NeuroImage 157 (2017) 598–611

Contents lists available at ScienceDirect

NeuroImage

journal homepage: www.elsevier.com/locate/neuroimage

How do self-interest and other-need interact in the brain to determine altruistic behavior?

Jie Hu[a,b], Yue Li[a,b], Yunlu Yin[a,b], Philip R. Blue[a,b], Hongbo Yu[a,b], Xiaolin Zhou[a,b,c,d,e,*]

[a] Center for Brain and Cognitive Sciences, Peking University, Beijing 100871, China
[b] School of Psychological and Cognitive Sciences, Peking University, Beijing 100871, China
[c] Key Laboratory of Machine Perception (Ministry of Education), Peking University, Beijing 100871, China
[d] Beijing Key Laboratory of Behavior and Mental Health, Peking University, Beijing, China
[e] PKU-IDG/McGovern Institute for Brain Research, Peking University, Beijing 100871, China

ABSTRACT

Altruistic behavior, i.e., promoting the welfare of others at a cost to oneself, is subserved by the integration of various social, affective, and economic factors represented in extensive brain regions. However, it is unclear how different regions interact to process/integrate information regarding the helper's interest and recipient's need when deciding whether to behave altruistically. Here we combined an interactive game with functional Magnetic Resonance Imaging (fMRI) and transcranial direct current stimulation (tDCS) to characterize the neural network underlying the processing/integration of self-interest and other-need. At the behavioral level, high self-risk decreased helping behavior and high other-need increased helping behavior. At the neural level, activity in medial prefrontal cortex (MPFC) and right dorsolateral prefrontal cortex (rDLPFC) were positively associated with self-risk levels, and activity in right inferior parietal lobe (rIPL) and rDLPFC were negatively associated with other-need levels. Dynamic causal modeling further suggested that both MPFC and rIPL were extrinsically connected to rDLPFC; high self-risk enhanced the effective connectivity from MPFC to rDLPFC, and the modulatory effect of other-need on the connectivity from rIPL to rDLPFC positively correlated with the modulatory effect of other-need on individuals' helping rate. Two tDCS experiments provided causal evidence that rDLPFC affects both self-interest and other-need concerns, and rIPL selectively affects the other-need concerns. These findings suggest a crucial role of the MPFC-IPL-DLPFC network during altruistic decision-making, with rDLPFC as a central node for integrating and modulating motives regarding self-interest and other-need.

Introduction

When seeing a child who knowingly cannot swim falling into a river, most people will immediately jump into the water to help. But, if you know that jumping into the river may not save the child or may even lead to the drowning of you both, would you still jump in without hesitation? Humans often help others in need at a cost to themselves (Moll et al., 2006; Batson et al., 2007), but at times we are not willing to help others in need (Warneken and Tomasello, 2009; Bode et al., 2015). A variety of factors have been proposed to affect whether individuals are willing to help, such as the individuals' ability, affective state, perception of responsibility, and knowledge of existing social norms (Penner et al., 2005). Among these factors, other-oriented emotional responses (e.g., empathy) and self-oriented motives (e.g., cost of helping) are two fundamental factors underlying the decision to engage or not in altruistic behavior (Batson et al., 1983; Hein et al., 2010; Tusche et al., 2016).

A compelling explanation for altruistic behaviors is the empathy-altruism hypothesis (Toi and Batson, 1982; Batson and Shaw, 1991; Penner et al., 2005). A stronger empathic concern causes individuals to give up more money to help others (FeldmanHall et al., 2015). If individuals are unable to recognize the other's need (Warneken et al., 2007; FeldmanHall et al., 2013) or do not value the welfare of the other (Batson et al., 2007; Hein et al., 2010), they show less empathic concern and are less willing to help the other (Batson et al., 2007). These findings support the empathy-altruism hypothesis, which claims that empathic emotions for others' suffering, such as sympathy and compassion, elicit an individual's altruistic motives to relieve or attenuate the distress of others (Batson and Shaw, 1991; Penner et al., 2005).

* Correspondence to: Department of Psychology, Peking University, Beijing 100871, China.
E-mail address: xz104@pku.edu.cn (X. Zhou).

http://dx.doi.org/10.1016/j.neuroimage.2017.06.040
Received 21 December 2016; Accepted 16 June 2017
Available online 20 June 2017
1053-8119/

However, in many contexts, individuals do not always help others in need: if helping behavior entails personal costs or a potential loss to the helper (Warneken and Tomasello, 2009; Bode et al., 2015), like negatively affecting the potential helper's good mood (Isen and Simmonds, 1978) or coming at a financial cost to the helper (Moll et al., 2006), concerns for self-interest could override the participant's altruistic motives and reduce helping behavior (Batson et al., 1983). These findings are in line with the cost-reward models of altruistic behavior, which posit that individuals weigh the costs and rewards of their helping behavior and make a decision that will maximize their benefits and minimize their costs (e.g. Hamilton's Inclusive Fitness Model, Hamilton, 1964a, 1964b; Penner et al., 2005).

Although theories and empirical evidence suggest that helping behavior is a deliberative response determined by the consideration of the need of the recipient and the interest of the helper (Heinsohn and Legge, 1999; Fehr and Krajbich, 2014), a key question which has not been addressed is how people simultaneously weigh or integrate the altruistic and the egoistic motives when deciding whether to help others, especially under contexts in which the consequence of helping is unknown.

A survey of previous neuroimaging studies shows that altruistic behavior is possibly supported by three sets of brain regions: 1) empathy-related regions, such as the anterior insula (AI) and the temporoparietal junction (TPJ); 2) reward-related regions, such as the ventral tegmental area (VTA), the subgenual anterior cingulate cortex (sgACC), the caudate, and the ventral striatum (VS); and 3) cognitive control-related regions, such as the dorsolateral prefrontal cortex (DLPFC) and the dorsal part of ACC. Specifically, stronger activity of the empathy-related regions reflects stronger empathic concern for others' feelings and facilitates helping behavior (Mathur et al., 2010; Hein et al., 2010; Waytz et al., 2012; FeldmanHall et al., 2015). The activity of the reward-related regions is positively associated with helping behavior (Moll et al., 2006; FeldmanHall et al., 2015; Hu et al., 2015), suggesting that helping others is a kind of social reward and provides the helper with satisfaction and pleasure (Harbaugh et al., 2007). Finally, activity in the cognitive control-related regions correlates with altruistic helping behavior (FeldmanHall et al., 2015; Hu et al., 2015), suggesting that these regions are engaged in inhibiting egoistic motives (Knoch et al., 2006; Ruff et al., 2013) or modulating altruistic and egoistic motives (Nihonsugi et al., 2015) when an altruistic decision is made.

Most of these studies, however, focused on assessing participants' altruistic behavior by measuring how much money they were willing to spend (FeldmanHall et al., 2013, 2015) or how much pain they were willing to take (Hein et al., 2010, 2011) to alleviate others' suffering. These studies did not differentiate helpers' concerns for their own interests and for the welfare of others during altruistic helping. Also, the measurements of altruistic behavior in these studies are the output of the trade-off between self-interest and other-regarding motives, which limits the ability to understand the unique effects of self-interest and other-regard on altruistic behavior. To more precisely assess individuals' altruistic tendencies, it is critical to distinguish individuals' concerns for self-interest from their other-regarding motives. Moreover, in past research, when participants considered how much money to spend or how much electric shock to receive to help others, the helping behavior always increased the welfare of the recipient. In real life, however, the helper is not always sure whether the helping behavior actually helps the recipient in the end. To more precisely reveal the neural basis of altruistic behavior, it is therefore important to separate the helping motives from the consequence of help.

The aim of the current study is to examine 1) how individuals process the information of self-interest and other-need and integrate these two dimensions of motives in a risky situation, and 2) how this integration or trade-off is implemented in the brain to determine altruistic behavior. Here, we sought to deepen our understanding of the neural basis of altruistic helping in the following ways. First, we developed a novel paradigm in which participants were required to weigh both their own probability and others' probability of being punished when deciding whether to help others. Second, we attempted to identify the brain regions associated with representing the motives of self-interest and other-need, and used dynamic causal modeling (DCM) to clarify how these regions interact as a functional network to support the helping decision. Finally, to further examine the role of the brain regions involved in altruistic behaviors, we examined the causal relationship between the regions of interest and individuals' self- and other-regarding tendencies using transcranial direct current stimulation (tDCS).

In the novel interactive paradigm, the participant played a dice game with a randomly chosen partner in which both of them rolled three dice in total. In each trial of the game, each player's goal was to roll at least 9 points; if a player failed to roll 9 points or more in any game trial, he/she would receive a punishment (i.e., 1 s of unpleasant noise administration). Before the third dice was revealed, the player with more points was given an opportunity to transfer 1 point to the player who had fewer points. That is, the player with the score advantage could help decrease the disadvantaged player's risk of being punished by increasing his/her own risk of being punished. Thus, the behavioral measure of altruistic helping was defined as the participant donating 1 point to the other player. We manipulated the self-risk and other-need by varying the probability each person would be punished after the first two dice outcomes were revealed (for detailed information, see *Materials and methods*). This manipulation allowed us not only to differentiate participants' concerns for self-interest and other-need, but also to examine altruistic behavior under a context in which helping motives were separated from the consequence of the help.

We hypothesized that when the participant had a relatively high risk of being punished, compared with no risk, he/she would be more concerned with his/her own welfare and less likely to help others. Moreover, when the other player had a relatively high risk of being punished, compared with a relatively low risk of being punished, the participant would have a stronger concern for the recipient's welfare and be more likely to offer help. At the neural level, we were interested in identifying the brain regions involved in self-risk and other-need processing. Our aim was to test whether reward-related regions (i.e. VTA, sgACC, and VS) are engaged in self-risk processing, and whether empathy-related regions (i.e. AI, and TPJ) are engaged in other-need processing. More generally, some other regions, like medial prefrontal cortex (MPFC) and parietal cortices, are involved in value processing (Pinel et al., 2004; Kahnt et al., 2014; Sul et al., 2015) and high-level social cognitive processes, especially in other-regarding tasks (Decety and Lamm, 2007; Donaldson et al., 2015). Therefore, we also examined whether these regions (i.e. MPFC and parietal cortices) are implicated in self-risk and/or other-need processing. Moreover, DLPFC, especially the right DLPFC (rDLPFC), has been repeatedly implicated in inhibiting selfish motives (Knoch et al., 2006; Baumgartner et al., 2011; Ruff et al., 2013; Strang et al., 2014; Zhu et al., 2014) and in modulating other-regarding motives (Nihonsugi et al., 2015) to promote prosocial behavior. Given that DLPFC serves as a critical region for integrating and modulating information from different sources (Buckholtz and Marois, 2012), we expected that DLPFC is involved in the integration of self-risk and other-need processing when making altruistic decisions.

Materials and methods

Participants

Twenty-one right-handed undergraduate and graduate students (age range 18–25 years, mean = 20.67, s.d. = 1.88, 14 females) participated in a pilot behavioral experiment. Thirty right-handed undergraduate and graduate students participated in the fMRI experiment. Four participants were excluded due to excessive head movement ($> \pm 3$ mm in translation and/or $> \pm 3°$ in rotation) and one

participant was excluded because he fell asleep during scanning. The remaining 25 participants were aged between 18 and 26 years (mean = 21.56, s.d. = 2.45; 12 females). In addition, two groups of right-handed undergraduate and graduate students participated in the tDCS experiments, with fifty-six participants (age range 18–26 years, mean = 21.41, s.d. = 2.21, 36 females) in the tDCS over rDLPFC and fifty-eight participants (age range 17–24 years, mean = 20.12, s.d. = 2.00, 39 females) in the tDCS over rIPL. No participant had a history of psychiatric, neurological, or cognitive disorders. Informed written consent was obtained from each participant before the experiments. The study was in accordance with the Declaration of Helsinki and was approved by the Ethics Committee of the Department of Psychology, Peking University.

Design and procedures

Interactive dice game

In the newly developed paradigm, the participant played a dice game with a randomly chosen partner in which both of them rolled three dice in total. In each trial of the game, each player's goal was to roll at least 9 points; if a player failed to roll 9 points or more in any game trial, he/she would receive a punishment (i.e., one-second of unpleasant noise administration). There was no reward for attaining higher than 9 points. Before the third dice was revealed, the player with more points was given an opportunity to transfer 1 point to the player who had fewer points. We manipulated the self-risk and other-need by parametrically and independently varying the probability that each player would be punished after the rolling of the first two dice. In designing and selecting specific trials for the study, we attempted to avoid the collinearity problem between the two variables (see below). Note also, the current manipulation allowed us to define the self-risk as “high self-risk” or “no self-risk” when the participant had high or zero probability of being punished after transferring 1 point to the other player, and to define the other-need for help as “high other-need” or “low other-need” when the recipient had a high or low probability of being punished after the first two dice were revealed (see below).

For critical trials, the participant's total number of points on the first two dice ranged from 11 to 6, with his/her risk of being punished after transferring 1 point to the other player ranging from 0 (i.e., having 11, 10, and 9 points) to 0.50 (i.e., having 6 points). The levels of self-risk (0, 0.17, 0.33, 0.50) were calculated by computing the probability of having 9 or more points total on the 3 dice after helping the other player (i.e., giving one point to the other player). The other player's total number of points on the first two dice ranged from 7 to 2, with their risk of being punished ranging from 0.17 to 1.00 (0.17, 0.33, 0.50, 0.67, 0.83, 1.00). The levels of other-need were calculated by computing the probability of having 9 or more points after rolling the third dice without the potential transfer of 1 point from the participant. For filler trials, the participant's probability of being punished ranged from 0.17 to 1.00, and the other player's probability of being punished ranged from 0 to 0.50. In the filler trials, the participant him/herself gained fewer points than the other player after the first two dice were rolled, and had nothing to do but to wait for the other player to make the decision.

For both the pilot and fMRI experiments, upon arriving at the laboratory, the participant met three confederates and was told that they would play an interactive dice game together through intranet. Then, the three confederates were led to another room to prepare for the experiment. The participant was told that, in each trial, he/she was randomly paired with one confederate (the other player) and each got to roll three dice. The participant was informed that the other two players who were not paired with him/her in the current trial would be paired with each other and perform the same task. The participant was also told that all the players’ personal information (e.g. photos) and decision in each trial (e.g. help or not help) would not be shown to each other during the game. That is, the participant could be paired with any one of the three confederates in each trial, but he/she did not know the other player's identity. Such an anonymous manipulation prevented potential reputation concerns. To examine the participants’ altruistic helping, we focused our analysis on the critical trials in which the participant rolled more points than the confederate and was given the opportunity to help; trials in which the participant rolled fewer points (i.e., the potential recipient of help) were included as fillers.

To avoid any confounding effects resulting from the outcome of the last dice (e.g., whether the altruistic helping successfully protected the recipients from punishments or whether the altruistic helping led the helper to receive punishments), we did not present the participant with feedback regarding the outcome of the third dice roll in the scanner. The participant was told that after scanning, the computer would randomly select 20 trials and reveal the outcome of the third dice trial-by-trial to the participant. He/she and the other players would receive punishments according to the outcomes of the 20 trials. This schedule was to ensure that the participant treated each trial in equal terms.

Before task instructions, the participant heard 30 one-second noise clips of varying loudness in randomized order, and rated the unpleasantness on a visual analog scale (VAS) (Price et al., 1994; Park et al., 2011). The very left extreme of the VAS was labeled as 0 (not unpleasant at all), and the very right extreme was labeled as 10 (extremely unpleasant). The noise stimuli were delivered by a pair of AKG K271 MKII headphones and controlled by Presentation software (Neurobehavioral System Inc.). The participant was then informed that noise punishments were tailored to the ratings of each person, and that each person's punishment administration in the game would be equivalent to the noise clip rated as 8 out of 10 by that particular participant.

Following the task instructions and a test of comprehension of the task instructions, participants performed two task sessions in the behavioral laboratory for the pilot experiment, or underwent two task fMRI scanning sessions for the fMRI experiment. For the pilot experiment, each session consisted of 70 trials (35 critical and 35 filler trials); and for the fMRI experiment, each session consisted of 96 trials (12 for each of the four critical conditions and 48 filler trials) and lasted for about 20 min. See Fig. 1 for details of the trial sequence. The experiment was administered by Presentation software (Neurobehavioral System Inc.) to control the presentation and timing of stimuli. Unknown to the participants, the sequences of the trials were predetermined and pseudorandomized with the restriction that no more than 3 consecutive trials were of the same critical conditions. Before the formal fMRI scanning, the participant performed 10 trials of the dice game to get familiar with the task outside the scanner.

To collect convergent evidence for our arguments and/or to facilitate data analyses, we also categorized the experimental trials into factorial designs. In the pilot experiment, we categorized the levels of self-risk into two groups (no self-risk: 0; high self-risk: 0.17, 0.33, 0.50) and levels of other-need into two groups (low other-need: 0.17, 0.33, 0.50; high other-need: 0.67, 0.83, 1.00). This categorization allowed the experiment to be formulated as a two-by-two within-participant factorial design, with the first factor referring to the participant's risk involved in helping (no risk vs. high risk) and the second factor referring to the other player's need for help (low need vs. high need). This gave rise to four critical conditions: no self-risk and low other-need (NS_LO), no self-risk and high other-need (NS_HO), high self-risk and low other-need (HS_LO), and high self-risk and high other-need (HS_HO). Specifically, in the pilot experiment, there were 70 critical trials in total, with 36 trials for the no self-risk conditions and 34 trials for the high self-risk conditions. These 70 trials can also be categorized as low other-need vs. high other-need. Fig. 2A (left panel) presents the details of the trial set. In addition, there were 70 filler trials in total in the pilot experiment. The distribution for filler trials was the same as that for critical trials except that it was the participant who rolled fewer points in the filler trials.

A problem with the trial setup in the pilot experiment was that there

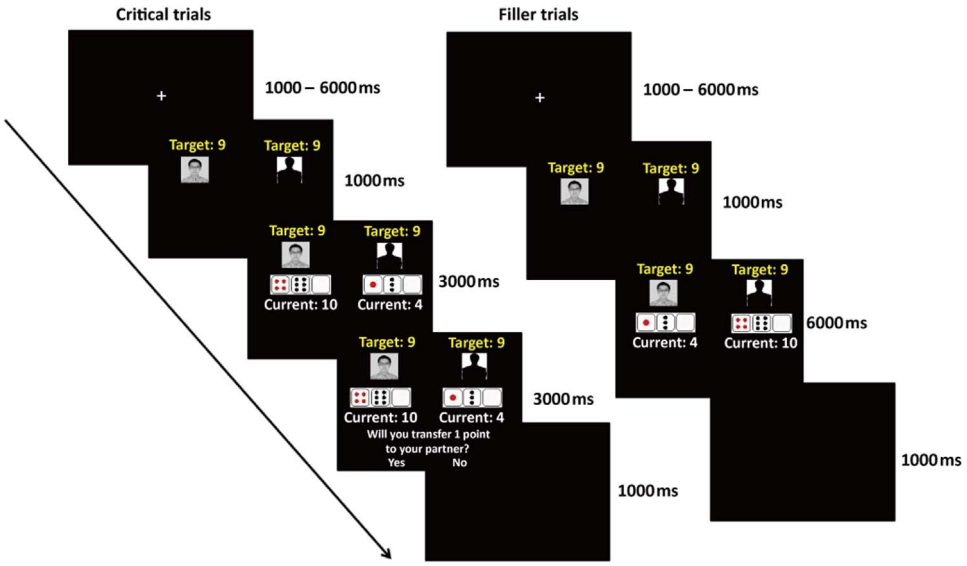

Fig. 1. Schematic diagram of the experiment of the fMRI experiment. Each trial began with a fixation sign at the center of the screen for either 1000, 1500, 2000, 2500, 3000, 3500, 4000, 4500, 5000, 5500, or 6000 ms. Then, the participant's own portrait and a faceless silhouette representing the other player, together with each player's goal for each trial (obtaining 9 points) were presented on the left and right side of the screen, respectively, for 1000 ms, suggesting to the participant that the computer had paired him/her with a player and was rolling the first two dice. The positions of these two figures were counterbalanced between participants. Next, the outcomes of the first two dice for each person were presented for 3000 ms. For the critical trials, in which the participant had more points than the other player, the sentence "Will you transfer 1 point to your partner?" and the options "Yes" and "No" in Chinese were then presented on the lower part of the screen. The participant was instructed to press buttons corresponding to "Yes" or "No" within 3000 ms to make the choice. Once the participant pressed the corresponding button, a box was placed around the chosen option and this screen was presented for the remainder of the 3000 ms before moving on to the next trial. For the filler trials, in which the participant had fewer points than the other player, the outcomes of the first two dice were presented for 6000 ms, and the participant was not required to make any response. After a 1000 ms blank screen, the next trial started.

was a systematic confound: in most HS_LO trials, transferring one point would reverse the relative advantageous status of the two players (i.e., transferring one point would lead to a higher risk for the helper than for the other player to be punished). In the other three conditions, transferring would not lead to such a reverse of relative advantageous status for the two players. Therefore, in the fMRI experiment, we adjusted the trial distribution to prevent this confound.

In the fMRI experiment, there were three requirements for the critical trials: 1) the participant always had more points than the other player, 2) transferring one point would not lead to a higher risk for the participant than for the other player, and 3) the amounts of trials for the four critical conditions were equal. To fulfill these requirements, we had to exclude over half of the "high self-risk and low other-need" trials in the pilot experiment; as a result, the amount of trials in the "high self-risk & low other-need" condition was much lower than the other conditions. To balance the amount of trials in each condition and to make the outcome distribution approximate to the theoretical distribution as closely as possible, we decided to define trials with other-need of 0.67 as "low", rather than as "high". Specifically, we categorized the levels of self-risk into two groups (no self-risk: 0; high self-risk: 0.17, 0.33, 0.50) and levels of other-need into two groups (low other-need: 0.17, 0.33, 0.50, 0.67; high other-need: 0.83, 1.00). There were 96 critical trials in total, with 24 trials for each of the four experimental conditions (see Fig. 2A, right panel for information of the trial set). Note that, although the pilot experiment and the fMRI experiment had different definitions for the trial set (i.e., the other-need of 0.67 was defined as "high" in the pilot experiment and as "low" in the fMRI experiment), the two experiments showed the same pattern of effects of self-interest and other-need (Fig. 2B). Moreover, the logistic regression analysis of the fMRI behavioral data provided convergent evidence for the results of the factorial analyses (see *Behavioral results*), indicating that the behavioral effects of the manipulated factors were not influenced by different categorizations of trials.

There were two reasons for having wide distribution of trials in this study: 1) to make the task more natural to the participants and 2) to enable parametric analysis of the fMRI data. As a result, such a trial setup allowed the task to have an equal number of trials (24 trials) in each of the 4 critical conditions in the factorial design. In addition, there were 96 filler trials in total in the fMRI experiment. The distribution for filler trials was the same as that for critical trials except that it was the participant who rolled fewer points. Thus in the current design, in half of the trials, the participant rolled more points than the other player and in the other half, the participant rolled fewer points. Importantly, no participant reported any oddity concerning the experienced frequencies of different outcomes in the experiment.

Loss aversion task

One could argue that in our dice game, participants' altruistic behavior was influenced by their risk-taking or loss aversion tendencies. Therefore, after the scanning in the fMRI experiment, to measure individuals' risk-taking behavior and loss aversion tendencies, we asked the participants to perform a gamble task (Tom et al., 2007; De Martino et al., 2010) in which they decided whether to accept a set of gambles with equal probability (50%) of winning or losing a variable amount of money. The task procedure and the trial-set matrix were exactly the same as the loss aversion task used in De Martino et al. (2010). The participant was told that, after the task, one trial would be randomly chosen and the final payment would be determined by his/her actual decision. That is, the amount earned or lost in the loss aversion task would be added to or deducted from the basic payment of 150 RMB (about 23 US dollars) for taking part in the scanning and the loss aversion task. The participants were paid the final payment in cash after the experiment.

To assess individuals' sensitivity to potential gain and loss during gambling, we fitted a logistic regression with the acceptability of gambling as the dependent variable and the amount of potential gain and loss as independent variables. In accordance with Tom and colleagues (2007), we calculated the behavioral loss aversion λ by dividing the absolute value of weight on loss by the weight on gain. To examine whether individuals' helping behavior in our dice game was influenced by their loss aversion tendency under risk, we conducted correlation analyses between the log-transformed loss aversion λ (Tom

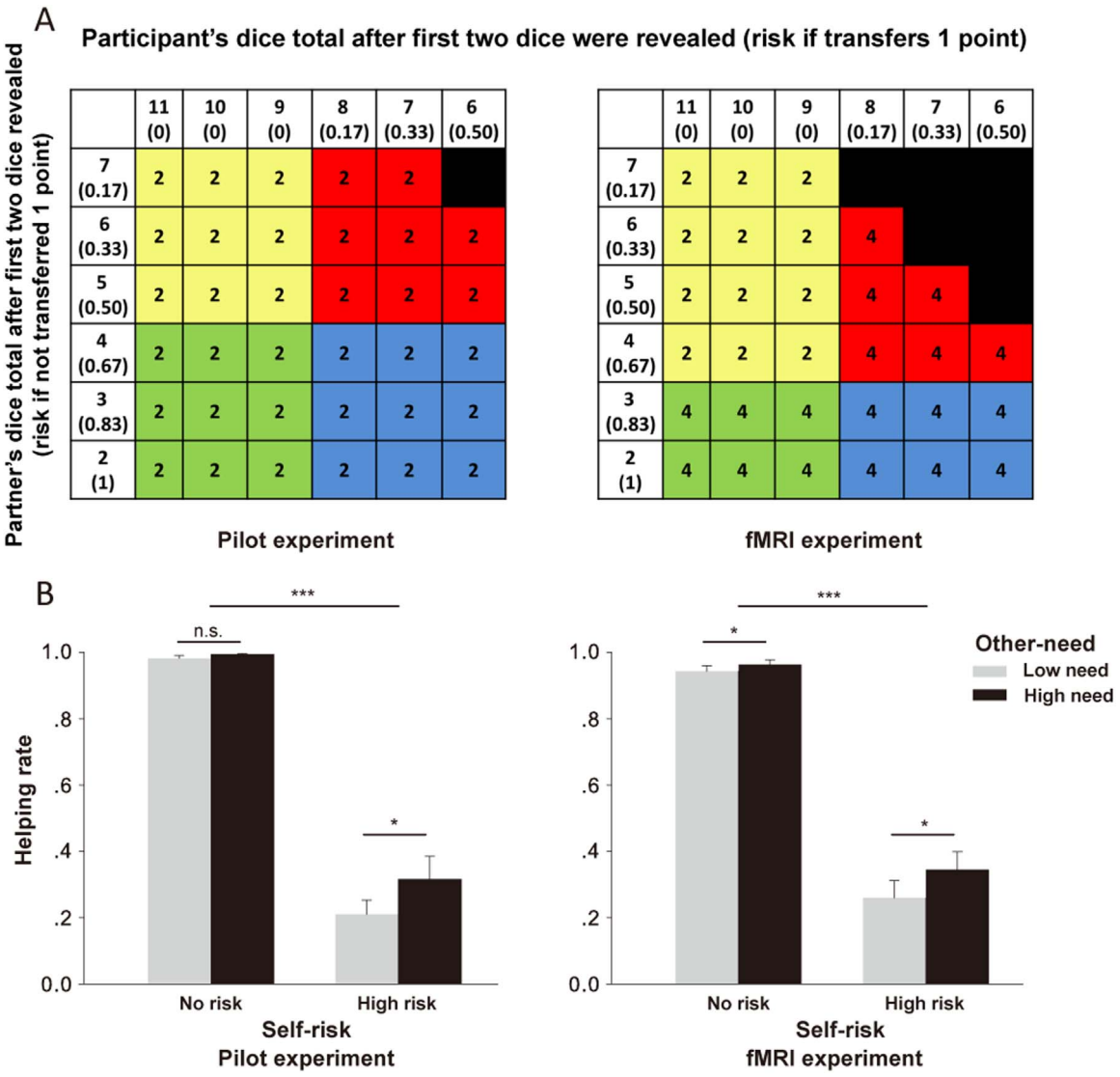

Fig. 2. Trial setup and behavioral results for pilot and fMRI experiment. (A) First row: Top number indicates first-two dice outcome for the participant; bottom number in parentheses indicates probability of receiving punishment if the participant chooses to transfer 1 point. First column: Top number indicates first-two dice outcome for the other player; bottom number in parentheses indicates the probability of receiving punishment if not transferred 1 point. The number in each cell indicates the trial amount in each possible participant-partner dice outcome (e.g., The first yellow cell indicates that there are 2 trials in which the participant rolls 11 with a 0% probability of getting punished if he/she chooses to transfer 1 point and in which the partner for that round rolls a 7 and will have a 17% probability of receiving punishment if the participant does not transfer 1 point). Yellow cells indicate trials in the NS_LO condition; red cells indicate trials in the HS_LO condition; green cells indicate trials in the NS_HO condition; and blue cells indicate trials in the HS_HO condition. Black cells indicate trials which do not exist in the current experiment as these cells would result in the participant having a higher probability of being punished than the other player (pilot experiment, left panel) or the participant having a higher probability of being punished than the other player if the participant transfers 1 point (fMRI experiment, right panel). (B) The helping rate is depicted as a function of self-risk and other-need.

et al., 2007) and the helping rate in each of the four critical conditions. The results showed that there was no significant association between individuals' loss aversion and the helping rate (r ranging from − 0.15 to − 0.21, $ps > 0.351$), suggesting that participants' risk-taking tendency did not influence their altruistic behavior in the current study.

Behavioral data analysis

To examine the hypothesis regarding the behavioral effects of self-risk and other-need, for both the pilot and fMRI experiments, we first examined the effects of self-risk and other-need by performing 2 (self-risk: no risk vs. high risk) × 2 (other-need: low need vs. high need) repeated measures ANOVAs on participants' helping rates. Moreover, for the fMRI experiment, to more precisely scrutinize the factors driving the behavioral effects, we also fitted a set of logistic regression models to each participant's helping decisions (help: 1, not help: 0). Model 1 only included self-risk as predictor, and Model 2 only included other-need as predictor. One could argue that the observed behavioral effects were driven by inequity aversion in the current design. Specifically, the difference between the participant's own probability of being punished and the other player's probability of being punished could have motivated the participant to transfer one point to the partner to reduce the inequality between the two players. Therefore, in Model 3, we included the difference between the probability of the other player being punished and the probability of the participants themselves being punished as predictor. Alternatively, one could argue that the behavioral effects were driven by efficiency concerns (i.e. not transferring the point may help the participant him/herself achieve the target more easily). Therefore, in Model 4, we included the ratio of the other player's probability of being punished to the participant's probability of being punished as predictor. A higher ratio would signal a larger probability for the other player to be punished, relative to the probability for the participant to be punished. This relative probability could reflect participants' concern for efficiency. In Model 5, we

included both self-risk and other-need as predictors. In Model 6, we also included the interaction between self-risk and other-need in addition to self-risk and other-need as predictors. We assessed model evidence by employing the Akaike Information Criterion (AIC; Burnham and Anderson, 2004). We only included the trials in which the amplitude of self-risk was higher than 0 in the logistic regressions; this was done for the following two reasons: 1) there was a categorical difference between no self-risk (i.e., self-risk = 0) and high self-risk (i.e., self-risk > 0) trials; 2) the ANOVAs showed that, when there was no self-risk, the participant showed a strong choice bias (i.e., transferring the point to the other player in about 95% of the trials) and a small variance (i.e., participants were not influenced by the amplitude of other-need), which was different from the behavioral pattern in high self-risk trials.

MRI data acquisition

Imaging data were collected using a GE-MR750 3.0 T scanner with a standard head coil at Peking University, Beijing, China. T2*-weighted echo-planar images (EPI) with blood oxygenation level-dependent (BOLD) contrast were acquired in 35 axial slices parallel to the AC-PC line with an interslice gap of 4 mm, allowing for full-brain coverage. Images were acquired in an interleaved order, with a repetition time (TR) of 2000 ms, an echo time (TE) of 30 ms, a flip angle of 90°, and a field of view (FOV) of 200 mm × 200 mm, and 3 mm × 3 mm × 4 mm voxels.

fMRI preprocessing

Preprocessing of the fMRI images was done using Statistical Parametric mapping software SPM8 (Wellcome Trust Department of Cognitive Neurology, London, UK), which was run through Matlab (Mathworks). For each session, the first five volumes were discarded to allow for stabilization of magnetization. Then, the remaining images were slice-time corrected, motion-corrected, re-sampled to 3 × 3 × 3 isotropic voxel, normalized to Montreal Neurological Institute (MNI) space, and spatially smoothed using an 8-mm FWHM Gaussian filter. Data were filtered using a high-pass filter with 1/128 Hz cutoff frequency. We performed statistical analyses based on general linear model (GLM), first at the participant level and then at the group level.

General linear model analyses

At the neural level, the first question we would like to address is which brain regions processed information or motives regarding self-interest and other-need. Therefore, we first carried out parametric analyses with the level of self-risk and other-need as parametric regressors to identify the brain regions associated with representing the motives of self-interest and other-need.

In the parametric model (GLM 1), we separately modeled the first-two-dice outcome, fixation, partner pairing, and motor response in the critical trials and the first-two-dice outcome, fixation, and partner pairing in the filler trials with boxcar functions spanning the whole event. The regressor of the first-two-dice outcome in critical trials was modulated by the level of self-risk and other-need. The duration for the first-two-dice outcome was equal to the time from the onset of the first-two-dice outcome screen to the time point at which the participant pressed the corresponding button. All regressors of interest and no interest were convolved with a canonical hemodynamics response function (HRF). The six rigid body parameters were also included as regressors of no interest to account for head motion artifacts. The participant-specific estimates of the parametric regressors at each voxel were then fed into a second-level group level one-sample *t*-tests treating participants as a random variable. For the GLM analysis, whole brain results are reported using a voxel-level uncorrected threshold of $p < 0.001$ with cluster-level family wise error (FWE) corrected $p < 0.05$ unless otherwise stated.

Regions showing increased activations with the increase of self-risk and decreased activations with the increase of other-need were identified by parametric analyses. To extract regional activation strength, two more models were built in which the presentation onsets of the first-two-dice outcome corresponding to each self-risk level (self-risk level: 0, 0.17, 0.33, 0.50, GLM 2) and to each other-need level (other-need: 0.17, 0.33, 0.50, 0.67, 0.83, 1.00, GLM 3) were modeled as separate regressors (GLM 2: 4 regressors of interest and GLM 3: 6 regressors of interest). These two models also included the fixation, partner pairing, and motor response in critical trials and the onsets of the first-two-dice outcome, the fixation, and partner pairing in filler trials with boxcar functions spanning the whole event. All regressors of interest and no interest were convolved with a canonical hemodynamics response function (HRF). The six rigid body parameters were also included as regressors of no interest. Note that, GLM 2 and GLM 3 were used to show the activation patterns for different levels of self-risk and other-need; these models did not perform any additional inferential analysis of the data. These supplementary GLM analyses were provided purely to estimate the non-standardized effect sizes (beta weights) for each level of perceived self-risk and other-need.

Effective connectivity analysis

To address whether and how the regions related to self-risk and other-need interact with each other to influence altruistic helping behavior, we investigated the effective connectivity between brain regions based on self-risk and other-need processing. Here we used DCM implemented in SPM8 to build and compare different neural connectivity models. The DCMs were defined by three sets of parameters (Friston et al., 2003): 1) the average connectivity which represents endogenous or extrinsic connectivity between VOIs not influenced by experimental variables; 2) the modulatory connectivity represents the effect of critical experimental conditions on the extrinsic connectivity; and 3) the driving input represents the influence of critical experimental conditions on regions of interest in the model.

In the DCM analyses, we selected the volumes of interest (VOIs) based on the parametric analysis, and extracted the first principal component (or eigenvariate) of the time series from each VOI (3 mm spheres centered on the group-level peak coordinates) in each participant. In accordance with Kiebel and colleagues (Kiebel et al., 2007), we set the slice timings of all the VOIs as the reference slice in the slice-time correction during preprocessing (i.e. the 35th slice, the middle slice in the interleaved order). As most of the regions showed significant associations with self-risk and other-need were located in right hemisphere, we constructed our models with the regions in the right hemisphere. Specifically, three volumes were extracted based on the peaks of the contrast of "positive association with self-risk", "negative association with other-need," and the conjunction of these two contrasts (see *Results* section for parametric results). The first VOI was associated with other-need and located in the right inferior parietal lobe (rIPL: 57, – 37, 52). IPL is involved in number comparison (Dehaene et al., 1999, 2003; Pinel et al., 2004), with greater activations for close relative to far comparisons of numerical magnitude (Pinel et al., 2001; Chiao et al., 2009). Therefore, the negative association between rIPL activity and the magnitude of other-need may reflect the recognition of others' risk to be punished by comparing others' number of points and their ultimate goal (i.e., attaining 9 points). The second VOI was associated with self-risk and located in the medial prefrontal cortex (MPFC: 9, 50, 34). MPFC is suggested to convey risk signals during risky decision tasks (Van Duijvenvoorde et al., 2015; Xue et al., 2009). Thus, the positive association between the MPFC activity and the magnitude of self-risk may reflect the encoding of the participant's probability of being punished. The third VOI was associated with both the self-risk and the other-need and located in rDLPFC (48, 11, 22). DLPFC is thought to serve as an integration-and-selection node which is critical for selecting a context-appropriate response from multiple

response options (Buckholtz and Marois, 2012). The associations between rDLPFC activity and both self-risk and other-need may indicate that this region was responsible for integrating information from both sources and making a final response. Therefore, we assumed that the effects of critical conditions would input from rIPL and MPFC, and then these two regions would convey the information of self-risk and other-need to rDLPFC to influence participants' decisions.

We also performed DCM analyses based on the VOIs defined individually. Specifically, we first took the coordinate of the peak voxel in each region of interest (ROI) identified in the group-level analyses as a landmark, and searched for individual peak voxel which survived the $p < 0.05$ threshold around the landmark within 8-mm radius distance and within the same anatomical regions in each ROI. Then we defined the individual VOI as 3-mm spherical volumes centered at the individual peak voxel. However, we failed to find the peak voxel of rIPL in 1 participant and the peak voxel of rDLPFC in 1 participant even when we lowered the threshold to $p < 0.1$. Therefore, we excluded these 2 participants, and performed DCM analyses based on individual specific VOIs for the remaining 23 participants. The results were largely the same as the results of the DCM analyses based on VOIs defined by group-level peak voxels. Here we report the DCM results based on VOIs defined by group-level peak voxel. Results of the DCM based on individual VOIs are reported in the Supplementary materials.

To clarify how rIPL and MPFC interacted with rDLPFC to process and weigh self-risk and other-need, we built and compared 9 families of models differing in the direction of extrinsic connectivity (bilateral or unilateral), the presence or absence of the extrinsic connectivity between any two of the three regions (i.e. between MPFC and rDLPFC, between rIPL and rDLPFC, or between rIPL and MPFC). In the current study, we included four critical experimental conditions - HS_HO, HS_LO, NS_HO, and NS_LO - as the driving inputs and the modulators to the connectivity between different brain regions. For the bilateral models, we built 4 families (Model Families 1–4) varying in the presence of the extrinsic connectivity between any two of the three regions. For the unilateral models, since we assumed that rDLPFC was an integration region, we built 5 families (Model Families 5–9) with connectivies from rIPL and/or MPFC to rDLPFC, which varied in the presence of the extrinsic connectivity between any two of the three regions. Moreover, since we assumed that rIPL processed information regarding other-need and that MPFC processed information regarding self-risk, in our model space, the inputs of critical conditions were from rIPL and MPFC in all the models. Within each family, models differed in the modulatory effect on the extrinsic connectivity. See Table S2 in the Supplementary materials for specific modulatory effects in each model. The models and model families were then compared at the group level using the Bayesian Model Selection (BMS), a random-effect analysis (i.e. assuming that the model structure might vary across participants) implemented in SPM8. This approach is robust to outliers (Stephan et al., 2009) and estimates the model evidence of each model which accounts for the trade-off between model simplicity and model fitness (Penny et al., 2004). To compare model families, the exceedance probabilities were used. The exceedance probability describes the probability of each model being more likely to stand than any other model. When comparing model families, the exceedance probabilities are calculated for each model family (Penny et al., 2010). Model parameters within the winning family were then assessed using Bayesian model averaging for each participant. The significance of these connectivity estimates at the between-participant level was assessed using classical (one-sample *t*) tests. Although Bayesian model comparison could have been used to assess the contribution of modulatory effects on effective connectivity, *t*-test provides a more intuitive assessment of the effects of interest, in relation to inter-subject variability. Thus, we reported the simpler *t*-tests on the modulatory effects of self-interest and other-need on effective connectivities.

Methods of tDCS experiments

Given that the fMRI experiment found that rDLPFC and rIPL may play critical roles in altruistic behavior, we further examined the causal roles of these two regions in altruistic helping behaviors by conducting two tDCS experiments. In both the rDLPFC and the rIPL experiments, the participants received either cathodal or sham tDCS to rDLPFC (rIPL) while playing the dice game. Fifty-six and fifty-eight different right-handed undergraduate and graduate students participated in the rDLPFC and rIPL tDCS experiments, respectively. Participants were randomly assigned to the cathodal (rDLPFC tDCS: n = 28, 18 females; rIPL tDCS: n = 29, 19 females) or control "sham" (rDLPFC tDCS: n = 28, 18 females; rIPL tDCS: n = 29, 20 females) groups in a double-blind manner, with the participants and the experimenter who introduced the instructions to the participants not knowing who received cathodal or sham stimulation.

High-Definition (HD) Stimulation was delivered by a multi-channel stimulation adapter (SoterixMedical, 4 × 1 - C3, New York) connected to a battery-driven stimulator (SoterixMedical, Model 1300-A, New York). Five Ag-AgCl sintered ring electrodes were connected to the skull with conductive gel, according to the international 10–20 system. To deliver stimulation on rDLPFC, we placed one central electrode on F4 (Sellaro et al., 2016), and four return electrodes on the locations corresponding to C4, FT8, Fp2, and Fz. To deliver stimulation on rIPL, we placed one central electrode on P4 (Cai et al., 2016; Ono et al., 2016), and four return electrodes on the locations corresponding to C4, Pz, O2, and P8. The four return electrodes formed a square and were spaced ~ 7.5 cm radially around the central electrode according to previous HD-tDCS studies (Villamar et al., 2013a, 2013b). Current polarity on the target brain area depended on the central electrode. Participants received a constant current of 2 mA intensity. tDCS started 8 min before the task and was delivered during the whole course of the dice game. For sham stimulation, the electrodes were placed at the same positions as cathodal stimulation, but the stimulator was only turned on for the initial 30 s. This method of sham stimulation has been shown to be reliable (Gandiga et al., 2006; Nihonsugi et al., 2015). For both cathodal and sham groups, participants experienced itchy or painful sensations at the beginning of the stimulation, and such uncomfortable feelings would disappear soon thereafter. At the beginning of the task, 8 min after the start of the stimulation, no participant reported any uncomfortable feelings. In the decision-making literature, cathodal stimulation has been suggested to effectively interrupt neural activity of cortical regions and substantially influence participants' behavior (Knoch et al., 2008; Mengarelli et al., 2015; Ruff et al., 2013). Given that cathodal stimulation increases the membrane potential of the neurons (hyperpolarization) and thus decreases the neuronal firing rate (Utz et al., 2010), the spike activity of the target regions (i.e. rDLPFC and rIPL) would be lower for the cathodal group relative to the sham group. The tDCS task was the same as the fMRI task, with the exception that the tDCS task included 96 critical trials and only 48 filler trials in total and lasted for ~ 18 min.

After the dice game, we asked the participants to perform non-social tasks (i.e. number comparison task and working memory task) to examine the effects of tDCS stimulations on individuals' cognitive abilities. To examine whether interrupting rDLPFC would influence participants' affect or working memory ability, we measured participants' affective states and working memory performance both before and after tDCS stimulation over rDLPFC. Affective states were assessed using the Positive and Negative Affect Schedule (PANAS; Watson et al., 1988) which included 10 kinds of positive affect and 10 kinds of negative affect; working memory capacity was assessed using a computerized Digit Span task (Wechsler Adult Intelligent Scale-Third Edition: WAIS-III, Wechsler, 1997; Lefebvre et al., 2005; Mackey et al., 2016). Working memory performance was reported as the maximum set size that the participant was able to recall correctly in each block.

To examine whether interrupting rIPL would influence participants' number processing ability, we included a number comparison task during the tDCS stimulation over rIPL. In this task, the participant was presented with a pair of double-digit numbers in each trial. They were

instructed to compare the magnitude of these two numbers and to indicate which number had a higher magnitude by pressing the corresponding key as soon and as accurately as possible. The numerical distance between the numbers in each pair was defined as close when the difference between the two numbers ranged from 1 to 3 (i.e. 37 vs. 38), as medium when the difference ranged from 4 to 6 (i.e. 12 vs. 18), and as far when the difference ranged from 7 to 9 (i.e. 40 vs. 48). According to the numerical distance effect, comparing numbers with a close distance between them should entail a longer response time than when comparing numbers with a far distance between them, and such an effect was tightly associated with the activity in bilateral parietal cortices (Pinel et al., 2001, 2004; Chiao et al., 2009).

Results

Behavioral results

Pilot experiment

A 2 (self-risk: no risk vs. high risk) × 2 (other-need: low need vs. high need) repeated measures analysis of variance (ANOVA) on participants' helping rates revealed a significant main effect of self-risk, $F(1, 20) = 172.70$, $p < 0.001$, η^2partial = 0.90, with the helping rate (mean ± SE) being higher for the no self-risk conditions (0.99 ± 0.004) than for the high self-risk conditions (0.26 ± 0.06). There was also a significant main effect of other-need, $F(1, 20) = 7.60$, $p = 0.012$, η^2partial = 0.28, indicating that participants' helping rate was higher for the high other-need conditions (0.65 ± 0.04) than for the low other-need conditions (0.60 ± 0.02). Importantly, there was a significant interaction between self-risk and other-need, $F(1, 20) = 6.08$, $p = 0.023$, η^2partial = 0.23. Tests for simple effects (Fig. 2B left panel) showed that when self-risk was high, the helping rate was higher when other-need was high (0.32 ± 0.07) than when other-need was low (0.21 ± 0.04), $p = 0.014$, η^2partial = 0.27; this effect was absent when self-risk was low (0.99 ± 0.004 vs. 0.98 ± 0.01), $p = 0.285$, η^2partial = 0.06.

fMRI experiment

For the fMRI experiment, we also conducted an ANOVA on the behavioral data at the group level. A 2 (self-risk: no risk vs. high risk) × 2 (other-need: low need vs. high need) repeated measures ANOVA on participants' helping rates revealed a significant main effect of self-risk, $F(1, 24) = 173.79$, $p < 0.001$, η^2partial = 0.88, with the helping rate (mean ± SE) being higher for the no self-risk conditions (0.95 ± 0.02) than for the high self-risk conditions (0.30 ± 0.05). There was also a significant main effect of other-need, $F(1, 24) = 7.53$, $p = 0.011$, η^2partial = 0.24, indicating that participants' helping rate was higher for the high other-need conditions (0.65 ± 0.03) than for the low other-need conditions (0.60 ± 0.03). Importantly, there was a significant interaction between self-risk and other-need, $F(1, 24) = 4.63$, $p = 0.042$, η^2partial = 0.16. Tests for simple effects (Fig. 2B right panel) showed that when self-risk was high, the helping rate was higher when other-need was high (0.35 ± 0.05) than when other-need was low (0.26 ± 0.05), $p = 0.017$, η^2partial = 0.22; this effect was smaller when self-risk was low (0.96 ± 0.01 vs. 0.94 ± 0.02), $p = 0.044$, η^2partial = 0.16. Although the trial distribution in the fMRI experiment was somewhat different from that in the pilot experiment, ANOVA revealed essentially the same pattern of behavioral effects. It is clear that the pattern of our results does not depend on the exact grouping of self-risk and other-need trials.

Next, we fitted and compared a set of logistic regression models to each participant's helping to further examine the factors (such as self-risk, other-need, inequity aversion, and efficiency) driving the behavioral effects. Table 1 summarizes the evidence for all of the models. The model with the lowest AIC is considered to be the best at explaining behavioral effects. As shown in Table 1, Model 5 had the lowest AIC. Thus, the model with self-risk and other-need (Model 5) is better than other models (i.e. the inequity aversion model (Model 3), the efficiency model (Model 4), and the interaction model (Model 6)) in explaining participants' helping behavior.

Table 1
Quality of model fits.

Models	Predictors	Number of parameters	AIC
1	Self-risk (β_{self})	1	879
2	Other-need (β_{other})	1	1021
3	Self-risk − other-need (β_{diff})	1	953
4	Other-need/(Self-risk + 1/6) (β_{ratio})	1	926
5	Self-risk (β_{self}), other-need (β_{other})	2	764
6	Self-risk (β_{self}), other-need (β_{other}), self-risk * other-need ($\beta_{interaction}$)	3	782

We also conducted logistic regression analyses with all the trials, and the model comparison showed similar results, with the exception that the model with the interaction term (Model 6, AIC = 1224) has a lower AIC than the model without the interaction term (Model 5, AIC = 1320). However, given that the interaction between self-risk and other-need in the Model 6 would cause a problem of multicollinearity in regression analyses, we only included self-risk and other-need as parametric modulators in the GLM analyses of fMRI data.

fMRI results

Parametric results

In the GLM 1, we included self-risk and other-need as parametric modulators to the presentation of the first-two-dice outcome, and found significant positive correlations of activity with the magnitude of self-risk in rDLPFC and MPFC (Table 2 and Fig. 3). No region showed a significant negative correlation with self-risk. We also found significant negative correlations of activity with the magnitude of other-need in rDLPFC, the right inferior parietal lobe (IPL), the right middle frontal gyrus (MFG), and bilateral middle occipital gyrus (MOG; Table 2 and Fig. 3). No region showed a significant positive correlation with other-need. A whole-brain conjunction analysis showed that rDLPFC was the only region in which the activity significantly correlated with both self-risk and other-need at a more liberal threshold voxel-wise $p < 0.001$ uncorrected with a minimum cluster extent of 60 voxels (Table 2 and Fig. 3). These findings suggested that processing of self-risk and other-need may involve both distinct and overlapping brain regions.

In addition, we conducted regions of interest (ROIs) parametric

Table 2
Brain activation in the parametric contrast (GLM1, voxel-level uncorrected $p < 0.001$ and cluster-level FWE corrected $p < 0.05$).

Regions	Laterality	Peak MNI coordinates x	y	z	Max *T*-value	Cluster size (k)
Positive association with self-risk						
DLPFC	R	45	11	25	5.82	473
MPFC	L, R	9	50	34	4.21	339
Negative association with other-need						
DLPFC	R	51	11	22	5.81	121
IPL	R	57	− 37	52	4.69	126
MFG	R	30	26	55	4.55	140
MOG	R	42	− 73	34	4.38	153
	L	− 30	− 67	37	4.70	160
Conjunction: "Positive association with self-risk" and "Negative association with other-need" (voxel-wise $p < 0.001$ uncorrected with a minimum cluster extent of 60 voxels)						
DLPFC	R	48	11	22	5.11	82

MNI-coordinates are reported for peak activation.
R, right; L, left. DLPFC = dorsolateral prefrontal cortex, MPFC = medial prefrontal cortex, IPL = inferior parietal lobe, MFG = middle frontal gyrus, MOG = middle occipital gyrus.

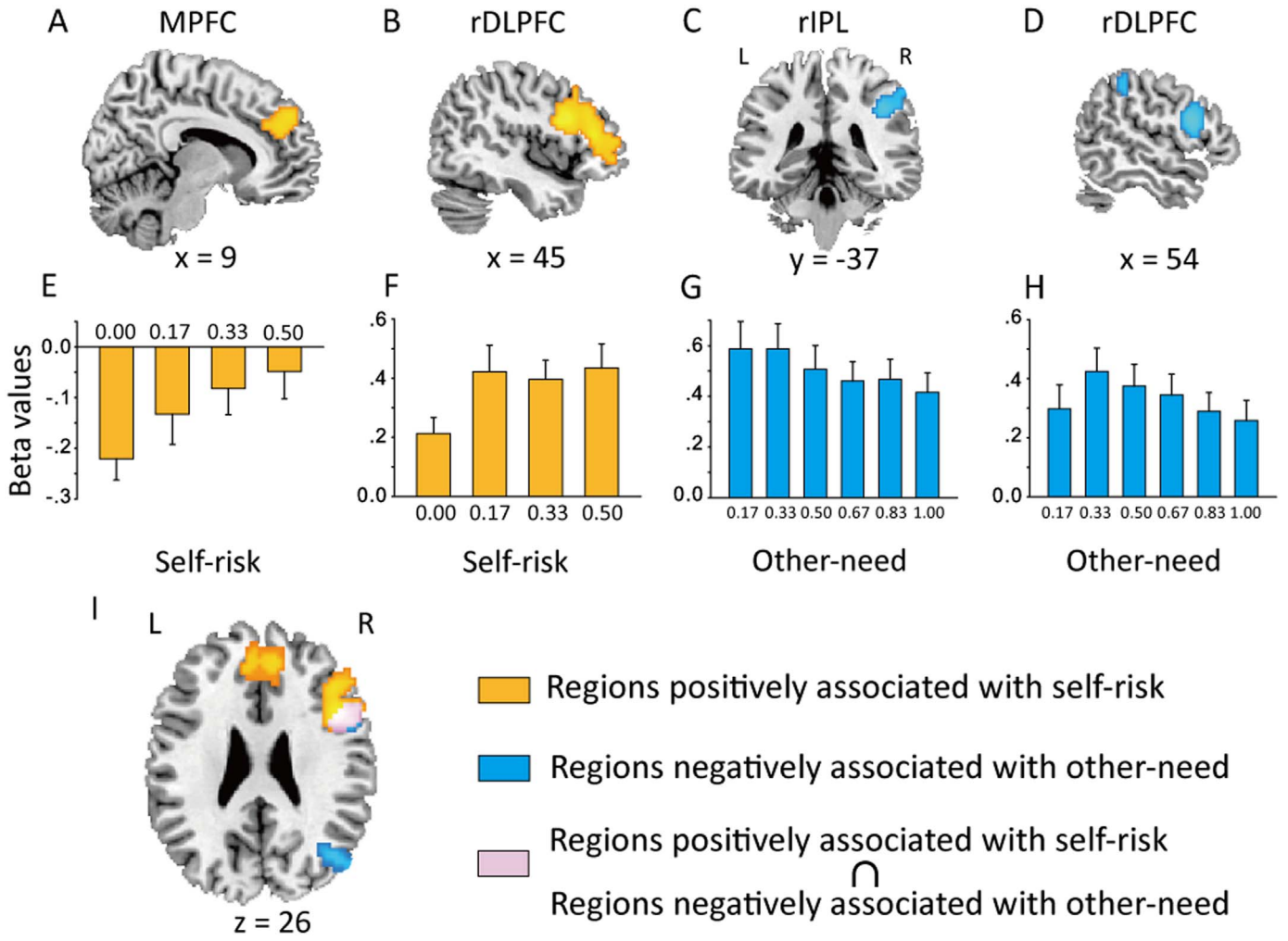

Fig. 3. Parametric analysis results. MPFC (**A**) and rDLPFC (**B**) showed positive associations with the magnitude of self-risk. rIPL (**C**) and rDLPFC (**D**) showed negative associations with the magnitude of other-need. Beta values corresponding to 4 levels of self-risk were extracted from MPFC (**E**, based on GLM 2) and rDLPFC (**F**, based on GLM 2), and beta values corresponding to 6 levels of other-need were extracted from rIPL (**G**, based on GLM 3) and rDLPFC (**H**, based on GLM 3). To avoid double-dipping the data, we merely show the data pattern here (**E-H**), and do not subject the beta values to further statistical analysis. The beta values were the averaged beta values across the voxels in a spherical regions within 3 mm radius and centered at the peak coordinate of the activation regions (MPFC: 9, 50, 34; rIPL: 57, – 37, 52; rDLPFC: 48, 11, 22). (**I**) Conjunction analysis showed that rDLPFC (violet voxels) was the only region both positively correlated with the magnitude of self-risk and negatively correlated with the magnitude of other-need. Activations were thresholded at voxel-wise $p < 0.001$ uncorrected and cluster-wise family wise error (FWE) corrected $p < 0.05$, with the exception of the conjunction analysis, in which the activation was thresholded at a more liberal threshold voxel-wise $p < 0.001$ uncorrected with a minimum cluster extent of 60 voxels. MPFC, medial prefrontal cortex; rDLPFC, right dorsolateral prefrontal cortex; rIPL, right inferior parietal lobe.

analyses using small volume correction with radius of 8 mm to explore whether other relevant regions from previous literature were also involved in representing self-risk and other-need in the current task. The self-interest related ROIs included right VTA (4, – 12, – 8; Moll et al., 2006), left ventral striatum (VS: – 2, 5, – 2; Moll et al., 2006), anterior medial prefrontal cortex (aMPFC: 0, 53, – 2; Moll et al., 2006), ventral medial prefrontal cortex (vMPFC: – 5, 39, – 3, Moll et al., 2006), sgACC (2, 28, – 6, FeldmanHall et al., 2015), and bilateral caudate/putamen (– 16, 12, 0; 16, – 20, 6; Hu et al., 2015). None of these ROIs showed a significant association with the level of self-risk. The other-regarding related ROIs included bilateral anterior insula (AI: – 36, 18, 1; 36, 29, – 8; Hein et al., 2010) and right temporoparietal junction (TPJ: 52, – 40, 4; FeldmanHall et al., 2015). Again, none of these ROIs showed a significant association with the level of other-need.

Note that we also conducted factorial analyses of the fMRI data, modeled after the behavioral data analyses. The pattern of effects was essentially the same as we reported above for the parametric analyses (see Supplementary materials).

Effective connectivity results

Fig. 4A summarizes the structures of the models. Each of the 9 structures formed a model family, with each family containing three to eight models differing in the modulatory effect, resulting in 34 models in total. Model comparison results suggested that the winning DCM model family had unilateral extrinsic connectivity from IPL and MPFC towards DLPFC, and its input was from IPL and MPFC, as shown in Fig. 5. One-sample *t*-tests showed that the extrinsic connectivity from MPFC to rDLPFC (0.11 ± 0.02, $t(24) = 4.45$, $p < 0.001$) and from rIPL to rDLPFC (0.19 ± 0.04, $t(24) = 5.00$, $p < 0.001$) were significantly positive (Fig. 5A).

Crucially, we examined the modulatory effects of different conditions on functional connectivity between regions. The 2 (self-risk: no risk vs. high risk) × 2 (other-need: low need vs. high need) repeated measures ANOVA revealed no significant main effect of either self-risk or other-need on the modulatory connectivity from rIPL to rDLPFC (*p*s > 0.082), but there was a significant main effect of self-risk on the modulatory connectivity from MPFC to rDLPFC, $F(1, 24) = 10.19$, $p = 0.004$, η^2partial = 0.30. The connectivity from MPFC to

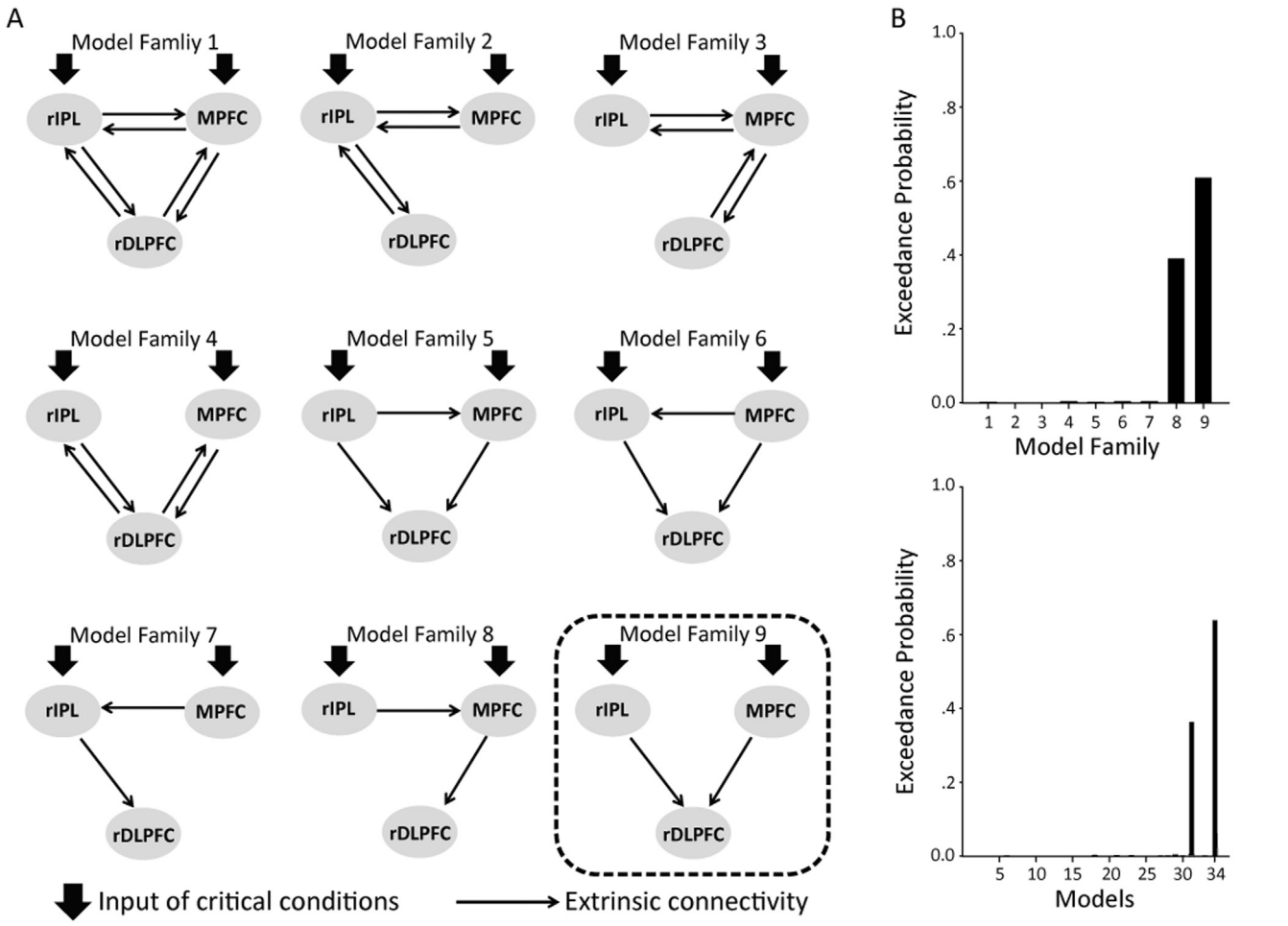

Fig. 4. The DCM analysis for the network consisting of rIPL (peak MNI: 57, – 37, 52), MPFC (peak MNI: 9, 50, 34), and rDLPFC (peak MNI: 48, 11, 22). (A) The structures of 9 model families (right hemisphere). Each model family contained three to eight models differing in the specific pathway(s) that was (were) modulated by the critical conditions (NS_LO, NS_HO, HS_LO, HS_HO). The structure of the winning family is highlighted in the dashed box. (B) The exceedance probabilities of model families (upper panel) and individual models (lower panel). MPFC, medial prefrontal cortex; rDLPFC, right dorsolateral prefrontal cortex; rIPL, right inferior parietal lobe.

rDLPFC was more strongly modulated by high self-risk (0.21 ± 0.04) than no self-risk (0.04 ± 0.03; see Fig. 5B). Moreover, the modulatory effect of other-need ((HS_HO – HS_LO) – (NS_HO – NS_LO)) on the connectivity from rIPL to rDLPFC positively correlated with the modulatory effect of other-need ((HS_HO – HS_LO) – (NS_HO – NS_LO)) on individuals' helping rate, r = 0.41, p = 0.044 (Fig. 5C), suggesting that when participants had to take on risk to help others, enhanced connectivity from rIPL to rDLPFC was associated with increased helping rate in the high other-need condition relative to the low other-need condition. That is, the effective connectivity from rIPL to rDLPFC may provide the neural underpinnings for individual differences in other-regarding altruistic tendencies. This finding provides construct validity for the effective connectivity estimates in DCM in terms of behavioral phenotypes.

tDCS results

tDCS over rDLPFC

As with the fMRI experiment, we conducted ANOVAs for the behavioral data. Fig. 6A displays the helping rate as a function of tDCS stimulation, self-risk, and other-need. We conducted a 2 (tDCS stimulation type: cathodal vs. sham) × 2 (self-risk: no risk vs. high risk) × 2 (other-need: low need vs. high need) repeated measures ANOVA on participants' helping rate in different conditions, with the tDCS stimulation as a between-participant factor. Consistent with the fMRI experiment, this analysis revealed a significant main effect of self-risk, $F(1, 54) = 453.39$, $p < 0.001$, η^2partial = 0.89, and a significant interaction between self-risk and other-need, $F(1, 54) = 26.49$, $p < 0.001$, $\eta^2 partial = 0.33$. Importantly, we found a significant interaction between tDCS stimulation type and self-risk, $F(1, 54) = 9.64$, $p = 0.003$, $\eta^2 partial = 0.15$, and a significant interaction between tDCS stimulation type and other-need, $F(1, 54) = 5.51$, $p = 0.023$, $\eta^2 partial = 0.09$. Simple effect analysis, on the one hand, showed that when self-risk was high, the helping rate was lower for the cathodal group (0.19 ± 0.05) than for the sham group (0.35 ± 0.05; $p = 0.014$, $\eta^2 partial = 0.11$); when there was no self-risk, there was no difference in helping rate between the cathodal group (0.95 ± 0.02) and the sham group (0.92 ± 0.02, $p = 0.254$, $\eta^2 partial = 0.02$). On the other hand, for the sham group, the helping rate was higher when other-need was high (0.69 ± 0.03) than when other-need was low (0.59 ± 0.03; $p < 0.001$, $\eta^2 partial = 0.38$). This effect was smaller for the cathodal group (high need, 0.59 ± 0.03 vs. low need, 0.55 ± 0.03; $p = 0.019$, $\eta^2 partial = 0.097$). There was neither a main effect of tDCS stimulation type nor a three-way interaction between tDCS stimulation type, self-risk, and other-need, $ps > 0.100$. These results suggested that rDLPFC causally affected both the effects of self-risk and other-need on participants' helping rates.

For the working memory test, a 2 (tDCS stimulation type: cathodal vs. sham) × 2 (test period: pre-test vs. post-test) × 2 (recall order: forward vs. backward) repeated measures ANOVA on participants' digit span performance revealed that there was neither a main effect of tDCS stimulation type nor an interaction between tDCS stimulation type and

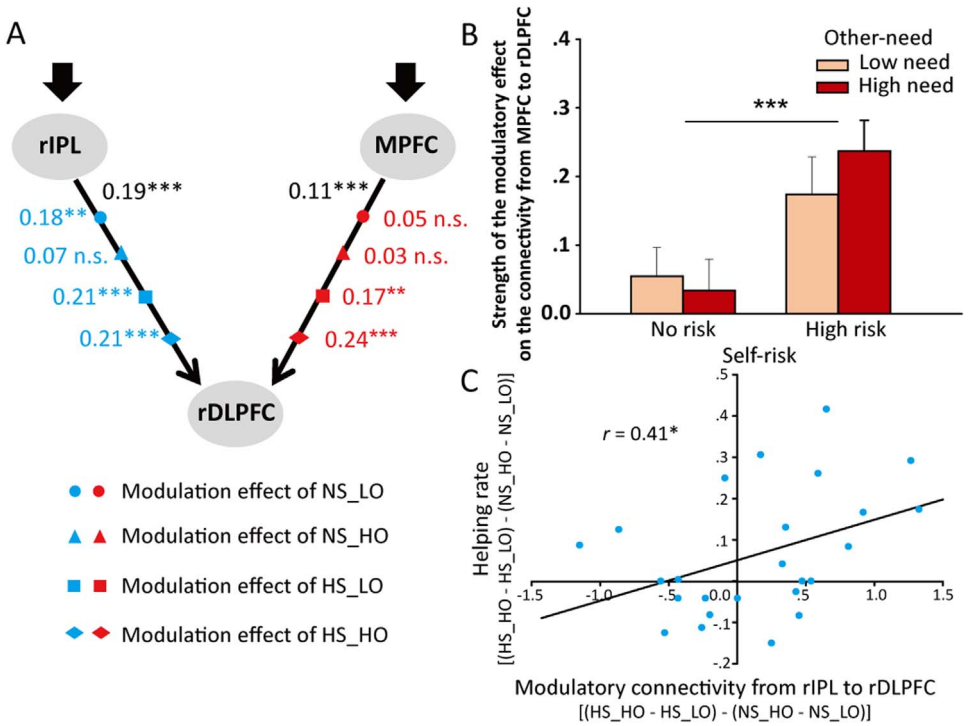

Fig. 5. DCM analysis results. (**A**) The estimated DCM parameters of the average model of the winning family (* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$, # $p < 0.1$). The numbers shown in black indicate the strength of the extrinsic connectivity, and the numbers shown in color indicate the strength of the modulatory effects of the experimental critical conditions on the connectivity from rIPL to rDLPFC (blue) and from MPFC to rDLPFC (red). Different shapes represent different conditions, with circle indicating NS_LO, triangle indicating NS_HO, square indicating HS_LO, and diamond indicating HS_HO. (**B**) The strength of the modulatory effect on the connectivity from MPFC to rDLPFC is depicted as a function of self-risk and other-need. (**C**) The modulatory effect of other-need ((HS_HO – HS_LO) – (NS_HO – NS_LO)) on the connectivity from rIPL to rDLPFC positively correlated with the modulatory effect of other-need ((HS_HO – HS_LO) – (NS_HO – NS_LO)) on individuals' helping rate.

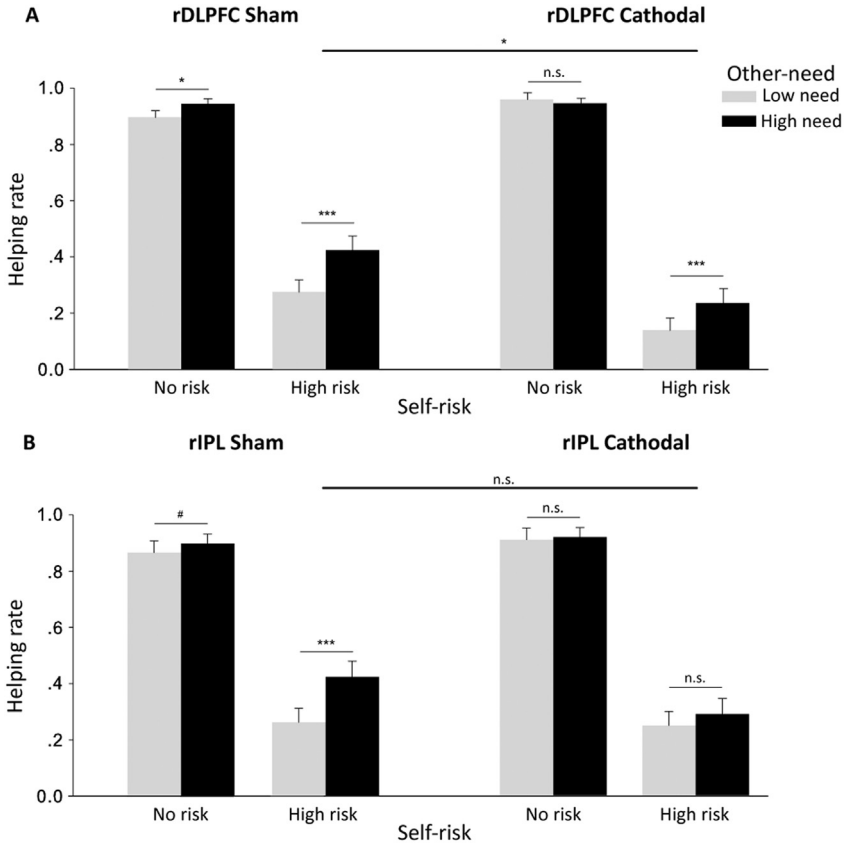

Fig. 6. tDCS results. The helping rate depicted as a function of self-risk, other-need, and tDCS stimulation type in the rDLPFC tDCS experiment (**A**) and in the rIPL tDCS experiment (**B**). * $p < 0.05$, *** $p < 0.001$, # $p < 0.1$, n.s. not significant.

Table 3
Participants' performance in Digit Span task and ratings in PANAS.

	Sham		Cathodal	
	Pre-test (Mean ± SE)	Post-test (Mean ± SE)	Pre-test (Mean ± SE)	Post-test (Mean ± SE)
WM performance: maximum set size in Digit Span task				
Forward recall	7.96 ± 0.34	8.41 ± 0.32	8.43 ± 0.34	8.66 ± 0.32
Backward recall	6.75 ± 0.33	7.12 ± 0.31	6.57 ± 0.33	7.26 ± 0.31
Affective state: ratings on PANAS				
Positive affect	2.92 ± 0.11	2.71 ± 0.13	3.02 ± 0.11	2.75 ± 0.13
Negative affect	1.85 ± 0.10	1.59 ± 0.09	1.87 ± 0.10	1.58 ± 0.09

other factors (i.e. test period and recall order), $ps > 0.327$. For the affective ratings, a 2 (tDCS stimulation type: cathodal vs. sham) × 2 (test period: pre-test vs. post-test) × 2 (affective valence: positive vs. negative) repeated measures ANOVA revealed neither a significant main effect of tDCS stimulation type nor an interaction between tDCS stimulation type and other factors, $ps > 0.659$. For descriptive statistics, see Table 3.

tDCS over rIPL

Fig. 6B displays the helping rate as a function of tDCS stimulation, self-risk, and other-need. A 2 (tDCS stimulation type: cathodal vs. sham) × 2 (self-risk: no risk vs. high risk) × 2 (other-need: low need vs. high need) repeated measures ANOVA with the tDCS stimulation as a between-participant factor revealed a significant main effect of self-risk, $F(1, 56) = 295.97$, $p < 0.001$, $\eta^2 partial = 0.84$, a significant main effect of other-need, $F(1, 56) = 19.14$, $p < 0.001$, $\eta^2 partial = 0.26$, and a significant interaction between self-risk and other-need, $F(1, 56) = 8.17$, $p = 0.006$, $\eta^2 partial = 0.13$. Importantly, we found a significant interaction between tDCS stimulation type and other-need on helping rate, $F(1, 56) = 6.50$, $p = 0.014$, $\eta^2 partial = 0.10$. Test of simple effects showed that in the sham group, the helping rate was higher when other-need was high (0.66 ± 0.04) than when other-need was low (0.56 ± 0.04), $p < 0.001$, $\eta^2 partial = 0.3$; this effect was absent in the cathodal group (0.61 ± 0.04 vs. 0.58 ± 0.04, $p = 0.202$, $\eta^2 partial = 0.029$). There was no main effect of tDCS stimulation type, no interactions between tDCS stimulation type and self-risk, and no three-way interaction between tDCS stimulation type, self-risk, and other-need on the helping rate, $ps > 0.101$. These results suggested that rIPL only causally affects the effect of other-need on participants' helping rates.

For the number comparison task, our findings replicated the classical numerical distance effect and showed that tDCS did not influence participants' average response times in the number comparison task and did not influence the numerical distance effect. Specifically, A 2 (tDCS stimulation type: cathodal vs. sham) × 3 (numerical distance: close vs. medium vs. far) repeated measures ANOVA on participants' accuracy rate in the number comparison task revealed a significant main effect of numerical distance, $F(2, 112) = 21.67$, $p < 0.001$, $\eta^2 partial = 0.23$, with accuracy rate being highest for the far distance condition (0.992 ± 0.003), intermediate for the medium distance condition (0.985 ± 0.003 ms), and lowest for the close distance condition (0.963 ± 0.005). The differences between conditions were all significant, $ps < 0.05$. There was neither a significant main effect of tDCS stimulation type nor an interaction between the type and numerical distance on accuracy rate, $ps > 0.1$.

The same pattern was observed on response times. We excluded the missing trials and calculated participants' mean response times for each condition. ANOVA results showed that there was a significant main effect of numerical distance, $F(2, 112) = 120.39$, $p < 0.001$, $\eta^2 partial = 0.68$, with response times being longest for the close distance condition (780 ± 16 ms), intermediate for the medium distance condition (726 ± 14 ms), and lowest for the far distance condition (686 ± 15 ms). The differences between conditions were all significant, $ps < 0.001$. No other effects were found.

Taken together, the results from the number comparison task and the dice game appeared to suggest that interrupting rIPL activity may substantially modulate participants' other-regarding tendencies but may have little effect on basic number comparison processing.

Discussion

Combining a novel dice game with fMRI and tDCS, we investigated the neural processing of self-risk and other-need and their bearing on altruistic helping behavior in an interactive context. Participants were less likely to help others in the high self-risk conditions than in the no self-risk conditions and were more likely to help others when other-need was high than when other-need was low. At the neural level, we observed that the processing of self-risk and other-need involved both distinct and overlapping brain regions, with MPFC associating with the level of self-risk, rIPL associating with the level of other-need, and rDLPFC associating with both self-risk and other-need levels. Importantly, MPFC, rIPL, and rDLPFC functioned as a network to influence individuals' altruistic behavior. Moreover, tDCS experiments in which rDLPFC and rIPL were interrupted provided causal evidence that these two regions played distinct roles in the activation of altruistic motives in helping behavior.

Our behavioral results provided robust evidence to support the hypothesis that altruistic helping is a reflective response interactively influenced by the interests of the helper and the need of the recipient (Batson et al., 1983; Heinsohn and Legge, 1999). In the current study, the participants were uncertain of whether their helping decision would actually help the recipients; nevertheless, their helping behaviors were substantially modulated by both the level of self-risk and the level of other-need. On the one hand, individuals' concerns for self-interest decreased their concerns for others' welfare, causing them to behave in a more self-interested manner (Isen and Simmonds, 1978). On the other hand, although high self-risk reduced helping behavior, greater other-need increased participants' helping rates, even when the participants themselves were in danger of being punished.

At the neural level, MPFC was involved in self-risk processing, and the level of self-risk modulated the effective connectivity between MPFC and rDLPFC. The involvement of MPFC in risk processing has been shown in several lines of research. Firstly, MPFC lesions lead to deficits in adaptive decision making under risk (Fellows and Farah, 2005). Secondly, fMRI studies show that making risky decisions, compared with safe decisions, generates stronger activation in the MPFC (Matthews et al., 2004). Thirdly, activity in MPFC is positively modulated by experienced risk, and activity in this region also reflects individual differences in risk preference (Xue et al., 2009). Moreover, here the effective connectivity from MPFC to rDLPFC was more strongly modulated by high self-risk than by no self-risk, indicating that MPFC may not only encode self-risk but also convey this signal to rDLPFC, triggering the latter to modulate the concerns for self-risk and other-need. In line with our findings, recent studies highlighted the role of the functional connectivity between MPFC and DLPFC in value computation, which is critical for goal-directed and economic decision making (Baumgartner et al., 2011). In the current setup, the increased effective connectivity from MPFC to rDLPFC may reflect stronger motives to keep the participants themselves safe. Thus, it is likely that MPFC serves as a critical region for the perception of self-risk in an interactive context and functions with DLPFC to reflect self-interest motives.

In contrast, in the current study, rIPL was involved in other-need processing. It is likely that the participants in the current study recognized others' need for help by comparing others' number of points with the ultimate goal (9 points). Given that rIPL activity was greater for comparison of numbers with close numerical distance than

for comparison with far numerical distance (Pinel et al., 2001, 2004), the enhanced/reduced rIPL activity for lower/higher other-need may reflect the calculation and recognition of others' need for help. Moreover, the DCM analysis revealed that the modulatory effect of other-need on the connectivity from rIPL to rDLPFC positively correlated with the modulatory effect of other-need on individuals' helping rate, suggesting that the efficient connectivity between these two regions may underpin individuals' other-regarding tendencies. Consistently, both the structural connection and the functional coupling between the inferior parietal cortex and the lateral prefrontal cortex have been found to support high-level cognitive processes, such as attentional control and empathy (Caspers et al., 2011; Mars et al., 2012; Wang et al., 2015). Although we observed that tDCS stimulation over rIPL had no effect on participants' number comparison performance, we cannot rule out the possibility that rIPL is engaged in calculation of other-need in the current study. One possible explanation for not observing an effect of tDCS on participants' number comparison performance was that comparing the values of two numbers in the current task was much easier and involved relatively coarse processing than calculating and weighing other's risk of being punished in the dice game, and the relative coarse number comparison ability remained intact after tDCS stimulation. Therefore, based on the fMRI results, we speculated that rIPL may recognize other-need via arithmetic computation or value processing and function in concert with rDLPFC to influence individuals' altruistic concerns (e.g. sympathy and empathic concern) in an interactive game.

Importantly, we further clarify the functions of rDLPFC and rIPL in altruistic behavior by distinguishing their causal roles using tDCS. On the one hand, given that the modulatory effect of other-need on effective connectivity from rIPL to rDLPFC was positively correlated with the modulatory effect of other-need on helping rates, tDCS cathodal stimulation over rDLPFC and rIPL may disrupt the interaction between these two regions and lead the participants to show less concern for other-need when deciding whether or not to help others. On the other hand, these two regions manifested distinct roles in modulating self-interest motives, with rDLPFC functioning to suppress the effect of self-interest, and with rIPL showing no influence on the effect of self-interest. Therefore, we argue that both rIPL and rDLPFC are necessary for individuals' altruistic motives, with rIPL selectively processing the level of other-need and signaling the information about other-need to rDLPFC, and with rDLPFC weighing and modulating the effect of other-need against the effect of self-interest before a final decision is made.

It is crucial to examine the role of rDLPFC in altruistic behavior in broader interactive contexts. DLPFC is generally implicated as a node integrating inputs from different resources and selecting an appropriate response from potential options (Miller and Cohen, 2001; Buckholtz and Marois, 2012). In neuroeconomic literature, DLPFC, especially rDLPFC, has been implicated in inhibiting selfish motives to promote norm compliance and altruistic behavior (Knoch et al., 2006; Ruff et al., 2013; Zhu et al., 2014) or modulating the effect of altruistic motives, rather than simply suppressing selfish motives (Nihonsugi et al., 2015). Taken together, previous findings and the current results suggest that rDLPFC, as part of a network involving at least rIPL and MPFC, integrates both the self-interest and other-regarding information and modulates the relative effects of these two aspects of motives while deciding whether or not to behave in an altruistic manner (Knoch et al., 2006; Ruff et al., 2013; Nihonsugi et al., 2015). Such an integration and modulation mechanism is fundamental to the reproduction/survival of animals and welfare of human beings both at the individual and the group/community levels, in that individuals have to first evaluate their own potential risk or cost and others' need, and then properly weigh the two dimensions of information to make a decision that maximizes the overall benefits or subjective utility (Heinsohn and Cockburn, 1994; Marquez et al., 2015; for a review, see Heinsohn and Legge (1999)).

The specific processes in our task may explain why we did not find other regions usually reported for processing self-interest and other-regarding motives during altruistic behaviors, such as VTA, caudate, and VS for self-interest motives, and AI and TPJ for other-regarding motives. On the one hand, the reward-related regions (e.g. VTA, caudate, and VS) identified in previous studies may reflect the experiences of reward and satisfaction derived from helping others (Moll et al., 2006; FeldmanHall et al., 2015; Hu et al., 2015). However, the self-interest motives in the current study were associated with self-risk, with a larger risk evoking stronger motives to avoid the punishments. Therefore, we observed activity in regions related with risk processing, rather than reward processing. On the other hand, the level of other-need in the current study was indexed by the strangers' probability of being punished, and the punishment was not delivered online. Thus, the stimuli in our study may not be as emotionally arousing as those in studies implicating empathy-related regions during altruistic behavior (e.g. potential electric shock delivered to in-group members or videos of suffering victims; Hein et al., 2010; FeldmanHall et al., 2015). Our task may involve more abstract representations of self-interest and other-regarding motives. Future studies should explore the brain networks of helping behaviors in other circumstances with different task features (e.g. money vs. pain), which would broaden our understanding of the neural mechanisms underlying the processing of self-interest and other-regarding motives during altruistic decision-making.

To conclude, by combining a novel interactive game with fMRI and tDCS techniques, we provided robust empirical evidence that both the helper's self-interest and the recipient's need for help affect the helper's altruistic behavior (Batson et al., 1983; Heinsohn and Legge, 1999). Neuroimaging results showed that MPFC, rIPL, and rDLPFC function as a neural network to support the altruistic decision making, and tDCS results further provided causal evidence that both rIPL and rDLPFC are necessary for individuals' other-regarding considerations. These findings not only shed light on the neural mechanisms of altruistic behavior, but may also have broader implications for understanding the neural deficits in individuals with autism, apathia, and antisocial personality disorder.

Acknowledgments

This work was supported by Natural Science Foundation of China (Grant 31630034) and the National Basic Research Program of China (973 Program: 2015CB856400).

Appendix A. Supplementary material

Supplementary data associated with this article can be found in the online version at doi:10.1016/j.neuroimage.2017.06.040.

References

Batson, C.D., O'Quin, K., Fultz, J., Vanderplas, M., Isen, A.M., 1983. Influence of self-reported distress and empathy on egoistic versus altruistic motivation to help. J. Pers. Soc. Psychol. 45 (3), 706–718.

Batson, C.D., Shaw, L.L., 1991. Evidence for altruism: toward a pluralism of prosocial motives. Psychol. Inq. 2 (2), 107–122.

Batson, C.D., Eklund, J.H., Chermok, V.L., Hoyt, J.L., Ortiz, B.G., 2007. An additional antecedent of empathic concern: valuing the welfare of the person in need. J. Pers. Soc. Psychol. 93 (1), 65–74.

Baumgartner, T., Knoch, D., Hotz, P., Eisenegger, C., Fehr, E., 2011. Dorsolateral and ventromedial prefrontal cortex orchestrate normative choice. Nat. Neurosci. 14 (11), 1468–1474.

Burnham, K.P., Anderson, D.R., 2004. Multimodel inference understanding AIC and BIC in model selection. Sociol. Methods Res. 33, 261–304.

Bode, N.W.F., Miller, J., O'Gorman, R., Codling, E.A., 2015. Increased costs reduce reciprocal helping behaviour of humans in a virtual evacuation experiment. Sci. Rep. 5, 15896.

Buckholtz, J.W., Marois, R., 2012. The roots of modern justice: cognitive and neural foundations of social norms and their enforcement. Nat. Neurosci. 15 (5), 655–661.

Cai, Y., Li, S., Liu, J., Li, D., Feng, Z., Wang, Q., Chen, C., Xue, G., 2016. The role of the frontal and parietal cortex in proactive and reactive inhibitory control: a transcranial direct current stimulation study. J. Cogn. Neurosci. 28 (1), 177–186.

Caspers, S., Eickhoff, S.B., Rick, T., von Kapri, A., Kuhlen, T., Huang, R., Shah, N.J., Zilles, K., 2011. Probabilistic fibre tract analysis of cytoarchitectonically defined human inferior parietal lobule areas reveals similarities to macaques. NeuroImage 58, 362–380.

Chiao, J.Y., Harada, T., Oby, E.R., Li, Z., Parrish, T., Bridge, D.J., 2009. Neural representations of social status hierarchy in human inferior parietal cortex. Neuropsychologia 47 (2), 354–363.
Decety, J., Lamm, C., 2007. The role of the right temporoparietal junction in social interaction: how low-level computational processes contribute to meta-cognition. Neuroscientist 13 (6), 580–593.
Dehaene, S., Spelke, E., Pinel, P., Stanescu, R., Tsivkin, S., 1999. Sources of mathematical thinking: behavioral and brain-imaging evidence. Science 284 (5416), 970–974.
Dehaene, S., Piazza, M., Pinel, P., Cohen, L., 2003. Three parietal circuits for number processing. Cogn. Neuropsychol. 20 (3), 487–506.
De Martino, B., Camerer, C.F., Adolphs, R., 2010. Amygdala damage eliminates monetary loss aversion. Proc. Natl. Acad. Sci. USA 107 (8), 3788–3792.
Donaldson, P.H., Rinehart, N.J., Enticott, P.G., 2015. Noninvasive stimulation of the temporoparietal junction: A systematic review. Neurosci. Biobehav. Rev. 55, 547–572.
Fehr, E., Krajbich, I., 2014. Social preferences and the brain. In: Glimcher, P.W., Fehr, E. (Eds.), Neuroeconomics: Decision Making and the Brain2nd ed.. Elsevier, Oxford, 193–218.
FeldmanHall, O., Dalgleish, T., Mobbs, D., 2013. Alexithymia decreases altruism in real social decisions. Cortex 49 (3), 899–904.
FeldmanHall, O., Dalgleish, T., Evans, D., Mobbs, D., 2015. Empathic concern drives costly altruism. NeuroImage 105, 347–356.
Fellows, L.K., Farah, M.J., 2005. Different underlying impairments in decision-making following ventromedial and dorsolateral frontal lobe damage in humans. Cereb. Cortex 15 (1), 58–63.
Friston, K.J., Harrison, L., Penny, W., 2003. Dynamic causal modelling. Neuroimage 19, 1273–1302.
Gandiga, P.C., Hummel, F.C., Cohen, L.G., 2006. Transcranial DC stimulation (tDCS): a tool for double-blind sham-controlled clinical studies in brain stimulation. Clin. Neurophysiol. 117 (4), 845–850.
Hamilton, W.D., 1964a. The genetical evolution of social behaviour. I. J. Theor. Biol. 7 (1), 1–16.
Hamilton, W.D., 1964b. The genetical evolution of social behaviour. II. J. Theor. Biol. 7 (1), 17–52.
Harbaugh, W.T., Mayr, U., Burghart, D.R., 2007. Neural responses to taxation and voluntary giving reveal motives for charitable donations. Science 316 (5831), 1622–1625.
Hein, G., Silani, G., Preuschoff, K., Batson, C.D., Singer, T., 2010. Neural responses to ingroup and outgroup members' suffering predict individual differences in costly helping. Neuron 68 (1), 149–160.
Hein, G., Lamm, C., Brodbeck, C., Singer, T., 2011. Skin conductance response to the pain of others predicts later costly helping. PLoS One 6 (8), e22759.
Heinsohn, R., Cockburn, A., 1994. Helping is costly to young birds in cooperatively breeding white-winged Choughs. Proc. R. Soc. B-Biol. Sci. 256, 293–298.
Heinsohn, R., Legge, S., 1999. The cost of helping. Trends Ecol. Evol. 14 (2), 53–57.
Hu, Y., Strang, S., Weber, B., 2015. Helping or punishing strangers: neural correlates of altruistic decisions as third-party and of its relation to empathic concern. Front. Behav. Neurosci. 9, 24.
Isen, A.M., Simmonds, S.F., 1978. The effect of feeling good on a helping task that is incompatible with good mood. Soc. Psychol. 41 (4), 346–349.
Kahnt, T., Park, S.Q., Haynes, J., Tobler, P.N., 2014. Disentangling neural representations of value and salience in the human brain. Proc. Natl. Acad. Sci. USA 111 (13), 5000–5005.
Knoch, D., Pascual-Leone, A., Meyer, K., Treyer, V., Fehr, E., 2006. Diminishing reciprocal fairness by disrupting the right prefrontal cortex. Science 314 (5800), 829–832.
Knoch, D., Nitsche, M.A., Fischbacher, U., Eisenegger, C., Pascual-Leone, A., Fehr, E., 2008. Studying the neurobiology of social interaction with transcranial direct current stimulation–the example of punishing unfairness. Cereb. Cortex 18 (9), 1987–1990.
Lefebvre, C.D., Marchand, Y., Eskes, G.A., Connolly, J.F., 2005. Assessment of working memory abilities using an event-related brain potential (ERP)-compatible digit span backward task. Clin. Neurophysiol. 116 (7), 1665–1680.
Mackey, W.E., Devinsky, O., Doyle, W.K., Meager, M.R., Curtis, C.E., 2016. Human dorsolateral prefrontal cortex is not necessary for spatial working memory. J. Neurosci. 36 (10), 2847–2856.
Marquez, C., Rennie, S.M., Costa, D.F., Moita, M.A., 2015. Prosocial choice in rats depends on food-seeking behavior displayed by recipients. Curr. Biol. 25 (13), 1736–1745.
Mars, R.B., Sallet, J., Schüffelgen, U., Jbabdi, S., Toni, I., Rushworth, M.F., 2012. Connectivity-based subdivisions of the human right "temporoparietal junction area": evidence for different areas participating in different cortical networks. Cereb. Cortex 22 (8), 1894–1903.
Mathur, V.A., Harada, T., Lipke, T., Chiao, J.Y., 2010. Neural basis of extraordinary empathy and altruistic motivation. NeuroImage 51 (4), 1468–1475.
Matthews, S.C., Simmons, A.N., Lane, S.D., Paulus, M.P., 2004. Selective activation of the nucleus accumbens during risk-taking decision making. Neuroreport 15 (13), 2123–2127.
Mengarelli, F., Spoglianti, S., Avenanti, A., di Pellegrino, G., 2015. Cathodal tDCS over the left prefrontal cortex diminishes choice-induced preference change. Cereb. Cortex 25 (5), 1219–1227.
Miller, E.K., Cohen, J.D., 2001. An integrative theory of prefrontal cortex function. Annu. Rev. Neurosci. 24, 167–202.
Moll, J., Krueger, F., Zahn, R., Pardini, M., de Oliveira-Souza, R., Grafman, J., 2006. Human fronto–mesolimbic networks guide decisions about charitable donation. Proc. Natl. Acad. Sci. USA 103 (42), 15623–15628.
Nihonsugi, T., Ihara, A., Haruno, M., 2015. Selective increase of intention-based economic decisions by noninvasive brain stimulation to the dorsolateral prefrontal cortex. J. Neurosci. 35 (8), 3412–3419.
Ono, K., Mikami, Y., Fukuyama, H., Mima, T., 2016. Motion-induced disturbance of auditory-motor synchronization and its modulation by transcranial direct current stimulation. Eur. J. Neurosci. 43 (4), 509–515.
Park, S.Q., Kahnt, T., Rieskamp, J., Heekeren, H.R., 2011. Neurobiology of value integration: when value impacts valuation. J. Neurosci. 31 (25), 9307–9314.
Penner, L.A., Dovidio, J.F., Piliavin, J.A., Schroeder, D.A., 2005. Prosocial behavior: multilevel perspectives. Annu. Rev. Psychol. 56, 365–392.
Penny, W.D., Stephan, K.E., Mechelli, A., Friston, K.J., 2004. Comparing dynamic causal models. Neuroimage 22, 1157–1172.
Penny, W.D., Stephan, K.E., Daunizeau, J., Rosa, M.J., Friston, K.J., Schofield, T.M., Leff, A.P., 2010. Comparing families of dynamic causal models. PLoS Comput. Biol. 6, e1000709.
Pinel, P., Dehaene, S., Rivière, D., LeBihan, D., 2001. Modulation of parietal activation by semantic distance in a number comparison task. NeuroImage 14 (5), 1013–1026.
Pinel, P., Piazza, M., Le Bihan, D., Dehaene, S., 2004. Distributed and overlapping cerebral representations of number, size, and luminance during comparative judgments. Neuron 41 (6), 983–993.
Price, D.D., Bush, F.M., Long, S., Harkins, S.W., 1994. A comparison of pain measurement characteristics of mechanical visual analogue and simple numerical rating scales. Pain 56, 217–226.
Ruff, C.C., Ugazio, G., Fehr, E., 2013. Changing social norm compliance with noninvasive brain stimulation. Science 342 (6157), 482–484.
Sellaro, R., Nitsche, M.A., Colzato, L.S., 2016. The stimulated social brain: effects of transcranial direct current stimulation on social cognition. Ann. N.Y. Acad. Sci. 1369 (1), 218–239.
Strang, S., Gross, J., Schuhmann, T., Riedl, A., Weber, B., Sack, A., 2014. Be Nice if You Have to - The Neurobiological Roots of Strategic Fairness. Soc. Cognit. Affect. Neurosci. 10, 790–796.
Stephan, K.E., Penny, W.D., Daunizeau, J., Moran, R.J., Friston, K.J., 2009. Bayesian model selection for group studies. Neuroimage 46, 1004–1017.
Sul, S., Tobler, P.N., Hein, G., Leiberg, S., Jung, D., Fehr, E., Kim, H., 2015. Spatial gradient in value representation along the medial prefrontal cortex reflects individual differences in prosociality. Proc. Natl. Acad. Sci. USA 112 (25), 7851–7856.
Toi, M., Batson, C.D., 1982. More evidence that empathy is a source of altruistic motivation. J. Pers. Soc. Psychol. 43 (2), 281–292.
Tom, S.M., Fox, C.R., Trepel, C., Poldrack, R.A., 2007. The neural basis of loss aversion in decision-making under risk. Science 315 (5811), 515–518.
Tusche, A., Bockler, A., Kanske, P., Trautwein, F.M., Singer, T., 2016. Decoding the charitable brain: empathy, perspective taking, and attention shifts differentially predict altruistic giving. J. Neurosci. 36 (17), 4719–4732.
Utz, K.S., Dimova, V., Oppenländer, K., Kerkhoff, G., 2010. Electrified minds: transcranial direct current stimulation (tDCS) and Galvanic Vestibular Stimulation (GVS) as methods of non-invasive brain stimulation in neuropsychology – a review of current data and future implications. Neuropsychologia 48, 2789–2810.
Van Duijvenvoorde, A.C., Huizenga, H.M., Somerville, L.H., Delgado, M.R., Powers, A., Weeda, W.D., Casey, B.J., Weber, E.U., Figner, B., 2015. Neural correlates of expected risks and returns in risky choice across development. J. Neurosci. 35 (4), 1549–1560.
Villamar, M.F., Wivatvongvana, P., Patumanond, J., Bikson, M., Truong, D.Q., Datta, A., Fregni, F., 2013a. Focal modulation of the primary motor cortex in fibromyalgia using 4×1-ring high-definition transcranial direct current stimulation (HD-tDCS): immediate and delayed analgesic effects of cathodal and anodal stimulation. J. Pain 14 (4), 371–383.
Villamar, M.F., Volz, M.S., Bikson, M., Datta, A., Dasilva, A.F., Fregni, F., 2013b. Technique and considerations in the use of 4 × 1-ring high-definition transcranial direct current stimulation (HD-tDCS). J. Vis. Exp. 77, e50309.
Wang, L., Yu, H., Hu, J., Theeuwes, J., Gong, X., Xiang, Y., Jiang, C., Zhou, X., 2015. Reward breaks through center-surround inhibition via anterior insula. Hum. Brain Mapp. 36 (12), 5233–5251.
Warneken, F., Hare, B., Melis, A.P., Hanus, D., Tomasello, M., 2007. Spontaneous altruism by chimpanzees and young children. PLoS Biol. 5 (7), e184.
Warneken, F., Tomasello, M., 2009. Varieties of altruism in children and chimpanzees. Trends Cogn. Sci. 13 (9), 397–402.
Watson, D., Clark, L.A., Tellegen, A., 1988. Development and validation of brief measures of positive and negative affect: the PANAS scales. J. Pers. Soc. Psychol. 54 (6), 1063–1070.
Waytz, A., Zaki, J., Mitchell, J.P., 2012. Response of dorsomedial prefrontal cortex predicts altruistic behavior. J. Neurosci. 32 (22), 7646.
Wechsler, D., 1997. WAIS-III: Administration and Scoring Manual. The Psychological Corporation, San Antonio.
Xue, G., Lu, Z., Levin, I.P., Weller, J.A., Li, X., Bechara, A., 2009. Functional dissociations of risk and reward processing in the medial prefrontal cortex. Cereb. Cortex 19 (5), 1019–1027.
Zhu, L., Jenkins, A.C., Set, E., Scabini, D., Knight, R.T., Chiu, P.H., King-Casas, B., Hsu, M., 2014. Damage to dorsolateral prefrontal cortex affects tradeoffs between honesty and self-interest. Nat. Neurosci. 17 (10), 1319–1321.