Self	Other			
1	500	0	498	44		x	
2	498	44	492	87		x	
3	492	87	483	129		x	
4	483	129	470	171		x	
5	470	171	354	354			x
6	354	354	171	470			x
7	498	44	500	0	x		
8	492	87	498	44	x		
9	483	129	492	87	x		
10	470	171	483	129	x		
11	354	354	470	171			x
12	171	470	354	354			x
13	-500	0	-498	-44	x		
14	-498	-44	-492	-87	x		
15	-492	-87	-483	-129	x		
16	-483	-129	-470	-171	x		
17	-470	-171	-354	-354			x
18	-354	-354	-171	-470			x
19	-498	-44	-500	0		x	
20	-492	-87	-498	-44		x	
21	-483	-129	-492	-87		x	
22	-470	-171	-483	-129		x	
23	-354	-354	-470	-171			x
24	-171	-470	-354	-354			x
Sum	0	0	0	0			

The temporal order in which the trials were presented was randomized for each subject. Whether an alternative would appear on the screen as (Alternative A) or (Alternative B) was randomized on every trial. The computer would always choose what is designated here as (Alternative A).

never see each other or get any information about the other subject other than through computer-interfaced interaction during the public good game. The study was divided into two main parts: (1) a scanning phase and (2) a post-scanning part in which subjects filled out exit questions and an empathy questionnaire. Afterward, subjects were paid out. Average earnings totaled to around 45 euros per subject.

The scanning phase consisted of three scanning blocks. In the first block, we administered a pre-interaction DOT to test neural responses to different monetary distributions. In the second block, the subject pair played a two-person public good game. In the third block subjects were administered a post-interaction DOT. Right after the post-interaction DOT, subjects played the final rounds of the public good game. After scanning, subjects completed a general exit questionnaire and the balanced emotional empathy scale (BEES; Mehrabian and Epstein, 1972; Mehrabian, 1997).

#### DISTRIBUTIONAL OUTCOME TEST

Before and after public good game interaction, the DOT was administered, containing 24 test items (see **Table 1**). These test items were chosen because behavioral pilot testing showed that they discriminated most strongly between subjects when given

the ability to freely choose among them. In the DOT however, subjects would not make a choice between the two payoff combinations themselves, but the computer would “randomly” pick an alternative. Unbeknownst to subjects, computer choices were predetermined such that the chosen alternatives and the resulting payouts were the same between participants and between the pre- and post-interaction DOT. Also unbeknownst to subjects, both the sum of the chosen distributions and the sum of the alternative distributions was zero, both for allocations to self and allocations to other.

At the start of each trial, subjects were asked to carefully inspect the payoff combinations presented by the computer, and to assess which alternative they preferred. Upon pressing a button, the alternatives would start to alternately highlight at a decreasing rate, until settling down on one of them (the “computer choice”). No changes were made to the display during the ensuing 8 s, so that a clean measure of a subject’s neural response to the “computer choice” could be taken. To keep subjects engaged in this otherwise passive task, they were asked to subsequently indicate whether they agreed or disagreed with the computer choice using a simple yes/no button press, after which the test would continue to the next trial.

Importantly, the pre- and post-interaction DOT were identical in all respects, except for one difference. In the pre-interaction DOT, subjects were informed that the “other” was a random subject from the experiment, while in the post-interaction DOT, subjects were informed that the “other” was one’s (anonymous) interaction partner during the preceding public good game. Hence differences between neural responses to the post- and pre-interaction DOT acted as a post–pre-interaction contrast, isolating differential neural responses to sharing as a result of the intermediate public good game interaction. Subtracting the pre-interaction DOT from the post-interaction DOT removes effects that cannot be related to the intermediate public good game interaction.

Therefore, we subtracted the normalized pre-interaction DOT fMRI signal from the post-interaction DOT signal, to isolate the component of the DOT response that may have been impacted by the intermediate public good game interaction (see below for analysis details). This is a non-standard fMRI analysis step, but pertinent to the type of information we wanted to extract from our design. We were interested in isolating changes in neural components of sharing preference as a result of experience-induced change over time (i.e., as a result of public good game interaction). The difference signal of neural responses acquired at different moments in time – while keeping all other variables the same – provided this measure.

**PUBLIC GOOD GAME**

The intermediate public good game was a finite, non-linear, self-paced public good game described elsewhere (van Dijk et al., 2002). The game consisted of 29 rounds between two subjects. On every round, subjects were given 12 monetary units, which they could divide freely between their private account and a public account. The payoff consequences of contributions to the public account were made explicit by use of an

on-screen payoff table (see **Table 2**). Payoffs in this table were given by  $14*(X + Y) + 32*(12 - X) - (12 - X)^2 - 160$ , where  $X$  represents a subject’s own contribution to the public account, while  $Y$  denotes the other subject’s contribution to the public account. The Pareto-optimum was at 10 tokens in the public good, but players could decide to invest more than 10 tokens, increasing the other’s payoff at one’s own expense. Likewise, participants could invest less than the Nash-equilibrium (three tokens), resulting in a payoff below Nash for the other at one’s own expense.

Prior to the public good game, extensive on-screen instructions were given, followed by nine multiple-choice questions to check understanding. At round 25, the public good game was unexpectedly interrupted for the post-interaction DOT. Right after the post-interaction DOT subjects finished the last four rounds of the public good game. At the end of the experiment, payoffs were converted to real currency and paid out along with the other payouts.

**POST-SCAN QUESTIONNAIRES FOR EMPATHY AND INTERPERSONAL TIES**

After scanning, subjects completed an exit questionnaire and an empathy scale questionnaire. Interpersonal ties were measured during the exit questionnaire. Subjects were asked to indicate the interpersonal tie they felt with the subject they had interacted with during the public good game. The question was:

*Please indicate on the scale below how much you like or dislike the participant you were paired with in part 2 and 3 of the experiment. Circle the number of your choice, where 1 stands for “very unpleasant person” and 9 stands for “very nice person.”*

Below was a 9-point scale running from “very unpleasant person” to “very nice person.” Note that a value below five signals a negative tie, whereas a value above five signals a positive tie. Ratings emanated from previous cooperation success, as there

**Table 2 | Payoff table during the public good game.**

		Other													
		Y	0	1	2	3	4	5	6	7	8	9	10	11	12
You	Public	Private	12	11	10	9	8	7	6	5	4	3	2	1	0
	0	12	80	94	108	122	136	150	164	178	192	206	220	234	248
	1	11	85	99	113	127	141	155	169	183	197	211	225	239	253
	2	10	88	102	116	130	144	158	172	186	200	214	228	242	256
	3	9	89	103	117	131	145	159	173	187	201	215	229	243	257
	4	8	88	102	116	130	144	158	172	186	200	214	228	242	256
	5	7	85	99	113	127	141	155	169	183	197	211	225	239	253
	6	6	80	94	108	122	136	150	164	178	192	206	220	234	248
	7	5	73	87	101	115	129	143	157	171	185	199	213	227	241
	8	4	64	78	92	106	120	134	148	162	176	190	204	218	232
	9	3	53	67	81	95	109	123	137	151	165	179	193	207	221
	10	2	40	54	68	82	96	110	124	138	152	166	180	194	208
	11	1	25	39	53	67	81	95	109	123	137	151	165	179	193
12	0	8	22	36	50	64	78	92	106	120	134	148	162	176	

*X denotes a participant’s own contribution to the public account, Y denotes the other’s contribution to the public account. Private is what is left in a participant’s private account given that contribution level. The cell where X and Y cross shows one’s own payoff given those contribution levels.*

were no encounters between partners other than during anonymous interaction in the public good game. In addition, the exit questionnaire contained some questions on interaction strategy and control questions to check whether there had indeed been no contact with the other subject other than through computer-mediated interaction. To assess trait empathy, subjects completed the BEES (Mehrabian and Epstein, 1972; Mehrabian, 1997) after they had completed the exit questionnaire.

### PAST COOPERATION SUCCESS AND FUTURE INVESTMENT BEHAVIOR

As a measure of previous cooperation success, we calculated the other's average contribution to the public good in the last five rounds prior to the second DOT (round 21–25; van Dijk et al., 2002). Future investment behavior was defined as a player's own average contribution to the public good in the rounds following the second DOT (round 26–28). The final round (round 29) was not included in the average because of the well-known “end” effect (Ledyard, 1995). A *t*-test confirmed that contribution in round 29 differed significantly from contribution in round 28 [ $t(28) = 2.22$ ,  $p = 0.03$ ], while this was not the case between round 28 and 27 [ $t(28) = 0.81$ ,  $p = 0.42$ ] or between round 27 and 26 [ $t(28) = 0.85$ ,  $p = 0.40$ ]. There was also no change between round 25 and round 26, either in one's own contribution [ $t(28) = 0.57$ ,  $p = 0.57$ ] or in the other's contribution to the public good [ $t(28) = 1.22$ ,  $p = 0.23$ ], showing that the interruption by the second DOT did not significantly impact subsequent public good game behavior. Control analyses confirmed that changing the size of the interval over which interaction success or future investment behavior was measured did not change the results.

### IMAGE ACQUISITION AND PRE-PROCESSING

Images were acquired on a Philips 3T Intera scanner. The functional recordings were acquired using a T2\*-weighted sequence [40 coronal slices; flip angle (FA), 80°; echo time (TE), 30 ms; repetition time (TR), 2.3 s; slice thickness, 3 mm; field of view (FOV), 220 mm × 220 mm; in-plane voxel resolution, 2.3 mm × 2.3 mm]. Sessions ended with the acquisition of a high-resolution anatomical image using a T1 turbo field echo sequence [182 coronal slices; FA, 8°; TE, 4.6 ms; TR, 9.6 s; slice thickness, 1.2 mm, FOV, 256 mm × 256 mm; in-plane voxel resolution, 1 mm × 1 mm]. Pre-processing and data analysis was performed using the fMRI expert analysis tool (FEAT), v5.98 from the FMRIB's software library (FSL) package (<http://www.fmrib.ox.ac.uk/fsl>). fMRI images were motion corrected, slice-time aligned, aligned to the structural image of the subject, and spatially smoothed using a Gaussian kernel of 5 mm and high-pass temporally filtered using a Gaussian envelope of 50 s. Anatomical brains were extracted from the structural images, and transformed to the standard space of the Montreal Neurological Institute (MNI) using FMRIB's non-linear image registration tool (FNIRT). Finally, the functional data were co-registered to the MNI brain using non-linear parameters obtained from FNIRT.

### DATA ANALYSIS

For each subject, general linear models (GLMs) were specified separately for the pre- and post-interaction DOT. A predictor was created for each trial, corresponding to the moment at which the

“computer choice” (see **Table 1**) was presented to the subject. All predictors were convolved with a standard double gamma hemodynamic response function. The resulting GLMs were applied to the (pre-processed) fMRI signals that were acquired during the pre- and post-interaction DOT.

Next, the mean percent signal change on each trial of the pre- and post-interaction DOT was extracted for each subject and each region of interest (ROI). Trials were averaged depending on the sharing context (see **Table 1**): for each subject and each ROI, an average was created for trials in which the other obtained monetary gains at the expense of oneself (“Other gains at expense of Self”), and another average was created for trials in which oneself gained at the expense of the other (“Self gains at expense of Other”), separately for the pre- and post-interaction DOT. Subsequently, the resulting pre-interaction DOT averages were subtracted from the post-interaction DOT averages of corresponding sharing contexts, isolating the part of the neural response that was affected by the intermediate public good game experience, while subtracting out the bare neural responses to the monetary divisions in these sharing contexts (which were equal in both DOTs).

The post- minus pre-public good game interaction DOT difference signals from these sharing contexts (“Self gains at expense of Other” and “Other gains at expense of Self”) for each of the ROIs were correlated against the empathy and interpersonal tie measures, as well as against the measures for past cooperation success and future investment behavior. Because we computed a lot of correlations, one needs to correct for multiple comparisons, as some correlations will turn up significant at the conventional 0.05 level by chance alone. We corrected for multiple comparisons by means of the false discovery rate (FDR) correction, which fixes the expected proportion of incorrectly rejected null hypotheses (type I errors) given the number of tests that were performed (Benjamini and Hochberg, 1995). All reported statistics in the experiment are two-tailed.

To summarize, the subtraction method and the FDR correction each correct for a potential source of spurious correlation: (1) The subtraction of the pre-interaction DOT from the post-interaction DOT isolates the part of the signal that can be related to the intermediate public good game. The subtraction method removes correlations between the DOT and our dependent measures that are not related to the intermediate public good game. Such correlations may be coincidental but “real” within the specific sample that we measured from, or they may even be real in the population at large. Regardless, the subtraction method removes them. (2) The FDR correction thresholds the resulting post- minus pre-interaction correlations such that they cannot be attributed to measurement errors (noise). The correlations that are identified using the combination of these two methods can therefore be attributed to the intermediate public good game, and cannot be attributed to chance.

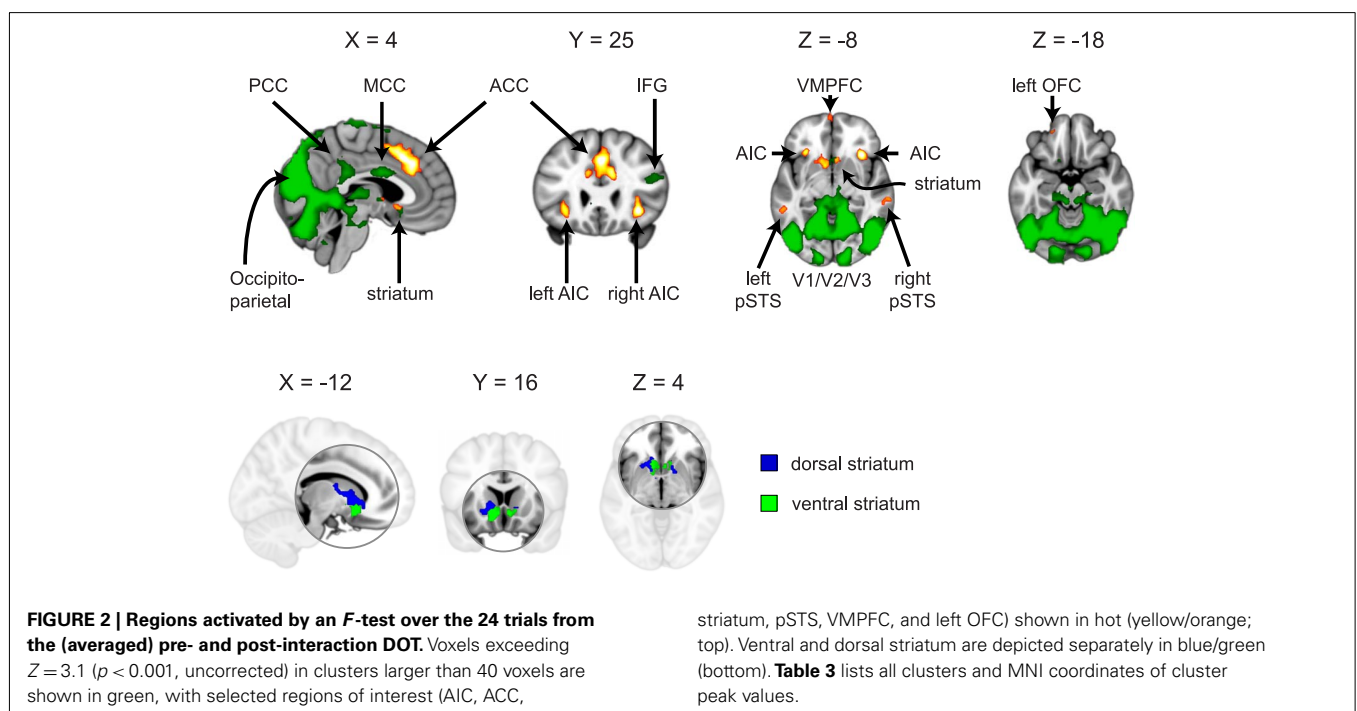
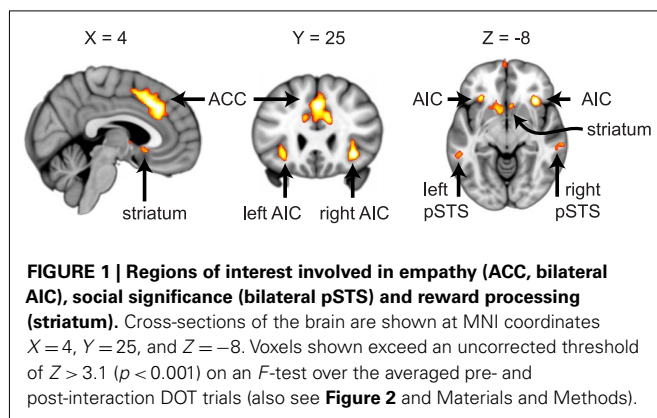
To establish whether empathy and interpersonal ties might interact to predict behavior, we carried out a regression analysis. We used the BEES and liking ratings as measures for empathy and social ties respectively. An interaction term of the two was added. All terms were regressed onto the subject's future contributions in the three rounds following the second DOT.

## SELECTION OF REGIONS OF INTEREST

For ROI selection, the results from the lower level DOT analysis were combined using a fixed effects higher-level analysis, in which corresponding trials from the pre- and post-interaction DOT were matched. Note that in this particular analysis, trials from the pre- and post-interaction DOT were combined (effectively averaged), rather than subtracted. Subsequently, the resulting individual subject statistics were analyzed at a group level using FMRIB's Local Analysis of Mixed Effects (FLAME1). At this stage, an  $F$ -test was performed over all 24 trials, allowing us to test in a non-directional way in what regions the combined pre- and post-interaction DOT contributed significantly to the fMRI signal at a group level. By using a non-directional  $F$ -test, the statistics were indifferent to the direction of hemodynamic changes between different types of trials within the DOT. Clusters with  $Z$ -values higher than 3.1 (uncorrected) were considered for inclusion as ROIs into the final analysis. From these clusters,

ROIs were selected on anatomical grounds that are known to be involved in reward processing (striatum; Rilling et al., 2002; Moll et al., 2006; Haber and Knutson, 2010), empathy [anterior insular cortex (AIC) and anterior cingulate cortex (ACC)] (Wicker et al., 2003; Singer et al., 2004b; Singer and Lamm, 2009), social significance (pSTS; Singer et al., 2004a; Tankersley et al., 2007; Hampton et al., 2008), and social decision-making (VMPFC and OFC; Rilling et al., 2002; Decety et al., 2004). As striatal activity was part of a larger cluster, we extracted voxels from this cluster using the Harvard–Oxford Subcortical Structural Atlas included in the FSL package. Only voxels that had more than 50% probability of belonging to the nucleus accumbens (ventral striatum) or caudate (dorsal striatum) were included. No voxels exceeded the threshold in the putamen. See **Figure 1** for selected clusters. For the entire activation map of the  $F$ -test and the selection of ventral and dorsal striatum see **Figure 2**, for a list of all clusters see **Table 3**.

Because the DOT data used for ROI selection were also used during hypothesis testing, a possible concern could be selection bias resulting from the non-independence error or “double dipping” (Kriegeskorte et al., 2009). However, the statistical tests used during hypothesis testing were not affected by circularity. For one, the result statistics used during ROI selection were different from those used during hypothesis testing. ROI selection was done on the averaged pre- and post-interaction DOT, while during hypothesis testing the pre- interaction DOT was subtracted from the post-interaction DOT. More to the point however, hypothesis testing was done by correlating this difference signal against measures of empathy, interpersonal ties, and interaction success, while none of these behavioral measures were used when generating the  $F$ -statistics used for ROI selection. Critically, this ensured that none of our results could be explained by double dipping.





**Table 3 | Locus of peak activations in clusters resulting from an *F*-test over the 24 trials from the (averaged) pre- and post-interaction DOT ( $Z > 3.1$ , cluster size  $> 40$ ).**

Cluster index	Brain region	MNI coordinates of peak activation (mm)			Z score
		x	y	z	
1	<u>Ventromedial prefrontal cortex (VMPFC)</u>	-2	64	-10	4
2	<u>Left posterior superior temporal sulcus</u>	-56	-40	-6	4.43
3	<u>Left orbitofrontal cortex (OFC)</u>	-22	48	-26	5.32
4	<u>Right posterior superior temporal sulcus</u>	62	-26	-10	4.73
5	Near right lateral ventricle	26	-44	16	6.11
6	<u>Left anterior insular cortex (AIC)</u>	-30	26	-6	6.31
7	Left precuneus	-8	-52	28	4.42
8	Left central opercular cortex (middle part of the Sylvian fissure)	-66	-14	8	5
9	Left angular gyrus	-50	-72	30	4.26
10	Left parietal opercular cortex (posterior part of the Sylvian fissure)	-46	-36	16	6.37
11	Left postcentral gyrus	-38	-22	50	7.91
12	<u>Right anterior insular cortex (AIC)</u>	34	24	-6	8.21
13	Left superior temporal gyrus (posterior)	64	-30	14	7.05
14	Mid cingulate cortex (MCC)	4	0	28	6.62
15	Posterior cingulate cortex (PCC)	2	-34	26	6.88
16	Left early visual cortex (V1, V2, V3)	-16	-88	-20	8.21
17	Right early visual cortex (V1, V2, V3)	20	-90	-18	8.21
18	Left Precentral Gyrus	-44	-2	26	8.21
19	<u>Anterior cingulate cortex (ACC)</u>	2	24	38	8.21
20	Occipitoparietal cortex, extending into precentral gyrus and inferior frontal gyrus on the right, and into cerebellum, brain stem, thalamus, and <u>striatum</u> ventrally	0	-78	20	8.21

Regions of interest used in subsequent analyses are underlined.

## RESULTS

### ORTHOGONALITY OF TRAIT EMPATHY AND INTERPERSONAL TIES

Trait empathy was measured post-experimentally using the BEES (Mehrabian and Epstein, 1972; Mehrabian, 1997). Interpersonal ties were measured in an exit questionnaire in which subjects were asked to rate how much they liked or disliked their interaction partner. To assess whether these measures were somehow correlated, we calculated their correlation coefficient. This showed that our empathy and our interpersonal ties measure were unrelated ( $r = -0.06$ ,  $p = 0.76$ ).

### NEURAL CORRELATES OF EMPATHY DURING IMPOSED SHARING

In order to assess the presence of neural correlates of trait empathy during imposed sharing, we correlated BEES scores with the post- minus pre-interaction neural DOT responses during the “Self gains at expense of Other” and “Other gains at expense of Self” sharing contexts, in all regions of interest (see **Table 4**). The FDR-corrected significance level at which these correlations were evaluated is 0.0043, as marked by three asterisks (see bottom of **Table 4** on the next page). After correction for multiple comparisons, the BEES scores showed correlations in AIC [ $r = 0.61$ ,  $p = 0.0004$ ] (**Figure 3**, top), ACC ( $r = 0.64$ ,  $p = 0.0002$ ; **Figure 3**, middle), and the striatum [ $r = 0.53$ ,  $p = 0.0043$ ] (**Figure 3**, bottom). This shows that responses in these areas were stronger for empathic subjects than for non-empathic subjects after, but

not before interaction (**Figure 4**). Notably, empathy only correlated with neural responses to sharing when the other gained at one’s own expense, and not the other way around (all  $p > 0.3$ ).

### NEURAL CORRELATES OF INTERPERSONAL TIES DURING IMPOSED SHARING

Similar to the empathy analyses, we correlated liking ratings with post- minus pre-interaction neural DOT responses in both sharing contexts for all regions of interest (see **Table 4**). After correction for multiple comparisons, only the pSTS correlated significantly with our interpersonal tie measure, and only when oneself gains at the expense of the other ( $r = 0.58$ ,  $p = 0.0011$ ), and not vice versa ( $r = 0.32$ ,  $p = 0.0928$ ; **Figure 5**, top). This suggests that gaining at another person’s expense invokes activation in the pSTS, where the degree of activation correlates positively with the interpersonal tie one has developed with that person during previous interaction. Receiving reward at another’s expense may invoke a tie-related response in the pSTS, where a strong response corresponds with a positive interpersonal tie with the other person, and a weak response corresponds with a negative tie. Again, the correlation is specific to the post- minus pre-interaction DOT, emphasizing the relevance of the intermediate interaction in driving the response (**Figure 6**).

**Table 4 | Correlations for sharing with measures of empathy, social ties, past cooperation success, and future investment behavior.**

Region of interest	Self gains at the expense of other		Other gains at the expense of self	
	Pearson	p-Value	Pearson	p-Value
<b>CORRELATIONS WITH BALANCED EMOTIONAL EMPATHY SCALE (BEES)</b>				
Anterior insular cortex (AIC)	0.03	0.8582	0.61	0.0004***
Left AIC	-0.06	0.7632	0.63	0.0003***
Right AIC	0.16	0.4079	0.46	0.0127*
Striatum	-0.18	0.3507	0.52	0.0043***
Dorsal striatum (caudate)	-0.10	0.5940	0.55	0.0020***
Left dorsal striatum	-0.17	0.3831	0.43	0.0203*
Right dorsal striatum	-0.01	0.9485	0.59	0.0008***
Ventral striatum (nucleus accumbens)	-0.23	0.2304	0.39	0.0344*
Left ventral striatum	0.00	0.9914	0.16	0.3941
Right ventral striatum	-0.37	0.0495*	0.48	0.0089**
Posterior superior temporal sulcus (pSTS)	-0.20	0.2999	0.28	0.1396
Left pSTS	-0.21	0.2639	0.27	0.1645
Right pSTS	-0.12	0.5230	0.19	0.3299
Anterior cingulate cortex (ACC)	0.14	0.4607	0.64	0.0002***
Left orbitofrontal cortex (OFC)	-0.14	0.4680	0.18	0.3472
Ventromedial prefrontal cortex (VMPFC)	-0.04	0.8170	0.38	0.0447*
<b>CORRELATIONS WITH 9-POINT LIKE SCALE (INTERPERSONAL TIE)</b>				
Anterior insular cortex (AIC)	0.19	0.3270	-0.12	0.5420
Left AIC	0.18	0.3566	-0.08	0.6744
Right AIC	0.14	0.4591	-0.14	0.4571
Striatum	0.09	0.6362	0.11	0.5562
Dorsal striatum (caudate)	0.21	0.2718	0.11	0.5875
Left dorsal striatum	0.21	0.2695	0.21	0.2786
Right dorsal striatum	0.18	0.3585	-0.01	0.9613
Ventral striatum (nucleus accumbens)	-0.07	0.7293	0.10	0.5887
Left ventral striatum	-0.15	0.4435	0.23	0.2341
Right ventral striatum	0.01	0.9594	-0.06	0.7757
Posterior superior temporal sulcus (pSTS)	0.58	0.0011***	0.32	0.0928
Left pSTS	0.40	0.0312*	0.01	0.9485
Right pSTS	0.57	0.0013***	0.54	0.0027***
Anterior cingulate cortex (ACC)	0.26	0.1684	-0.08	0.6835
Left orbitofrontal cortex (OFC)	-0.45	0.0149*	0.27	0.1630
Ventromedial prefrontal cortex (VMPFC)	-0.12	0.5337	-0.06	0.7670
<b>CORRELATIONS WITH CONTRIBUTION OF OTHER IN ROUND 21-25 (PREVIOUS COOPERATION SUCCESS)</b>				
Anterior insular cortex (AIC)	-0.13	0.5160	-0.20	0.3105
Left AIC	-0.04	0.8382	-0.24	0.2062
Right AIC	-0.21	0.2732	-0.09	0.6486
Striatum	0.29	0.1209	0.01	0.9705
Dorsal striatum (caudate)	0.34	0.0684	0.02	0.9200
Left dorsal striatum	0.43	0.0201*	0.11	0.5619
Right dorsal striatum	0.19	0.3227	-0.07	0.7080
Ventral striatum (nucleus accumbens)	0.17	0.3664	-0.01	0.9707
Left ventral striatum	0.11	0.5854	0.01	0.9729
Right ventral striatum	0.19	0.3132	-0.02	0.9257
Posterior superior temporal sulcus (pSTS)	0.63	0.0003***	0.38	0.0413*
Left pSTS	0.68	0.0000***	0.22	0.2610
Right pSTS	0.38	0.0436*	0.42	0.0245*
Anterior cingulate cortex (ACC)	0.08	0.6826	-0.13	0.5135

(Continued)



Table 4 | Continued

Region of interest	Self gains at the expense of other		Other gains at the expense of self	
	Pearson	<i>p</i> -Value	Pearson	<i>p</i> -Value
Left orbitofrontal cortex (OFC)	−0.19	0.3304	−0.01	0.9397
Ventromedial prefrontal cortex (VMPFC)	0.04	0.8315	0.11	0.5656
<b>CORRELATIONS WITH CONTRIBUTION OF SELF IN ROUND 26–28 (FUTURE INVESTMENT BEHAVIOR)</b>				
Anterior insular cortex (AIC)	0.01	0.9629	0.14	0.4710
Left AIC	0.08	0.6689	0.11	0.5857
Right AIC	−0.10	0.5981	0.16	0.4176
Striatum	0.29	0.1209	0.33	0.0809
Dorsal striatum (caudate)	0.33	0.0793	0.34	0.0754
Left dorsal striatum	0.38	0.0420*	0.43	0.0185*
Right dorsal striatum	0.22	0.2444	0.19	0.3207
Ventral striatum (nucleus accumbens)	0.19	0.3285	0.27	0.1574
Left ventral striatum	0.14	0.4651	0.22	0.2530
Right ventral striatum	0.19	0.3279	0.22	0.2499
Posterior superior temporal sulcus (pSTS)	0.58	0.0009***	0.47	0.0108*
Left pSTS	0.60	0.0006***	0.29	0.1293
Right pSTS	0.39	0.0371*	0.48	0.0081**
Anterior cingulate cortex (ACC)	0.16	0.4054	0.21	0.2844
Left orbitofrontal cortex (OFC)	−0.12	0.5450	0.05	0.7980
Ventromedial prefrontal cortex (VMPFC)	−0.13	0.4989	−0.12	0.5288

*N* = 29.

\**p* ≤ 0.05.

\*\**p* ≤ 0.01.

\*\*\**p* ≤ 0.0043, False discovery rate (FDR) corrected value of 0.05 over all correlations in the table.

#### RELATIONSHIP TO PAST COOPERATION SUCCESS AND FUTURE INVESTMENT BEHAVIOR

To determine the degree to which these sharing-related responses are driven by past cooperation success, and/or predict cooperation in the continuation of the public good game after the second DOT, we correlated all post- minus pre-DOT responses with past cooperation success and future behavior (Table 4, second part). Previous cooperation success was defined as the other's mean contribution to the public good in the five rounds prior to administering the post-interaction DOT (van Dijk et al., 2002), while future cooperation was defined as one's own mean contribution to the public good in the three rounds after the second DOT (see Materials and Methods). Past cooperation success (Figure 5, middle) as well as future cooperation (Figure 5, bottom) in the public good game correlated only with the pSTS, again particularly for outcomes that entail gains at the expense of the other (past success:  $r = 0.63$ ,  $p = 0.0003$ ; future cooperation:  $r = 0.58$ ,  $p = 0.0009$ ).

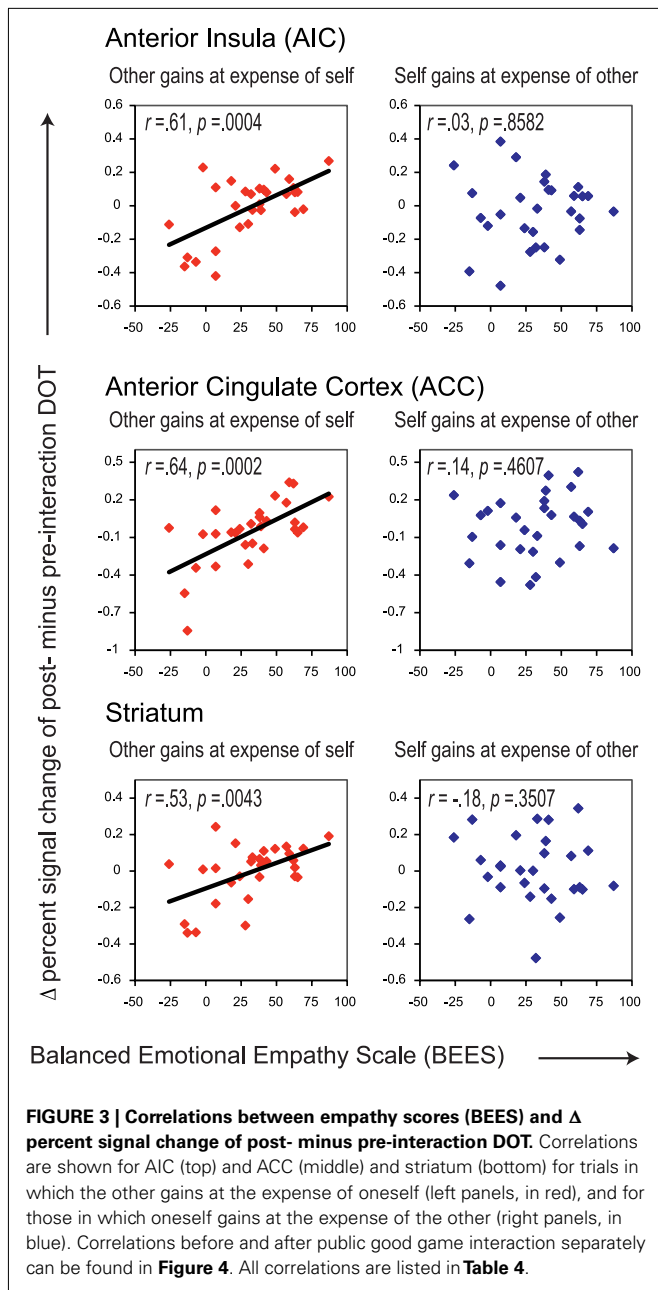
#### REGRESSION ANALYSIS OF EMPATHY-TIE INTERACTION

Finally, we wanted to determine whether our empathy and interpersonal ties measures – or their interaction – might be used to predict cooperative behavior directly. We carried out a regression analyses, in which we regressed our behavioral measures for empathy and interpersonal ties, plus an interaction term, onto the average contribution to the public good in the rounds following the second DOT (our behavioral measure for cooperative

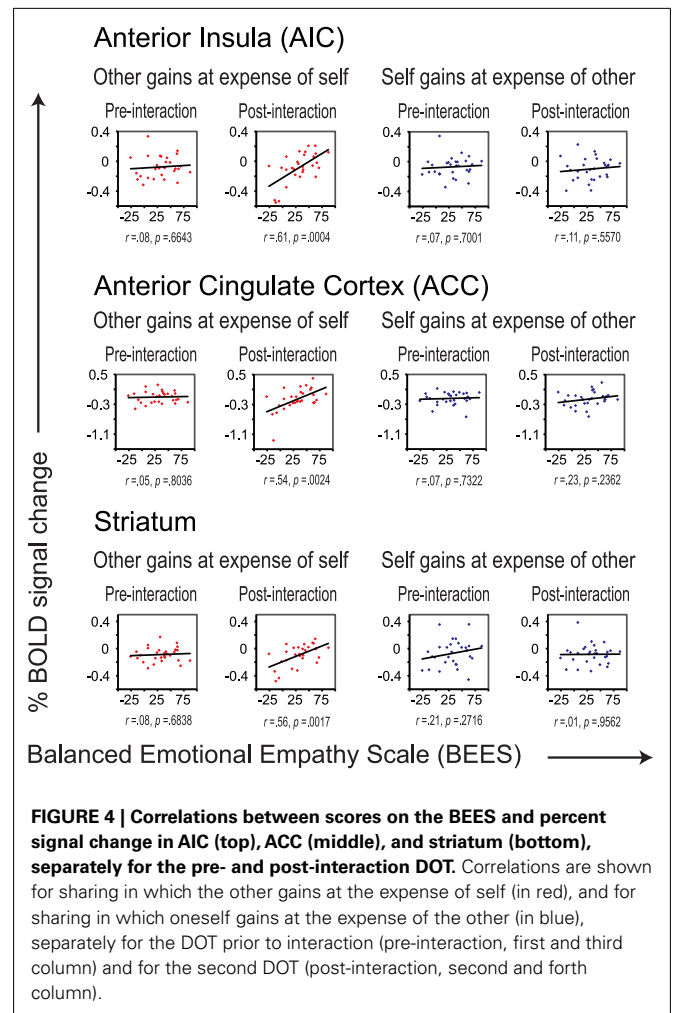
behavior in the public good game). The overall model was significant: [ $F(3,25) = 3.59$ ,  $p = 0.0277$ ]. Moreover, it was driven only by our interpersonal ties measure ( $p = 0.004$ ), and not by the BEES ( $p = 0.184$ ) or by the interaction of the two ( $p = 0.171$ ). This result confirms the influence of interpersonal ties – rather than empathy – in establishing prosocial behavior during interaction.

#### DISCUSSION

We isolated the influence of public good game interaction on neural responses in different sharing contexts, by subtracting normalized fMRI signals that were separated in time by the interaction. We investigated the post- minus pre-interaction difference signal in these sharing contexts (other gains at one's own expense and vice versa), subtracting out the raw neural responses to the division of monetary resources themselves. This somewhat unconventional analysis method allowed us to uniquely isolate the influence of intermediate public good game interaction on neural responses during imposed sharing, and determine the relation of these responses to empathy, interpersonal ties, past cooperation success, and future prosocial investment behavior. We found neural correlates of empathy and interpersonal tie measures in these interaction-related fMRI signals, showing that our neural correlates of these measures depend on the intermediate public good game interaction. Moreover, we show that the neural correlate of the interpersonal tie measure is uniquely related to past cooperation success and future prosocial investment behavior.

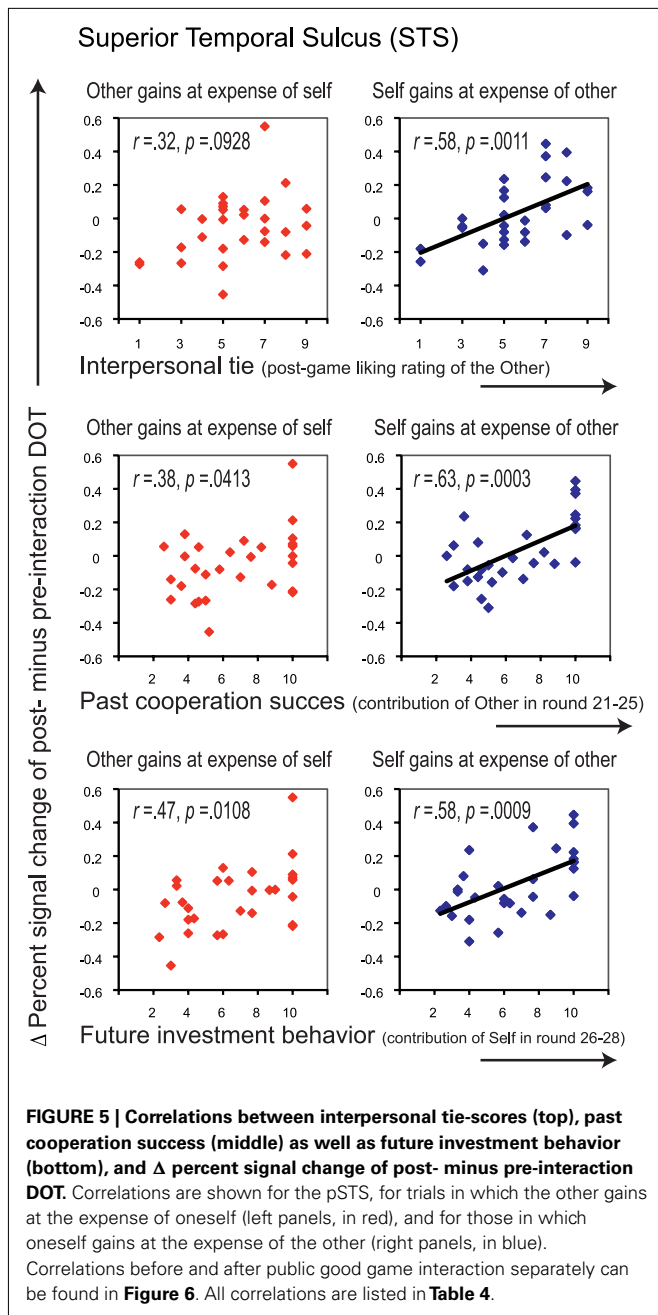


Although neural correlates of empathy only emerged after having interacted with the other participant (**Figure 4**), these correlates could not be related to previous cooperation success or to future choice behavior (**Table 4**). Thus, while these correlates only emerged as a result of interaction, they do not seem to result from specific choice outcomes, or result in specific choices. This suggests that the relationship between trait empathy and economic interaction is more generic. One possibility may be that trait empathy is only expressed when the other person turns out to be a “real” actor in the world. Bear in mind that during the pre-interaction DOT, subjects are told that the “other” is a random anonymous subject from the experiment and that there is no interaction with that participant. The combination of anonymity and a lack of interaction



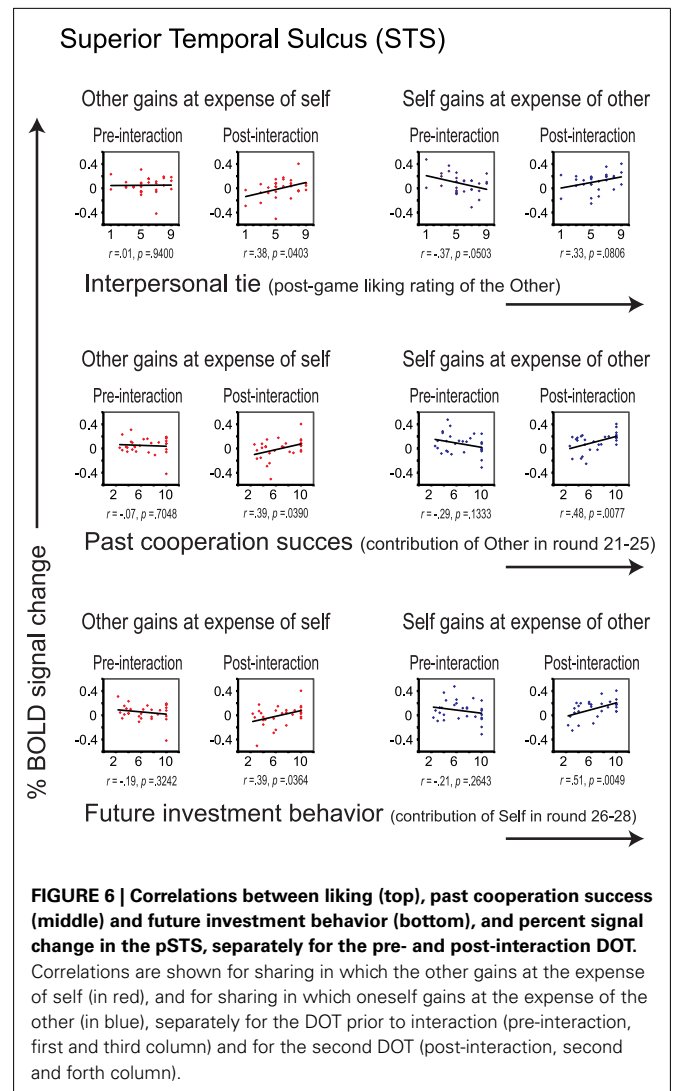
may lead to dehumanization, lessening empathic responses during the pre-interaction DOT (Cehajic et al., 2009).

The pSTS correlate of the interpersonal tie on the other hand, exhibits a more specific interaction-related pattern. Like empathy, it only emerges as a result of public good game interaction. However, this signal could also be linked to previous cooperation success and to future cooperative choice behavior in subsequent rounds of the public good game. Moreover, tie-related activity only correlated with pSTS activation in the context of receiving a monetary reward at the expense of the other. This seems to link pSTS activation to positive choice outcomes. Possibly, the degree to which self-favoring outcomes engender activation in the pSTS depends directly on past cooperation success, extending previous findings on the role of the pSTS in tracking the effects of one’s actions on other agents’ decisions (Hampton et al., 2008). Similarly, Haruno and Kawato (2009) have shown that the degree to which subjects are able to exploit other agent’s strategies is uniquely predicted by STS activity. We suggest that pSTS activation during forced sharing correlates with the interpersonal tie one has developed with another participant during previous interaction, even though this activation is captured outside and apart from the interaction. This alludes to a long-term representation of the



interpersonal tie that outlasts representations that are temporally contained within the time frame of the interaction itself.

Taken together, our results suggest that during economic interaction, tie-related mechanisms are more closely involved in experience-based choice than trait empathy. As we established that our trait empathy measure and our interpersonal tie measure were uncorrelated, it seems unlikely that trait empathy plays a direct role in the formation of interpersonal ties and choice behavior during economic interaction. However, trait empathy might play a modulating role in determining the influence of interpersonal ties on experience-based choice. A regression analysis of our empathy and interpersonal ties measures together with an interaction term established that only interpersonal ties predicted



cooperative behavior. This suggests that neither trait empathy, nor its interaction with interpersonal ties, determines subsequent choice behavior.

Importantly, we measured empathy in our experiment by establishing the degree to which someone is able to experience empathic emotions. Therefore, only trait-like aspects of empathy were considered. However, the degree to which empathy is experienced at any given moment in time can change depending on the situational context. For example, a person may feel more empathy toward a person that behaves nicely than toward a person that does not (Singer et al., 2006). As we did not measure the degree of empathy experienced during the experiment directly, we could not assess the influence of empathic experience on choice behavior directly. Other studies suggest that feelings of empathy or activity in empathy-related networks do play a role during prosocial behavior (Hein et al., 2010; Masten et al., 2011), which seems to be at odds with the findings we present here. Importantly however, these studies investigated non-interactive instances of prosocial behavior, in which the other is not expected to reciprocate the good deed bestowed onto them. Experiments like these often use verbal

descriptions or passive viewing of characters. In our experiment, participants were involved in actual interaction with interaction partners. So even if feelings of empathic concern play a role during choice behavior, our results suggest that in interactive experience-based contexts such as the one investigated here, this influence cannot be traced back to a stable empathic trait.

We find correlations between trait empathy and the AIC and ACC only when the other gains at one's own expense, and not when resource transfer is in the opposite direction. In the DOT, negative (resource-losing) events for oneself go hand in hand with positive (resource-gaining) events for the other. Hence, outcomes that benefit the other at one's own expense can either be interpreted as "painful" (resource-losing) events for oneself, as positive (resource-gaining) events for the other, or as a combination of the two. AIC and ACC have been implicated in empathy for both positive (Jabbi et al., 2007) and negative (Wicker et al., 2003) events, but also in a more general brain network known as the "pain matrix," thought to be involved in the perception and appraisal of pain (Singer et al., 2004b). Following reverse inference, one interpretation might be that there is a correlation between trait empathy and one's own "pain" responses when having to share. If true, one would expect to also see them in the pre-interaction DOT. This is not the case (see **Figure 4**), suggesting that the correlations are driven by empathic concern for the other's benefit as a result of having interacted with that person, rather than by pain over one's own loss. In line with this interpretation, we not only found correlations in AIC and ACC, but also in striatal circuitry typically associated with reward processing (Moll et al., 2006; Haber and Knutson, 2010). Although speculative, another person's benefit may register more strongly as a reward signal for individuals with higher empathy levels.

This finding is interesting in light of common theorizing about empathy. Empathy is generally thought to be "other-oriented," being able to tune into or share the affective experiences of others (positive or negative; de Waal, 2008; Singer and Lamm, 2009). Here, we have the special situation where a positive event for the other is accompanied by a negative event for oneself and vice versa (sharing). We only see correlations with empathy when the other gains at one's own expense and not the other way around. This suggests that the "other-orientation" of empathy is modulated by the direction of the resource transfer involved in sharing. Empathic subjects engage empathy-related structures (AIC, ACC) during sharing – even when these events are painful for themselves – but only when their own suffering helps the other along, and not the other way around. Although speculative, empathic subjects may use their ability to empathize to help alleviate, counterbalance, compensate, or justify their own suffering when having to share, while empathy-related responses do not manifest when the tables are turned, possibly as a result of joy over one's own profit.

More generally, the expression of neural correlates of empathy and interpersonal ties in different sharing contexts may tell us something about their function in prosocial behavior. While the interpersonal tie correlate seems to be expressed more strongly in a sharing context where oneself benefits at the other's expense, the network structure involved in empathy seems to be engaged more strongly when one suffers loss that benefits another. This may point to different functional contributions of these regions

to prosocial behavior. The pSTS may be used to keep track of the utility that other agents constitute in the short run, enabling immediate prosocial behavior during interaction. The empathy network on the other hand, may work to overcome negative emotions that are associated with loss. Although our study suggests that this is not translated into short-term investment choices during economic interaction, it may enable cooperative behavior in the long run, countering the adverse effects of absorbing temporary loss associated with social interactions.

Summarizing, previous cooperation success and future cooperation did not correlate with activations in structures related to empathy (AIC, ACC), reward (striatum), or social decision-making (OFC, VMPFC). However, previous cooperation success and future cooperation did correlate with pSTS activation, which we found to be related to interpersonal ties. This suggests that interpersonal ties and trait empathy may recruit different neural networks, of which only the former drives cooperative behavior during interaction. Moreover, individual differences in pSTS activation predicted individual differences in prosocial investment behavior in subsequent rounds of the public good game. This predictive value is important, as it suggests that the pSTS plays a role in keeping track of the dynamically evolving interpersonal tie underlying the willingness to invest in a shared public good. We conclude that during experience-based economic interaction, networks are engaged that keep track of interpersonal ties over longer periods of time, possibly enabling choice–outcome estimation in social settings, with the pSTS emerging as a key element.

A word of caution is also in order. The regions that were highlighted in our experiment have previously been related to many elementary operations, of which we discuss only a few. For example, the pSTS has also been linked to perception of biological motion (Perrett et al., 1989; Bonda et al., 1996), perception of intentional action (Castelli et al., 2000; Saxe et al., 2004) as well as speech perception (Demonet et al., 1992; Mottron et al., 2006). AIC has been implicated in eye movements (Anderson et al., 1994), speech production (Dronkers, 1996; Ackermann and Riecker, 2004), higher-order learning (Seymour et al., 2004), interoceptive awareness (Critchley et al., 2004; Khalsa et al., 2009), and even consciousness at large (Craig, 2009). Likewise, the ACC has been implicated in interoception (Critchley et al., 2004; Khalsa et al., 2009), reward based decision-making (Bush et al., 2002), error detection and conflict monitoring (Gehring et al., 1993), and various related cognitive and emotional processes (Lane et al., 1998; Bush et al., 2000). The striatum has traditionally been implicated in planning and modulation of movement (Rolls, 1994), but recently more prominently in reward processing (Moll et al., 2006; Haber and Knutson, 2010), novelty-based choice behavior (Wittmann et al., 2008), and higher-order learning (Seymour et al., 2004). Together, this paints a picture of a highly complex network in which regions work together to achieve certain functions, plausibly providing different functions in different contexts and at different moments in time. Although we have attempted to isolate empathy and tie-related mechanisms, the idea that these mechanisms are solely implemented in the regions that we highlight is most likely an oversimplification. More likely, these functions are achieved through interactions between and within

regions, executing various more elementary subprocesses to provide more complex higher-order functions. Our study does not allow us to tease apart empathy and tie-related mechanisms into their elementary processes. However, it does provide pointers as to where these higher-order functions seem to be expressed most strongly, and what their role is in interactive experience-based choice. Relatedly, it deserves mention that we did not have a control

group in which the other participant in the post-interaction DOT was either a non-human player (computer) or another anonymous participant from the experiment. Although it is unlikely that the specific activation patterns we observed can be attributed to non-specific effects such as game repetition, boredom or exhaustion, future studies employing a similar design might do well to include such controls.

## REFERENCES

- Ackermann, H., and Riecker, A. (2004). The contribution of the insula to motor aspects of speech production: a review and a hypothesis. *Brain Lang.* 89, 320–328.
- Anderson, T. J., Jenkins, I. H., Brooks, D. J., Hawken, M. B., Frackowiak, R. S. J., and Kennard, C. (1994). Cortical control of saccades and fixation in man – a pet study. *Brain* 117, 1073–1084.
- Batson, C. D. (1991). *The Altruism Question: Toward a Social Psychological Answer*. Hillsdale, NJ: L. Erlbaum.
- Benjamin, Y., and Hochberg, Y. (1995). Controlling the false discovery rate – a practical and powerful approach to multiple testing. *J. R. Stat. Soc. Series B Stat. Methodol.* 57, 289–300.
- Boddaert, N., Chabane, N., Gervais, H., Good, C. D., Bourgeois, M., Plumet, M. H., Barthelemy, C., Mouren, M. C., Artiges, E., Samson, Y., Brunelle, F., Frackowiak, R. S. J., and Zilbovicius, M. (2004). Superior temporal sulcus anatomical abnormalities in childhood autism: a voxel-based morphometry MRI study. *Neuroimage* 23, 364–369.
- Bonda, E., Petrides, M., Ostry, D., and Evans, A. (1996). Specific involvement of human parietal systems and the amygdala in the perception of biological motion. *J. Neurosci.* 16, 3737–3744.
- Bush, G., Luu, P., and Posner, M. I. (2000). Cognitive and emotional influences in anterior cingulate cortex. *Trends Cogn. Sci. (Regul. Ed.)* 4, 215–222.
- Bush, G., Vogt, B. A., Holmes, J., Dale, A. M., Greve, D., Jenike, M. A., and Rosen, B. R. (2002). Dorsal anterior cingulate cortex: a role in reward-based decision making. *Proc. Natl. Acad. Sci. U.S.A.* 99, 523–528.
- Castelli, F., Happe, F., Frith, U., and Frith, C. (2000). Movement and mind: a functional imaging study of perception and interpretation of complex intentional movement patterns. *Neuroimage* 12, 314–325.
- Cehajic, S., Brown, R., and Gonzalez, R. (2009). What do i care? Perceived ingroup responsibility and dehumanization as predictors of empathy felt for the victim group. *Group Process. Intergroup Relat.* 12, 715–729.
- Cooper, J. C., Kreps, T. A., Wiebe, T., Pirkel, T., and Knutson, B. (2010). When giving is good: ventromedial prefrontal cortex activation for others' intentions. *Neuron* 67, 511–521.
- Craig, A. D. (2009). How do you feel – now? The anterior insula and human awareness. *Nat. Rev. Neurosci.* 10, 59–70.
- Critchley, H. D., Wiens, S., Rotshtein, P., Ohman, A., and Dolan, R. J. (2004). Neural systems supporting interoceptive awareness. *Nat. Neurosci.* 7, 189–195.
- de Waal, F. B. M. (2008). Putting the altruism back into altruism: the evolution of empathy. *Annu. Rev. Psychol.* 59, 279–300.
- Decety, J., Jackson, P. L., Sommerville, J. A., Chaminade, T., and Meltzoff, A. N. (2004). The neural bases of cooperation and competition: an fMRI investigation. *Neuroimage* 23, 744–751.
- Demonet, J. F., Chollet, F., Ramsay, S., Cardebat, D., Nespoulous, J. L., Wise, R., Rascol, A., and Frackowiak, R. (1992). The anatomy of phonological and semantic processing in normal subjects. *Brain* 115, 1753–1768.
- Dronkers, N. F. (1996). A new brain region for coordinating speech articulation. *Nature* 384, 159–161.
- Eisenberg, N. (2000). “Empathy and sympathy,” in *Handbook of Emotions*, 2nd Edn, eds M. Lewis and J. M. Haviland-Jones (New York: The Guilford Press), 677–691.
- Farrow, T. F. D., Zheng, Y., Wilkinson, I. D., Spence, S. A., Deakin, J. F. W., Tarriner, N., Griffiths, P. D., and Woodruff, P. W. R. (2001). Investigating the functional anatomy of empathy and forgiveness. *Neuroreport* 12, 2433–2438.
- Fehr, E., and Fischbacher, U. (2003). The nature of human altruism. *Nature* 425, 785–791.
- Fehr, E., and Gächter, S. (2002). Altruistic punishment in humans. *Nature* 415, 137–140.
- Gehring, W. J., Goss, B., Coles, M. G. H., Meyer, D. E., and Donchin, E. (1993). A neural system for error-detection and compensation. *Psychol. Sci.* 4, 385–390.
- Haber, S. N., and Knutson, B. (2010). The reward circuit: linking primate anatomy and human imaging. *Neuropsychopharmacology* 35, 4–26.
- Hampton, A. N., Bossaerts, P., and O’Doherty, J. P. (2008). Neural correlates of mentalizing-related computations during strategic interactions in humans. *Proc. Natl. Acad. Sci. U.S.A.* 105, 6741–6746.
- Haruno, M., and Frith, C. D. (2010). Activity in the amygdala elicited by unfair divisions predicts social value orientation. *Nat. Neurosci.* 13, 160–161.
- Haruno, M., and Kawato, M. (2009). Activity in the superior temporal sulcus highlights learning competence in an interaction game. *J. Neurosci.* 29, 4542–4547.
- Hein, G., Silani, G., Preuschoff, K., Batson, C. D., and Singer, T. (2010). Neural responses to ingroup and outgroup members' suffering predict individual differences in costly helping. *Neuron* 68, 149–160.
- Isaac, R. M., and Walker, J. M. (1988). Group-size effects in public-goods provision – the voluntary contributions mechanism. *Q. J. Econ.* 103, 179–199.
- Jabbi, M., Swart, M., and Keysers, C. (2007). Empathy for positive and negative emotions in the gustatory cortex. *Neuroimage* 34, 1744–1753.
- Khalsa, S. S., Rudrauf, D., Feinstein, J. S., and Tranel, D. (2009). The pathways of interoceptive awareness. *Nat. Neurosci.* 12, 1494–1496.
- King-Casas, B., Tomlin, D., Anen, C., Camerer, C. F., Quartz, S. R., and Montague, P. R. (2005). Getting to know you: reputation and trust in a two-person economic exchange. *Science* 308, 78–83.
- Kriegeskorte, N., Simmons, W. K., Bellgowan, P. S. F., and Baker, C. I. (2009). Circular analysis in systems neuroscience: the dangers of double dipping. *Nat. Neurosci.* 12, 535–540.
- Lane, R. D., Reiman, E. M., Axelrod, B., Yun, L. S., Holmes, A., and Schwartz, G. E. (1998). Neural correlates of levels of emotional awareness: evidence of an interaction between emotion and attention in the anterior cingulate cortex. *J. Cogn. Neurosci.* 10, 525–535.
- Ledyard, J. O. (1995). “Public goods: a survey of experimental research,” in *The Handbook of Experimental Economics*, eds A. Roth and J. Kagel (Princeton, NJ: Princeton University Press), 111–181.
- Liebrand, W. B. G. (1984). The effect of social motives, communication and group-size on behavior in an N-person multi-stage mixed-motive game. *Eur. J. Soc. Psychol.* 14, 239–264.
- Masten, C. L., Morelli, S. A., and Eisenberger, N. I. (2011). An fMRI investigation of empathy for ‘social pain’ and subsequent prosocial behavior. *Neuroimage* 55, 381–388.
- Mathur, V. A., Harada, T., Lipke, T., and Chiao, J. Y. (2010). Neural basis of extraordinary empathy and altruistic motivation. *Neuroimage* 51, 1468–1475.
- Mehrabian, A. (1997). Relations among personality scales of aggression, violence, and empathy: validation evidence bearing on the risk of eruptive violence scale. *Aggress. Behav.* 23, 433–445.
- Mehrabian, A., and Epstein, N. (1972). A measure of emotional empathy. *J. Pers.* 40, 525–543.
- Moll, J., Krueger, F., Zahn, R., Pardini, M., De Oliveira-Souza, R., and Grafman, J. (2006). Human fronto-mesolimbic networks guide decisions about charitable donation. *Proc. Natl. Acad. Sci. U.S.A.* 103, 15623–15628.
- Mottronen, R., Calvert, G. A., Jaaskelainen, I. P., Matthews, P. M., Theisen, T., Tuomainen, J., and Sams, M. (2006). Perceiving identical sounds as speech or non-speech modulates activity in the left posterior superior temporal sulcus. *Neuroimage* 30, 563–569.
- Perrett, D. I., Harries, M. H., Bevan, R., Thomas, S., Benson, P. J., Mistlin, A. J., Chitty, A. J., Hietanen, J. K., and Ortega, J. E. (1989). Frameworks of analysis for the neural representation of animate objects and actions. *J. Exp. Biol.* 146, 87–113.
- Rakow, T., and Newell, B. R. (2010). Degrees of uncertainty: an overview and framework for future research on experience-based choice. *J. Behav. Decis. Mak.* 23, 1–14.
- Rilling, J. K., Gutman, D. A., Zeh, T. R., Pagnoni, G., Berns, G. S., and Kilts,

- C. D. (2002). A neural basis for social cooperation. *Neuron* 35, 395–405.
- Rolls, E. T. (1994). Neurophysiology and cognitive functions of the striatum. *Rev. Neurol. (Paris)* 150, 648–660.
- Saxe, R., Xiao, D. K., Kovacs, G., Perrett, D. I., and Kanwisher, N. (2004). A region of right posterior superior temporal sulcus responds to observed intentional actions. *Neuropsychologia* 42, 1435–1446.
- Seymour, B., O’doherly, J. P., Dayan, P., Koltzenburg, M., Jones, A. K., Dolan, R. J., Friston, K. J., and Frackowiak, R. S. (2004). Temporal difference models describe higher-order learning in humans. *Nature* 429, 664–667.
- Singer, T., Kiebel, S. J., Winston, J. S., Dolan, R. J., and Frith, C. D. (2004a). Brain responses to the acquired moral status of faces. *Neuron* 41, 653–662.
- Singer, T., Seymour, B., O’doherly, J., Kaube, H., Dolan, R. J., and Frith, C. D. (2004b). Empathy for pain involves the affective but not sensory components of pain. *Science* 303, 1157–1162.
- Singer, T., and Lamm, C. (2009). The social neuroscience of empathy. *Ann. N. Y. Acad. Sci.* 1156, 81–96.
- Singer, T., Seymour, B., O’doherly, J. P., Stephan, K. E., Dolan, R. J., and Frith, C. D. (2006). Empathic neural responses are modulated by the perceived fairness of others. *Nature* 439, 466–469.
- Sturmer, S., Snyder, M., Kropp, A., and Siem, B. (2006). Empathy-motivated helping: the moderating role of group membership. *Pers. Soc. Psychol. Bull.* 32, 943–956.
- Tankersley, D., Stowe, C. J., and Huettel, S. A. (2007). Altruism is associated with an increased neural response to agency. *Nat. Neurosci.* 10, 150–151.
- van Dijk, F., Sonnemans, J., and van Winden, F. (2002). Social ties in a public good experiment. *J. Public Econ.* 85, 275–299.
- Van Lange, P. A. M. (1999). The pursuit of joint outcomes and equality in outcomes: an integrative model of social value orientation. *J. Pers. Soc. Psychol.* 77, 337–349.
- van Winden, F., Stallen, M., and Ridderinkhof, K. R. (2008). On the nature, modeling, and neural bases of social ties. *Adv. Health Econ. Health Serv. Res.* 20, 125–159.
- Wicker, B., Keysers, C., Plailly, J., Royet, J. P., Gallese, V., and Rizzolatti, G. (2003). Both of us disgusted in My Insula: the common neural basis of seeing and feeling disgust. *Neuron* 40, 655–664.
- Wittmann, B. C., Daw, N. D., Seymour, B., and Dolan, R. J. (2008). Striatal activity underlies novelty-based choice in humans. *Neuron* 58, 967–973.
- commercial or financial relationships that could be construed as a potential conflict of interest.

Received: 30 November 2011; paper pending published: 17 December 2011; accepted: 10 February 2012; published online: 05 March 2012.

Citation: Fahrenfort JJ, van Winden F, Pelloux B, Stallen M and Ridderinkhof KR (2012) Neural correlates of dynamically evolving interpersonal ties predict prosocial behavior. *Front. Neurosci.* 6:28. doi: 10.3389/fnins.2012.00028

This article was submitted to *Frontiers in Decision Neuroscience*, a specialty of *Frontiers in Neuroscience*.

Copyright © 2012 Fahrenfort, van Winden, Pelloux, Stallen and Ridderinkhof. This is an open-access article distributed under the terms of the Creative Commons Attribution Non Commercial License, which permits non-commercial use, distribution, and reproduction in other forums, provided the original authors and source are credited.

**Conflict of Interest Statement:** The authors declare that the research was conducted in the absence of any