

# Disrupting the dorsolateral prefrontal cortex attenuates the difference in decision-making for altruistic punishment between the gain and loss contexts

**Yingjie Liu**

North China University of Science and Technology

**Hongbo Xing**

North China University of Science and Technology

**Yuan Gao**

North China University of Science and Technology

**Baxter DiFabrizio**

**He Wang** (✉ [wanghe429@163.com](mailto:wanghe429@163.com))

North China University of Science and Technology

---

## Research Article

**Keywords:** altruistic punishment, fairness, loss context, rTMS, DLPFC

**Posted Date:** March 20th, 2023

**DOI:** <https://doi.org/10.21203/rs.3.rs-2689936/v1>

**License:** © ⓘ This work is licensed under a Creative Commons Attribution 4.0 International License.

[Read Full License](#)

**Additional Declarations:** No competing interests reported.

---

**Version of Record:** A version of this preprint was published at Brain Topography on January 10th, 2024.

See the published version at <https://doi.org/10.1007/s10548-023-01029-9>.

# Abstract

Altruistic punishment are primary responses to social norms violations, its neural mechanism has also attracted extensive attention from researchers. In the present studies, we applied a low-frequency repetitive transcranial magnetic stimulation (rTMS) to the bilateral dorsolateral prefrontal cortex (DLPFC) while participants engaged in a modified Ultimatum Game (UG) (Study 1) and a modified third-party punishment game (Study 2) to probe the neural mechanisms affecting decision-making in altruistic punishment. Normally punishers intervene more often against and show more social outrage towards Dictators/Proposers who unfairly distribute losses than rather those who unfairly share gains. We found that disrupting the function of the left and right DLPFC with rTMS effectively obliterated this difference, making participants punish unfairly shared gains as often as they normally would punish unfairly shared losses. The inhibition of the DLPFC function will lead to the deviation of individual information integration ability and influence the moderating effect of gain and loss contexts on altruistic punishment. Our findings emphasize that DLPFC is closely related to altruistic punishment and provide causal neuroscientific evidence.

## 1 Introduction

The development of human society is based on social fairness norms. When the social fairness norms are violated, people will response to it. This "response" is usually proposed by the "second party" or third party observer involved in the specific interaction (Fehr & Fischbacher, 2004). The second parties are directly involved in the socially unfair event and are directly affected by the specific behavior that violated fairness norms (Halevy & Halali, 2015). Their rejection of an unfair offer is regarded as an punishment, the punishment is altruistic, because they will not get any benefit after rejecting the unfair offer (Fehr & Fischbacher, 2004). The third parties are an uninvolved outsider who happens to know about, observe, or assess and address the fairness violation, a party whose economic payoffs are generally unaffected by the specific bad act (Oswald et al., 2010). As such, second-party punishment and third-party punishment for the benefit of society at large can be thought of as "altruistic punishments," as they confer no direct material benefit on the adjudicator (Sääksvuori et al., 2011). Evolutionary models and empirical evidence suggest that altruistic punishment has been a vital force in the evolution of human fairness (Fehr & Fischbacher, 2003).

Researchers believe that altruistic punishment was one form of strong reciprocity established during evolution (House et al., 2020; Fehr & Fischbacher, 2003). Fairness and reciprocity norms motivate bystanders to act (almost reflexively) at a considerable cost to themselves (Org, 2010). Society relies on "strong reciprocators," people who act to enforce norms without direct rewards (Gintis, 2003; Fehr et al., 2002). According to the theory of strong reciprocity, long-term widespread fairness and cooperation are made possible by the presence of "strong reciprocators" keeping everyone in line (Buckholtz & Marois, 2012).

Nowadays, more and more researchers began to pay attention to the brain mechanism of maintaining fairness. Neuroimaging studies have identified, for example, the insula (Sanfey et al., 2003; Güroğlu et al., 2010; Kirk et al., 2011) and dorsal striatum, which are activated when someone rejects an unfair offer (second-party punishment) (de Quervain et al., 2004; Baumgartner et al., 2008). Moreover, rejecting unfair proposals requires a (small) personal cost and the suppression of selfish impulses, which is done by the DLPFC (Buckholtz et al., 2008; Baumgartner et al., 2011, 2012). When it comes to personal interest motives, the DLPFC may refer to the mediation of economic exchanges (Knoch et al., 2006, 2008). The DLPFC also plays a key role in third party punishment for maintaining fairness. It is concerned with central executive network and likely responsible for converting the social evaluation into a decision to act and punish (Buckholtz & Marois, 2012). The bilateral activity of the DLPFC could also perform an inhibitory action within the executive network when planning punishment behaviors (Krueger & Hoffman, 2016). The DLPFC plays a crucial, final-stage role, both in integrating mental state information, assessing the victim's harm, and selecting a suitable punishment (Ginther et al., 2016). And researchers have claimed that the region acts as a superordinate node supporting the integration of signals so individuals can select an appropriate punishment decision or action (Buckholtz et al., 2008, 2015; Treadway et al., 2014).

In recent years, researchers have also noted the function lateralization of DLPFC. The right DLPFC is recruited when participants decide whether to punish a partner by rejecting an unfair economic deal that the partner proposes (Sanfey et al., 2003). The right DLPFC is strongly activated by the decision to punish violations of norms based on a sensitive assessment of the transgressor's culpability (Buckholtz et al., 2008). And it has also been consistently associated with altruistic punishment, valuation judgments, and fairness (Moll et al., 2005; Guo et al., 2013). Other research suggests that the left DLPFC is involved in weighing and rejecting unfair offers, which supports the idea that the DLPFC incorporates some normative idea of in-born socially related fairness, by overriding one's impulse to take something rather than nothing at all (something a strictly rational actor might do) to enforce some sort of social fairness norm (Güroğlu et al., 2010). The left DLPFC was shown to evince high spikes in activation when a third party punished another person rather than when the participant was just observing an act of punishment (Buckholtz et al., 2008), and the left DLPFC has also been found to be related to executive function and impulse control (Ochsner et al., 2002; Figner et al., 2010; Barbey et al., 2012). Thus, both sides of DLPFC are likely involved in altruistic punishment.

However, social fairness cannot be separated from economic gains and losses. Most of the studies were on participants who adjudicated norm violations among unfairly shared gains (the gain context), people's reactions to violating social norms in loss context also need to be further explored. The latter situation commonly appears in human society, for example in the liquidation of a bankrupt enterprise. Previous studies have found that the functions of bilateral DLPFC are different in the gain and loss contexts. The left DLPFC activated in association with rejecting offers (vs. accepting them) in the loss context, but not in the gain context (Guo et al., 2013). Activations in DLPFC were enhanced in the loss context (Wu et al., 2014). In the third-party punishment, participants were more inclined to choose to altruistically punish Dictators in the loss context than in the gain context (Liu et al., 2017; 2019). But the neural signatures

associated with third-party punishment have not been fully elucidated. The decision-making in altruistic punishment across the gain and loss contexts might be related to functional DLPFC lateralization. Lesions in the left DLPFC can inhibit and hinder many foundational cognitive operations: concept formation skills, especially in terms of problem-solving skills; labeling attributes in verbal and visual domains; inhibiting previous responses; cognitive flexibility; and lesions in the DLPFC have been shown to weaken one's ability to elucidate and explain abstract concepts (Delis, 2001; Litzman & Markon, 2010). rTMS studies support the dissociation of the right DLPFC from the left's functions. Destruction of the DLPFC on the right reduces an individual's willingness to reject unfairness, even if they believe it to be very unfair (Knoch et al., 2006). An inhibited or underactive DLPFC may directly affect people's responses regarding enforcing normative fairness. We hypothesize that DLPFC disruption will lead to different altruistic punishment decisions when adjudicating malefactors in both the gain and loss contexts.

To pursue this research aim, of localizing the neural architecture of altruistic punishment and testing its behavior under disruption, we applied low-frequency transcranial magnetic stimulation to the participants' right DLPFC or their left DLPFC for 20 minutes, or we presented them with "sham TMS" who then acted as either second party (recipient) or third-party (judge/bystander) responders to events of social unfairness. Low-frequency rTMS for several minutes leads to a suppression of activity in the stimulated brain regions that outlast the duration of the rTMS train for about half of the duration of the stimulation (Baumgartner et al., 2011). Taken together, in the present research, we used a modified UG (Study 1) (Guo et al., 2013; Wu et al., 2014) and a modified third-party punishment game (Study 2) (Liu et al., 2017; 2019), and applied rTMS to investigate the neural mechanism generating differences between second-party punishment and third-party punishment decision-making between the gain and loss contexts. Specifically, we expected this research could provide help to illuminate how the norm violation being perceived would affect the neural processing going into an altruistic punishment.

## **2 Study 1 Disrupting Dorsolateral Prefrontal Cortex Influence On The Difference Of Second-party Punishment Decision-making Between The Gain And Loss Contexts**

### 2.1 Method

#### 2.1.1 Participants

A power analysis (G\*Power 3.1, effect size  $f^2 = 0.2$ , power=0.95,  $\alpha=0.05$ ) suggested that 54 participants could detect significant effects with 95% statistical power (Bezerra et al. 2019; Faul et al. 2007; Folwaczny et al. 2021). Finally, sixty healthy students participated in the study (24 males, mean age = 21.52, SD = 2.75, the age range is 18-26). Among them, 20 participants (8 males, mean age = 21.70, SD = 2.49) were stimulated with left DLPFC. 20 participants (8 males, mean age = 20.70, SD = 2.58) were stimulated with right DLPFC. 20 participants (8 males, mean age = 22.1, SD = 3.20) were assigned randomly to the sham stimulus group. There was no significant difference in age among the three groups,  $F(2, 57) = 0.67, p = 0.519$ . All participants were right-handed according to the Edinburgh

Handedness Inventory (Oldfield, 1971), had normal or corrected to normal vision, and had no clinical history of neurological or psychiatric disorders or any other specific contraindications to non-invasive brain stimulation (Rossi et al., 2009). Each participant completed the Adult Safety Screening Questionnaire as a safety check (Keel et al., 2001).

Before the experiment, the participants were informed of the current usage status of TMS in medicine, neuroscience, and other fields. The participants were then checked for whether they had any concerns. All participants volunteered to participate in the experiment and gave written informed consent before study procedures and after this prefatory explanation of what rTMS would entail. This study was approved by the Ethics Committee of Human Experiment of the local university and was carried out by the ethical standards of the revised Helsinki Declaration. The experimental operators received professional training in TMS operation and bilateral DLPFC positioning. At the end of the experiment, the participants were paid according to their winnings and debriefed.

### 2.1.2 Criteria for selecting participants

Participants were not allowed to participate in this study if they had or manifested any of the following conditions: (1) Severe dementia and organic disease; (2) metal implants; (3) had a personal or family history of traumatic brain injury or epilepsy; (4) obvious extrapyramidal adverse reactions; (5) illegal drug use; (6) a severe mental disorder; (7) a history of stroke; (8) a history of head injury or neurosurgery; (9) frequent or severe headaches; (10) had taken any medication that is associated with lowering the onset threshold; (11) were pregnant or sexually active and not using contraceptives.

### 2.1.3 Experimental design

Study 1 used a 2 (Context: Gain, Loss) × 3 (rTMS stimulation: Left DLPFC stimulation, Right DLPFC stimulation, Sham stimulation) mixed experimental design, in which rTMS stimulation was a between-subjects variable and Context was a within-subjects variable. The dependent variable measured was the rate at which these participants would reject unfair offers. All three groups completed the same task. The difference is that the participants in the left and right DLPFC groups received 1Hz rTMS stimulation before the experiment, receiving the rTMS stimulation on either their left or right DLPFC for a duration of 20 minutes. No magnetic field was emitted from the stimulus coil in the sham stimulus group, and the total duration of the sham stimulus was also 20 minutes.

### 2.1.4 Experiment task and procedure

We used a modified Ultimatum Game (UG) (Guo et al., 2013; Wu et al., 2014) presented as a computerized task. Before the experiment, we explained to participants that they would play a repeated UG in a computerized task with a group of anonymous students (Proposer, A) whom they did not know making proposals (offers). Participants would act as responders (B). In each round, a proposer (A) would make an offer about how to share a 100 monetary unit (MU) gain or loss with the respondent (B). Specifically, in the gain context, each proposer would distribute 100 MUs of income (approximately equal to USD

14.6) to the proposer/recipient pair in 90/10, 70/30, or 50/50 distribution schemes, these three conditions which respectively represented high unfairness, medium unfairness, or fairness. Upon observing the distribution, the participant would choose to accept or reject the offer. For example, at 70:30. If the participant chose to "accept" the allocation scheme, he was given the amount allocated in that scheme; if the participant chose to "reject" the offer, both proposer and responder would receive 0 MUs. In the loss context, the proposer distributed a *loss* of 100 MUs to the respondent in three conditions: -10/-90, -30/-70, or -50/-50, and trials were otherwise the same as the gain context trials. For example, -30/-70. If the participants chose to "accept" the distribution, the proposers would lose 30 MUs and the participants would lose 70 MUs. If the participants chose to "reject", both the proposers and the participants would lose 100 MUs. The participants' responses were independent in each trial, meaning they were faced with a different person in each trial.

The specific experimental trials progressed as follows: in each trial, a black fixation point was first presented for 500ms. It was followed by either the word "gain" or "loss" for 1000ms. Then, the screen displayed the distribution for 1500ms. Afterward, it displayed the player's options: "accept" or "reject." Participants pressed "F" to reject, and "J" to accept. There is no time limit at this stage. The placement of buttons and options was counterbalanced among participants. After the participant made a response, a white blank screen was displayed for 500ms. See Figure 2-1. Each participant completed 4 practice trials before the task. There were 24 trials in the formal experiment, and each condition was presented with 4 replications. The duration of the experiment is about 10 minutes. Each trial was presented in a pseudo-random order—that is, no more than three consecutive trials were all gains or losses. At the end of the experiment, the participants received a participation fee that was assessed based on the MUs they had accrued during a random trial of the experiment and were debriefed.

### 2.1.5 rTMS experimental parameters and DLPFC positioning method

We applied low-frequency (1 Hz) rTMS to the DLPFC for 20 min (1200 pulses) before subjects participated in the formal experiment (off-line paradigm), using a Magstim Rapid Magnetic Stimulator (Magstim, Winchester, MA, USA) and a commercially available figure-of-eight coil (70-mm diameter double-circle, air-cooled).

Before the formal implementation of the stimulation, the resting motor threshold (MT) of each participant was measured by applying 50% of the intensity of stimulation to the right primary motor cortex to obtain the minimum observable contraction of the muscles of their left hand. The intensity of the stimulation was measured by the resting motor threshold of the participant (Ren et al., 2015). At the time of formal stimulation, the intensity of the rTMS was set to 90% of the resting movement threshold of the participant (Hartwigsen et al., 2010). Each participant received 20 series of 1Hz rTMS, each of which contained 60 magnetic stimulation pulses, with an interval of 1s or 1200 consecutive 1Hz sub-stimulation pulses between the two series (Rossi et al., 2009). The localization of DLPFC referred to the positions of F3 and F4 in the international 10-20 EEG system (specifically located in Brodmann areas 9 and 46) (Koch et al., 2003; Mylius et al., 2013). The coordinates of the left DLPFC was (-42, 44, 30) and the coordinates of the

right DLPFC was (27, 39, 51). The junction of the two wings of the coil was placed in position on the F3 and F4, and the magnetic target was tangent to the scalp. The position and orientation of the coil placement for each subject were presented on the nylon positioning cap of the 10-20 system worn by the subject during the entire rTMS stimulation (Kearney- Ramos et al., 2018). The rTMS parameters were well within currently recommended guidelines (Rossi et al., 2009) and result in a suppression of excitability of the targeted cortical region for several minutes following completion of the rTMS train (Eisenegger et al., 2008; Baumgartner et al., 2011).

## 2.2 Results

### 2.2.1 Decision-making rate

Statistical analyses were conducted on unfair trials only (David et al., 2017; Liu et al., 2017, 2019, 2021) and we combined the two levels of unfair distribution in the analysis. We excluded fair trials (50:50/ -50: -50) because participants never choose to reject on fair trials ( $M_{\text{reject}} = 0$ ). In the experiment, we asked the participants to choose between rejection and acceptance. The probability of choosing to accept is determined by the probability of choosing to reject ( $P(\text{accept}) = 1 - P(\text{reject})$ ). We used a 2 (Context: Gain vs. Loss)  $\times$  3 (rTMS stimulation: Left DLPFC, Right DLPFC, Sham stimulation), which was analyzed by mixed ANOVA. The main effect of Context was significant ( $F(1, 54) = 22.77, p < 0.001, \eta_p^2 = 0.46$ ; see Table 2-1). Participants chose to reject more often in the loss context than they did in the gain context. The main effect of rTMS stimulation was not significant ( $F(1, 54) = 0.952, p = 0.398$ ; see Table 2-1). The interaction between Context and rTMS stimulation was significant ( $F(1, 54) = 3.46, p = 0.045, \eta_p^2 = 0.20$ , see Fig. 2-1). Simple effect analysis showed that in the gain context, the rejection rate of the Left DLPFC ( $M \pm SD = 0.80 \pm 0.24$ ) was significantly different from that of the Sham stimulus ( $M \pm SD = 0.58 \pm 0.24$ ),  $p = 0.048$ ; the rejection rate between the Left DLPFC and the Right DLPFC ( $M \pm SD = 0.74 \pm 0.24$ ), the Right DLPFC and the Sham stimulus were no significant difference,  $ps > 0.098$ , Tukey HSD correction. In the loss context, there was no significant difference in the rejection rate between the Left DLPFC ( $M \pm SD = 0.86 \pm 0.20$ ) and the Right DLPFC ( $M \pm SD = 0.88 \pm 0.20$ ), and the Right DLPFC also showed no significant difference from the Sham stimulus ( $M \pm SD = 0.84 \pm 0.20$ ),  $ps > 0.763$ , Tukey HSD correction. Meanwhile, when the left DLPFC was stimulated, there was no significant difference in rejection rate between the gain and loss contexts,  $p = 0.283$ . When the right DLPFC was stimulated, the rejection rate in the gain and loss contexts was significantly different,  $p = 0.023$ ; In the Sham stimulus, the rejection rate was significantly different between the gain and loss contexts, replicating prior findings,  $p < 0.001$ .

— INSERT TABLE 2-1 ABOUT HERE —

## 3 Study 2 Disrupting Dorsolateral Prefrontal Cortex Influence On The Difference Of Third-party Punishment Decision-making Between The Gain And Loss Contexts

### 3.1 Method

### 3.1.1 Participant

Using the same method as study 1, effect size  $f^2 = 0.2$ , power = .95 and  $\alpha = .05$ , the sample size of  $N = 54$ . We recruited sixty participants for the study (28 males, mean age = 19.50, SD = 1.87, the age range is 18-24). Among them, 20 participants (10 males, mean age = 19.70, SD = 2.35) were stimulated with left DLPFC. 20 participants (12 males, mean age = 19.10, SD = 1.52) were stimulated with right DLPFC. 20 participants (10 males, mean age = 19.70, SD = 1.77) in the sham stimulus group. There was no significant difference in age among the three groups,  $F(2, 54) = 0.33$ ,  $p = 0.724$ . The remaining requirements were the same as in Study 1.

### 3.1.2 Criteria for selecting participants

We used the same criteria for selecting participants as we did in Study 1.

### 3.1.3 Experimental design

Study 2 used a 2 (Context: Gain, Loss)  $\times$  3 (rTMS stimulation: Left DLPFC stimulation, Right DLPFC stimulation, Sham stimulation) mixed experimental design, in which rTMS stimulation was a between-subjects variable and Context was a within-subjects variable. The dependent variables for this experiment were the punishment rate (the likelihood of punishing) and punishment amount, the amount of MUs participants expended on punishment. The rest of the method mirrors Study 1.

### 3.1.4 Experiment task and procedure

We used a modified third-party punishment game (Liu et al., 2017, 2019) presented as a computerized task to assess participants' willingness and harshness in judging the unfair behavior of others. In the experiment, participants acted as a third party, observing a Dictator Game (DG). The DG trials were performed by another group of participants before the formal experiment. There were two people in the task, one the Dictator or the Proposer (A), and the other the recipient (B). Proposer (A) and the recipient (B) receive 100 MUs together, A decides how to distribute these MUs among the two of them, and B has no choice but to accept the proposal. Third-party (participant) received 50 MUs at the beginning of each round (Fehr & Fischbacher, 2004; Leliveld et al., 2012; Hu et al., 2015; Liu et al., 2017, 2019), and after observing the Dictator game, they could use the MUs to punish the proposer or choose to retain the MUs without acting on the situation they observed. Notably, the relationship between the MUs offered by the participants for punishment and the MUs reduced by the proposer was 1:3 (Fehr & Fischbacher, 2004). So, if a third party chose to punish the proposer and then offered 5 of their MUs to enact that punishment, the proposer would lose 15 MUs.

To be more specific, A and B appeared in each trial and jointly earned 100 MUs or lost 100 MUs together (approximately equal to USD 14.6). In the gain context, each proposer would distribute 100 MUs of income to the recipient in 90/10, 70/30, or 50/50 distribution, these three conditions which respectively represented high unfairness, medium unfairness, or fairness. Upon observing the distribution, participants would choose to punish the Dictator or to keep their MUs. If the participants chose the punishment, they



had to choose a specific amount of MUs to expend on punishment from their purse of 50 MUs, which was replenished with each trial. There were 6 options, including 5, 10, 15, 20, 25, and 30. In the loss context, the proposer would distribute a loss of 100 MUs to the recipient in three conditions: -10/-90, -30/-70, or -50/-50. Other processes mirrored those of the gain context. The responses of the participants in each trial were independent, that is, they were faced with two different people in each trial.

The specific experimental trials progressed as follows: in each trial, a black fixation point was first presented for 500ms. It was followed by either the word "gain" or "loss" for 1000ms. Then, the screen displayed the distribution for 1500ms. Afterward, it displayed the player's options: "punishment" or "keep". Participant pressed "F" for punishment, and "J" to keep. There is no time limit at this stage. The placement of buttons and options was counterbalanced among participants. If participants chose to punish the Dictator, a gauge for the specific amount of MUs to spend to enact his punishment appeared on the screen. It presented 6 options in total: participants could choose to spend 5, 10, 15, 20, 25, or 30 MUs on hurting the Dictator's earnings, options which corresponded to the numbers 1, 2, 3, 4, 5, and 6 on the numeric keypad. After the participant made a response, a white blank screen was displayed for 500ms before the next trial. See Figure 3-1. Each participant completed 4 practice trials before the task. There were 24 trials in the formal experiment, and each condition was presented with 4 replications. Each trial was presented in a pseudo-random order—that is, no more than three consecutive trials were all gains or losses. At the end of the experiment, the participants received a participation fee that was assessed based on the MUs they had accrued during a random trial of the experiment and were debriefed.

### 3.1.5 rTMS experimental parameters and DLPFC positioning method

The same procedure was implemented here as in Study 1.

## 3.2 Result

### 3.2.1 Decision-making rate

Consistent with study 1, we only analyzed the unfair trials (David et al., 2017; Liu et al., 2017, 2019, 2021). The probability of choosing to punish is determined by the probability of choosing to keep ( $P(\text{keep}) = 1 - P(\text{punish})$ ). 2 (Context: Gain vs. Loss)  $\times$  3 (rTMS stimulation: Left DLPFC stimulation, Right DLPFC stimulation, Sham stimulation) was analyzed by mixed ANOVA. The main effect of Context was significant ( $F(1, 54) = 4.35$ ,  $p = 0.045$ ,  $\eta_p^2 = 0.14$ ; see Table 3-1). Participants chose to punish the Dictator more often in the loss context than they did in the gain context. The main effect of rTMS stimulation was significant ( $F(1, 54) = 5.85$ ,  $p = 0.008$ ,  $\eta_p^2 = 0.30$ ; see Table 3-1). Tukey-HSD multiple comparison analysis found the punishment rate after stimulation of left DLPFC and right DLPFC was respectively significantly higher than that of Sham stimulation,  $ps < 0.010$ . The left DLPFC and right DLPFC did not elicit a significant difference in punishment rate between each other,  $p = 0.704$ . The interaction between Context and rTMS stimulation was significant ( $F(1, 54) = 7.45$ ,  $p = 0.003$ ,  $\eta_p^2 = 0.36$ , see Fig. 3-1). Simple effect analysis showed that, in the gain context, the difference of punishment rate between the Left

DLPFC ( $M \pm SD = 0.95 \pm 0.23$ ) and Right DLPFC ( $M \pm SD = 0.98 \pm 0.23$ ) was not significant,  $p = 0.910$ , but the difference between the Left DLPFC and Sham stimulus ( $M \pm SD = 0.60 \pm 0.23$ ) was significant,  $p = 0.001$ . The difference between the Right DLPFC and Sham stimulus was also significant,  $p = 0.002$ , Tukey HSD correction. In the loss context, the difference among the Left DLPFC, the Right DLPFC, and Sham stimulus were not significant,  $ps > 0.584$ , Tukey HSD correction. When the left DLPFC was stimulated, there was no significant difference in punishment rate between the gain and loss contexts,  $p = 0.290$ . When the right DLPFC was stimulated, the punishment rate in the gain and loss contexts was not significantly different,  $p = 0.646$ . In the Sham stimulus, the rejection rate was significantly different between the gain and loss contexts, replicating prior research,  $p < 0.001$ .

— INSERT TABLE 3-1 ABOUT HERE —

### 3.2.2 Decision-making amount

We only analyzed the magnitude of MUs spent on third-party punishment in unfair trials, as none was spent in fair trials. The main effect of Context was not significant ( $F(1, 54) = 2.71, p = 0.111$ ; see Table 3-1). The main effect of rTMS stimulation was significant ( $F(1, 54) = 9.462, p = 0.001, \eta_p^2 = 0.41$ ; see Table 3-1). Tukey-HSD multiple comparison analysis found that the punishment amount after stimulation of both the left DLPFC and right DLPFC respectively was significantly higher than the suggested punishment amount in the Sham stimulation,  $ps < 0.025$ ; the difference between the amount spent given either left DLPFC and right DLPFC stimulation was not significant  $p = 0.060$ . The interaction between Context and rTMS stimulation was also not significant ( $F(1, 54) = 2.21, p = 0.129$ , see Fig. 3-1).

## 4 Discussion

In the current research, we applied rTMS and used a modified UG (Study 1) (Guo et al., 2013; Wu et al., 2014) and a modified third-party punishment game (Study 2) (Liu et al., 2017; 2019) to investigate the neural mechanisms underlying the differences between second-party punishment and third-party punishment decision-making across the gain and loss contexts. The two of present researches confirmed previous findings that people are willing to pay the cost to enforce social norms (Fehr & Fischbacher, 2004; Leliveld et al., 2012). It also provided new neural evidence about how humans weigh the elements of social harm when choosing an altruistic intervention strategy that optimizes social welfare.

Second parties rejected unfair offers and third parties intervened more often in the context of unfairly shared losses than they did with unfairly shared gains. This result corroborated previous studies (Guo et al., 2013; Wu et al., 2014; Liu et al., 2017, 2019). Perhaps this means that social pressure around normative fairness is more urgent around mutually negative experiences, inducing a stronger desire to sanction unbrotherly behavior in the loss context than in the gain context (Zhou & Wu, 2011).

To further explore the neural mechanism of altruistic punishment, we applied rTMS to the left or right DLPFC of the participants and found that disrupting the function of the left, but not right DLPFC, significantly raised the likelihood that the second party would punish the Proposer in the gain context,

effectively making participants as sensitive to unfairly distributed gains as they were to unfairly distributed losses. However, a study found that disruption of the right DLPFC by rTMS substantially reduces people's willingness to reject unfair offers (Knoch et al., 2006). Although our result was inconsistent with theirs, we focused on how the rejection of unfairness changed in the gain and loss contexts. The participants had to consider the existence of the loss context, which led to the deviation of their behavior. The DLPFC acts as a key node that supports the integration of signals and relevant information to select the appropriate punishment decision (Buckholz et al., 2008, 2015; Treadway et al., 2014). Previous studies have shown that activation of the DLPFC is associated with implementing cognitive control (Sanfey et al., 2003; Knoch et al., 2006; Haushofer & Fehr, 2008). Damage to the left DLPFC was associated with increased sensitivity to and increased rejecting of unfair offers, providing support for the engagement of fairness-related left DLPFC activity in top-down executive control over impulses to demand parity (Güroğlu et al., 2010, 2011). It seems involved in sanctioning norm violators. In the loss context, the left DLPFC showed greater activations than it did in the gain context, as it also has in prior research (Guo et al., 2013). Therefore, on the one hand, the damage of left DLPFC may result in impaired cognitive control and the deviation of individual information integration ability, make people more sensitive to unfairness, and improve the awareness of altruistic punishment to maintain the social norms of fairness- we could say it sharpens one's sensitivity to unfairness in general. On the other hand, the damage of left DLPFC may just weaken the influence of situational information on someone's judgment of fairness, perhaps leading to an evolutionarily negative outcome.

We also found that the DLPFC played an important role across contexts on third-party punishment. The bilateral activity of the DLPFC could underlie inhibitory actions of our executive control network when we are trying to titrate out or balance our punishment strategies to fit the norm violation we are addressing (Krueger & Hoffman, 2016). In Study 2, we inhibited the function of the left DLPFC by rTMS and found participants punished all behaviors like they were unfairly shared losses, that is, harshly and frequently, when they typically would treat unfairly shared gains more permissively.

However, we found that disrupting the function of the right DLPFC would raise the rate of third-party punishment in the gain context to match that of the loss context. In the third-party punishment decision-making network, the DLPFC is thought to support goal-directed response selection (Buckholz & Marois, 2012; Tabibnia et al., 2008). Moreover, the right DLPFC is responsive to violations of social norms and is causally involved in adhering to social fairness norms (Ruff et al., 2014). The function of right DLPFC appears to apply equally to situations where the motive for punishment is unfair behavior in a dyadic economic exchange and when responding to the violation of an institutionalized social norm in a disinterested third-party context (Buckholz et al., 2008). Therefore, third-party punishment requires the integration of several brain functions. In addition to inhibiting the left DLPFC to reduce people's judgment of the impact of loss and gain context, inhibiting the right DLPFC will further affect people's affect people's information integration of unfair events and improve people's sensitivity to unfairness. It led to an increase in the punishment rate in the gain context.

Beyond that, we also found a consistent results, there was no effects of rTMS whatsoever in the loss context, in neither study. Participants were likely to perceive higher unfairness level and have stronger desire to sanction social norm violations in the loss context than in the gain context (Zhou & Wu, 2011; Guo et al., 2013). The switch from the gain context to the loss context effectively changed subjective perception of social norm violations as well as fairness considerations. In the loss context, people have taken more action against the unfair behavior. Another reason could be the emergence of a ceiling effect, because rejection and punishment in the loss context were high already. It is also possible that there was a ceiling effect in the experiment, because rejection and punishment were already high in the loss context. It showed that people's judgment of unfairness did not change in the loss context.

Importantly, the DLPFC was preferentially or more consistently activated by prosocial behaviour (Bellucci et al., 2020) and showed a causal role in mentalizing (Kalbe et al., 2010). Functional decoding analyses linked the DLPFC to roles in both cognitive and affective processing (Hackel et al., 2015; Mende-Siedlecki et al., 2013). The recruitment of the DLPFC allows the tracking of another person's behavior for inferences on her intentions and planning on the adequate course of actions, a vital feature of a theory of mind. Research has found that people's reactions to unfair distributions are not only related to an individual observer's cognition, but also to that individual's emotions (Fehr & Fischbacher, 2003; Sanfey et al., 2003). In this study, the DLPFC function inhibition elicited different behaviors from participants that effectively equalized their sensitivity to normative unfairness. In Study 1, they rejected offers more readily, even offers in the gain context, to the point where sensitivity in the gain context mirrored what is typically expectable in the loss context. In Study 2, participants acting as judges punished dictators in the gain context just as often and as harshly as they would in the loss context, which was an as significant increase from what one would expect otherwise (Liu et al., 2019). These different actions shared the same cause-a bias resulting from the integration of information about unfair events, resulting in a loss of the distinction of the behavioral features of a specific norm violation, and an overall increase in the sensitivity to norm violations. Our course of rTMS also seemed to undeniably increase one's willingness to spend to harm the Proposer or Dictator. Perhaps the DLPFC is thus not only activated and involved in the tempered assessment of the situational context (gain vs. loss), but it may also relate to the forcible maintenance of social norms, and inhibited the individual's self-interest tendency (Knoch et al., 2008; Knoch et al., 2006). The inhibition of the DLPFC function produced the deviation of individual information integration ability and reduced third parties' sensitivity to the cost of punishment. It improved the third parties' willingness to maintain social norms and judge the perpetrators. This study further suggested that third-party punishment needs to mobilize and integrate more cognitive resources suitably. Although DLPFC function is inhibited, it still increases the individual's willingness to altruistic punishment. The results also further support the Strong reciprocity theory (Fehr & Fischbacher, 2003; Gintis et al., 2003) and Indirect reciprocity theory (Rand & Nowak, 2013). People are willing to act against norms that benefit society and individuals.

Our study has several limitations. First, we focused on the function of DLPFC in altruistic punishment decision-making and the differences in altruistic punishment decision-making between the gain and loss context. However, DLPFC is not only related to executive control—but is also related to emotional

processing. Future studies could investigate whether the DLPFC alters participants' emotional responses (perhaps using skin conductance or some other instruments). Secondly, in the second party punishment, we found the effects only following stimulation of the left dlPFC, but there was no significant difference between left and right dlPFC. In future studies, we need to explore the role of the right DLPFC in depth through different techniques and methods. Thirdly, we used low-frequency rTMS in this study. However, some researchers have found that the intervention effects of rTMS vary at different frequencies: low-frequency rTMS usually refers to magnetic stimulation less than or equal to 1Hz, and high-frequency rTMS refers to magnetic stimulation with a frequency higher than 1Hz (Rossi et al., 2009). The low-frequency rTMs can cause a sustained inhibitory effect in the stimulated brain regions of healthy participants, while high-frequency rTMs usually lead to a facilitation effect in the stimulated regions for a sustained period (Maeda et al., 2000). Future studies should apply high-frequency rTMS to the DLPFC and other regions to further explore how these structures drive altruistic punishment.

## **5 Conclusion**

In line with previous research, this study also provided the causal evidence that the DLPFC plays a pivotal role in modulating altruistic decision-making. It affected the difference in the decision-making of second-party and third-party altruistic punishment in the gain and loss contexts. Study 1 revealed that disrupting the function of the left DLPFC significantly raised the likelihood that the second party would reject unfair offers, where they are usually primarily more sensitive to unfair proposals regarding shouldering losses. The stimulation of the left and right DLPFC by rTMS seemed to have heightened the sensitivity to unfairness, especially the response to unfairness in the gain context, resulting in the difference in behavioral response across the two contexts (gain and loss) diminished. In Study 2, we inhibited the function of bilateral DLPFC by rTMS and found that in the gain context, third parties were more sensitive to unfairly distributed gains than they normally are and punished more forcibly as if they were unfairly shouldered losses. Thus, the inhibition of the DLPFC function will lead to the deviation of individual information integration ability and influence the moderating effect of gain and loss contexts on altruistic punishment.

## **Declarations**

### **Conflict of Interest**

The authors declare that they have no conflict of interest.

### **Ethical Approval**

All procedures performed in studies involving human participants were by the ethical standards of the institutional and/or national research committee and with the 1964 Helsinki Declaration and its later amendments or comparable ethical standards. This article does not refer to any studies conducted with animals.

## Informed Consent

Informed consent was obtained from all individual participants included in the study.

## Data accessibility statement

The authors are willing to share these data. Please contact the authors and they will send the data via email.

## References

1. Barbey AK, Colom R, Solomon J, Krueger F, Grafman J (2012) An integrative architecture for general intelligence and executive function revealed by lesion mapping. *Brain* 135(Pt 4):1154–1164. <http://doi.org/10.1093/brain/aws021>
2. Baumgartner T, Götte L, Gügler R, Fehr E (2012) The mentalizing network orchestrates the impact of parochial altruism on social norm enforcement. *Hum Brain Mapp* 33(6):1452–1469. <http://doi.org/10.1002/hbm.21298>
3. Baumgartner T, Heinrichs M, Vonlanthen A, Fischbacher U, Fehr E (2008) Oxytocin shapes the neural circuitry of trust and trust adaptation in humans. *Neuron* 58(4):639–650. <http://doi.org/10.1016/j.neuron.2008.04.009>
4. 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. <http://doi.org/10.1038/nn.2933>
5. Bellucci G, Camilleri JA, Eickhoff SB, Krueger F (2020) Neural signatures of prosocial behaviors. *Neurosci Biobehavioral Reviews* 118:186–195. <http://doi.org/10.1016/j.neubiorev.2020.07.006>
6. Buckholtz JW, Marois R (2012) The roots of modern justice: Cognitive and neural foundations of social norms and their enforcement. *Nat Neurosci* 15(5):655–661. <http://doi.org/10.1038/nn.3087>
7. Buckholtz JW, Martin JW, Treadway MT, Jan K, Zald DH, Jones O, Marois R (2015) From blame to punishment: Disrupting prefrontal cortex activity reveals norm enforcement mechanisms. *Neuron* 87(6):1369–1380. <http://doi.org/10.1016/j.neuron.2015.08.023>
8. David B, Hu Y, Krüger F, Weber B (2017) Other-regarding attention focus modulates third party altruistic choice: An fMRI study. *Sci Rep* 7:43024. <http://doi.org/10.1038/srep43024>
9. de Quervain DJ-F, Fischbacher U, Treyer V et al (2004) The neural basis of altruistic punishment. *Science* 305(5688):1254–1258. <http://doi.org/10.1126/science.1100735>
10. de Souza Bezerra E, Diefenthaler F, Sakugawa RL, Cadore EL, Izquierdo M, Moro ARP (2019) Effects of different strength training volumes and subsequent detraining on strength performance in aging adults. *J Bodyw Mov Ther* 23(3):466–472. <https://doi.org/10.1016/j.jbmt.2019.01.010>
11. Delis DC, Kaplan E, Kramer JH (2001) *Delis-Kaplan Executive Function System (DKEFS)*, San Antonio, TX, *The Psychological Corporation*. ISBN: 978-0-387-79947-6

12. Eisenegger C, Treyer V, Fehr E, Knöck D (2008) Time-course of “off-line” prefrontal rTMS effects: A PET study. *NeuroImage* 42(1):379–384. <http://doi.org/10.1016/j.neuroimage.2008.04.172>
13. Faul F, Erdfelder E, Lang AG, Buchner A (2007) G\* Power 3: A flexible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behav Res Methods* 39(2):175–191. <https://doi.org/10.3758/bf03193146>
14. Fehr E, Fischbacher U, Gächter S (2002) Strong reciprocity, human cooperation, and the enforcement of social norms. *Hum Nat – Interdisciplinary Biosocial Perspective* 13:1–25. <http://doi.org/10.1007/s12110-002-1012-7>
15. Fehr E, Fischbacher U (2003) The nature of human altruism. *Nature* 425(6960):785–791. <http://doi.org/10.1038/nature02043>
16. Fehr E, Fischbacher U (2004) Third-party punishment and social norms. *Evol Hum Behav* 25:63–87. [http://doi.org/10.1016/S1090-5138\(04\)00005-4](http://doi.org/10.1016/S1090-5138(04)00005-4)
17. Figner B, Knöck D, Johnson EJ, Krosch AR, Lisanby SH, Fehr E, Weber EU (2010) Lateral prefrontal cortex and self-control in intertemporal choice. *Nat Neurosci* 13(5):538–539. <http://doi.org/10.1038/nn.2516>
18. Folwaczny A, Waldmann E, Altenhofer J, Henze K, Parhofer KG (2021) Postprandial lipid metabolism in normolipidemic subjects and patients with mild to moderate hypertriglyceridemia: effects of test meals containing saturated fatty acids, mono-unsaturated fatty acids, or medium-chain fatty acids. *Nutrients* 13(5):1737–1747. <https://doi.org/10.3390/nu13051737>
19. Ginther MR, Bonnie RJ, Hoffman MB, Shen FX, Simons KW, Jones OD, Marois R (2016) Parsing the behavioral and brain mechanisms of third-party punishment. *J Neurosci* 36(36):9420–9434. <http://doi.org/10.1523/JNEUROSCI.4499-15.2016>
20. Gintis H, Bowles S, Boyd R, Fehr E (2003) Explaining altruistic behavior in humans. *Evol Hum Behav* 24:153–172. [http://doi.org/10.1016/S1090-5138\(02\)00157-5](http://doi.org/10.1016/S1090-5138(02)00157-5)
21. Guo X, Li Z, Zhu L, Li J, Wang Q, Dienes Z et al (2013) Increased neural responses to unfairness in a loss context. *NeuroImage* 77:246–253. <http://doi.org/10.1016/j.neuroimage.2013.03.048>
22. Güroğlu B, van den Bos W, Rombouts SARB, Crone EA (2010) Unfair? It depends: Neural correlates of fairness in social context. *Social Cogn Affect Neurosci* 5(4):414–423. <http://doi.org/10.1093/scan/nsq013>
23. Hackel LM, Doll BB, Amodio DM (2015) Instrumental learning of traits versus rewards: dissociable neural correlates and effects on choice. *Nat Neurosci* 18(9):1233. <http://doi.org/10.1038/nn.4080>
24. Halevy N, Halali E (2015) Selfish third parties act as peacemakers by transforming conflicts and promoting cooperation. *Proceedings of the National Academy of Sciences*, 112, 6937–6942. <https://doi.org/10.1073/pnas.1505067112>
25. Hartwigsen G, Price CJ, Baumgaertner A, Geiss G, Koehnke M, Ulmer S et al (2010) The right posterior inferior frontal gyrus contributes to phonological word decisions in the healthy brain: Evidence from dual site tms. *Neuropsychologia* 48(10):3155–3163. <http://doi.org/10.1016/j.neuropsychologia.2010.06.032>

26. Haushofer J, Fehr E (2008) You shouldn't have: Your brain on others' crimes. *Neuron* 60(5):738–740. [10.1016/j.neuron.2008.11.019](https://doi.org/10.1016/j.neuron.2008.11.019)
27. House BR, Kanngiesser P, Barrett HC, Yilmaz S, Smith AM, Sebastian-Enesco C, Erut A, Silk JB (2020) Social norms and cultural diversity in the development of third-party punishment. *Proceedings of the Royal Society B: Biological Sciences*, 28720192794
28. <http://doi.org/10.1098/rspb.2019.2794>
29. Kalbe E, Schlegel M, Sack AT, Nowak DA, Dafotakis M, Bangard C, ..., Kessler J (2010) Dissociating cognitive from affective theory of mind: a TMS study. *cortex*, 46(6), 769–780. <https://doi.org/10.1016/j.cortex.2009.07.010>
30. Kearney-Ramos TE, Dowdle LT, Lench DH, Mithoefer OJ, Devries WH, George MS, ..., Hanlon CA (2018) Transdiagnostic effects of ventromedial prefrontal cortex transcranial magnetic stimulation on cue reactivity. *Biol Psychiatry: Cogn Neurosci Neuroimaging* 3(7):599–609. <http://doi.org/10.1016/j.bpsc.2018.03.016>
31. Keel JC, Smith MJ, Wassermann EM (2001) A safety screening questionnaire for transcranial magnetic stimulation. *Clin Neurophysiology: Official J Int Federation Clin Neurophysiol* 112(4):720–720. [http://doi.org/10.1016/S1388-2457\(00\)00518-6](http://doi.org/10.1016/S1388-2457(00)00518-6)
32. Kirk U, Downar J, Montague PR (2011) Interoception drives increased rational decision-making in meditators playing the ultimatum game. *Front NeuroSci* 5:49–60. <http://doi.org/10.3389/fnins.2011.00049>
33. Knoch D, Nitsche MA, 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. <http://doi.org/10.1093/cercor/bhm237>
34. 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. [10.1126/science.1129156](https://doi.org/10.1126/science.1129156)
35. Koch G, Olivieri M, Torriero S, Caltagirone C (2003) Underestimation of time perception after repetitive transcranial magnetic stimulation. *Neurology* 60:1844–1846. <http://doi.org/10.1212/WNL.60.11.1844>
36. Krueger F, Hoffman M (2016) The emerging neuroscience of third-party punishment. *Trends Neurosci* 39(8):499–501. <http://doi.org/10.1016/j.tins.2016.06.004>
37. Latzman RD, Markon KE (2010) The factor structure and age-related factorial invariance of the delis-kaplan executive function system (d-kefs). *Assessment* 17(2):172–184. <http://doi.org/10.1177/1073191109356254>
38. Leliveld MC, Dijk EV, Beest IV (2012) Punishing and compensating others at your own expense: The role of empathic concern on reactions to distributive injustice. *Eur J Social Psychol* 42:135–140. <http://doi.org/10.1002/ejsp.872>
39. Liu Y, Li L, Zheng L, Guo X (2017) Punish the perpetrator or compensate the victim? Gain vs. Loss context modulate third-party altruistic behaviors. *Front Psychol* 8:2066–2077. <http://doi.org/10.3389/fpsyg.2017.02066>



40. Liu Y, Wang H, Li L, Tian T, Li J, DiFabrizio B, Liu Y (2021) Social status affects how third parties assess unfairly shared losses and unfairly shared gains. *Scand J Psychol* 62(6):858–870
41. Liu Y, Wang H, Li L, Wang Y, Peng J, Baxter DF (2019) Judgments in a hurry: Time pressure affects how judges assess unfairly shared losses and unfairly shared gains. *Scand J Psychol* 60(3):203–212. <http://doi.org/10.1111/sjop.12532>
42. Maeda F, Keenan JP, Tormos JM et al (2000) Interindividual variability of the modulatory effects of repetitive transcranial magnetic stimulation on cortical excitability. *Exp Brain Res* 133:425–430. <http://doi.org/10.1007/s002210000432>
43. Mende-Siedlecki P, Cai Y, Todorov A (2013) The neural dynamics of updating person impressions. *Soc Cognit Affect Neurosci* 8(6):623–631. <http://doi.org/10.1093/scan/nss040>
44. Moll J, Zahn R, de Oliveira-Souza R, Krueger F, Grafman J (2005) Opinion: The neural basis of human moral cognition. *Nat Rev Neurosci* 6(10):799–809. <http://doi.org/10.1038/nrn1768>
45. Mylius V, Ayache SS, Ahdab R, Farhat WH, Zouari HG, Belke M et al (2013) Definition of DLPFC and M1 according to anatomical landmarks for navigated brain stimulation: Inter-rater reliability, accuracy, and influence of gender and age. *NeuroImage* 78:224–232. <http://doi.org/10.1016/j.neuroimage.2013.03.061>
46. Ochsner KN, Bunge SA, Gross JJ, Gabrieli JD (2002) Rethinking feelings: An fMRI study of the cognitive regulation of emotion. *J Cogn Neurosci* 14(8):1215–1229. <http://doi.org/10.1162/089892902760807212>
47. Oldfield RC (1971) The assessment and analysis of handedness: The Edinburgh inventory. *Neuropsychologia* 9(1):97–113. <http://doi.org/10.3152/147154602781766852>
48. Org Z (2010) Moral sentiments and material interests: the foundations of cooperation in economic life. *Am Anthropol* 109(2):380–381. <http://doi.org/10.1525/aa.2007.109.2.380>
49. Oswald ME, Orth U, Aeberhard M, Schneider E (2010) Punitive reactions to completed crimes versus accidentally uncompleted crimes. *J Appl Soc Psychol* 35:718–731. <https://doi.org/10.1111/j.1559-1816.2005.tb02143.x>
50. Rand DG, Nowak MA (2013) Human cooperation. *Trends Cogn Sci* 17(8):413–425. <http://doi.org/10.1016/j.tics.2013.06.003>
51. Ren W, Wang M, Zheng Z, Zhang Z (2015) The role of the right dorsolateral prefrontal cortex in time perception: TMS evidence. *Journal of Psychological Science*, 38(5), 1051–1057. <http://doi.org/CNKI:SUN:XLKX.0.2015-05-005>
52. Rossi S, Hallett M, Rossini PM, Pascual-Leone A (2009) Safety, ethical considerations, and application guidelines for the use of transcranial magnetic stimulation in clinical practice and research. *Clin Neurophysiol* 120(12):2008–2039. <http://doi.org/10.1016/j.clinph.2009.08.016>
53. Ruff CC, Fehr E (2014) The neurobiology of rewards and values in social decision making. *Nat Rev Neurosci* 15:549–562. <http://doi.org/10.1038/nrn3776>
54. Sääksvuori L, Mappes T, Puurtinen M (2011) Costly punishment prevails in intergroup conflict. *Proceedings of the Royal Society B: Biological Sciences*, 278(1723), 3428–3436.

<http://doi.org/10.1098/rspb.2011.0252>

55. Sanfey, Alan G, Rilling, James K, Aronson, Jessica A et al (2003) The neural basis of economic decision-making in the ultimatum game. *Science* 300:1755–1758.  
<http://doi.org/10.1126/science.1082976>
56. Tabibnia G, Satpute AB, Lieberman MD (2008) The sunny side of fairness. *Psychol Sci* 19(4):339–347. <http://doi.org/10.2307/40064936>
57. Treadway MT, Buckholtz JW, Martin JW, Jan K, Asplund CL, Ginther MR et al (2014) Corticolimbic gating of emotion-driven punishment. *Nat Neurosci* 17(9):1270–1275.  
<http://doi.org/10.1038/nn.3781>
58. Wu Y, Yu H, Shen B, Yu R, Zhou Z, Zhang G et al (2014) Neural basis of increased costly norm enforcement under adversity. *Soc Cognit Affect Neurosci* 9(12):1862–1871.  
<http://doi.org/10.1093/scan/nst187>
59. Zhou X, Wu Y (2011) Sharing losses and sharing gains: Increased demand for fairness under adversity. *J Exp Soc Psychol* 47(3):582–588. <http://doi.org/10.1016/j.jesp.2010.12.017>

## Tables

Table 2-1 The decision rate of the second-party punishment under different rTMS stimuli and the gain and loss contexts ( $M \pm SD$ )

		Rejection rate
Context	Gain	0.71 ± 0.24
	Loss	0.86 ± 0.20
rTMS stimulus	Left	0.83 ± 0.20
	Right	0.81 ± 0.20
	Sham	0.71 ± 0.20

Table 3-1 The decision rate and amount of the third-party punishment under different rTMS stimuli and the gain and loss contexts ( $M \pm SD$ )

		Punishment rate	Punishment amount
Context	Gain	0.85 ± 0.22	20.65 ± 4.71
	Loss	0.94 ± 0.13	22.15 ± 3.63
rTMS stimulus	Left	0.97 ± 0.14	21.60 ± 3.39
	Right	0.94 ± 0.14	24.60 ± 3.39
	Sham	0.77 ± 0.14	18.01 ± 3.39

# Figures

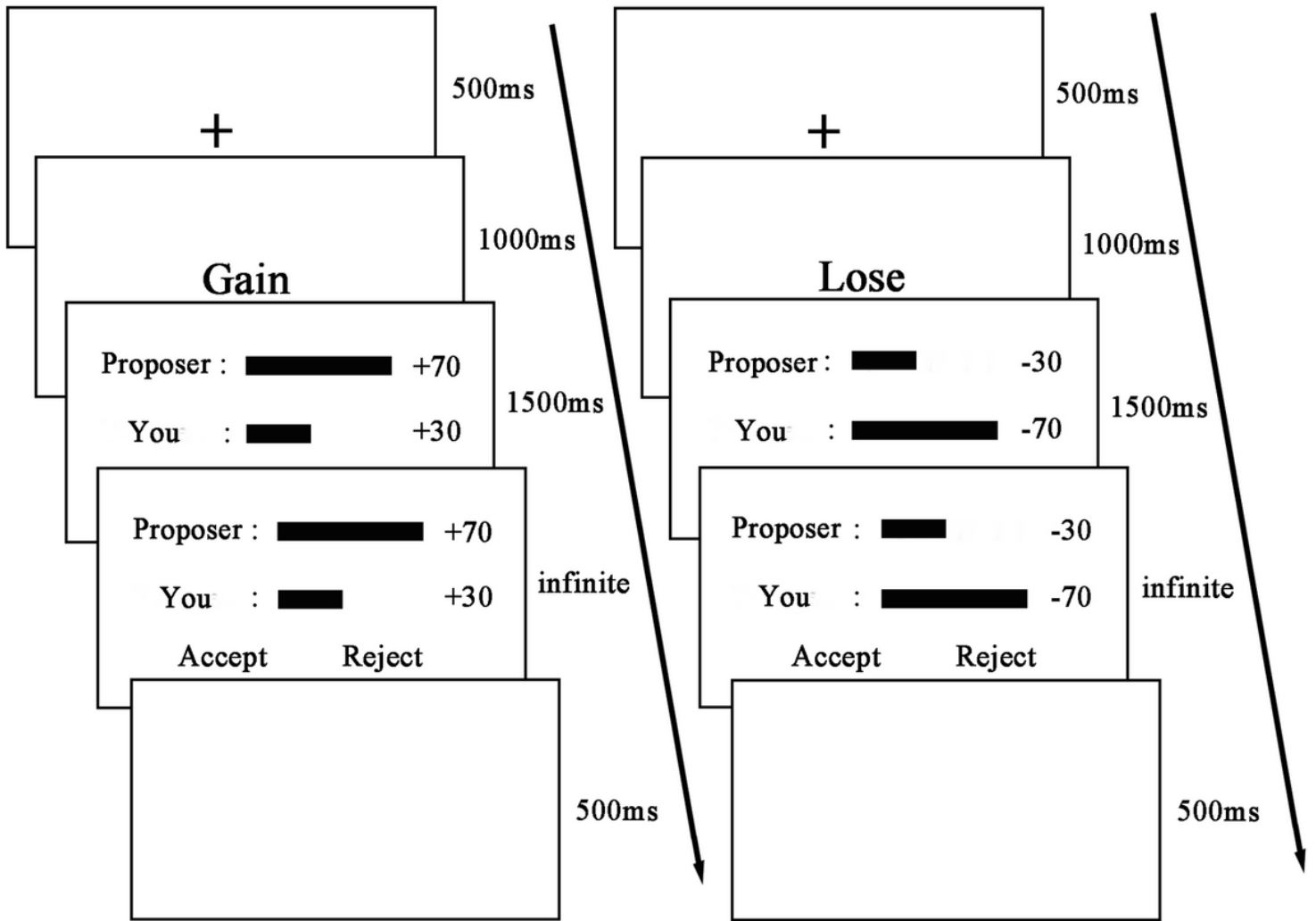
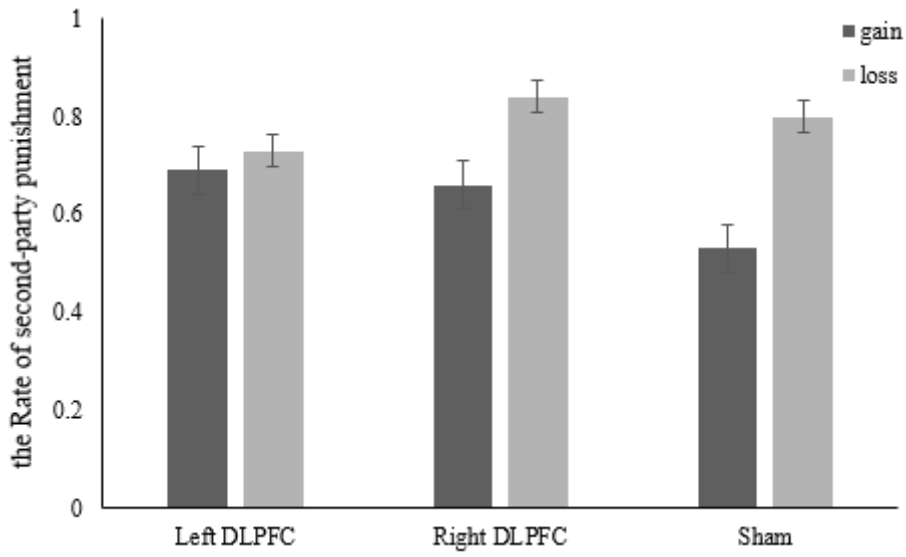


Figure 1

Flow chart of experimental trials in Study 1.



**Figure 2**

Different rTMS stimuli affect the decision-making rate of second-party punishment in the gain and loss context. Error bars = SEM.

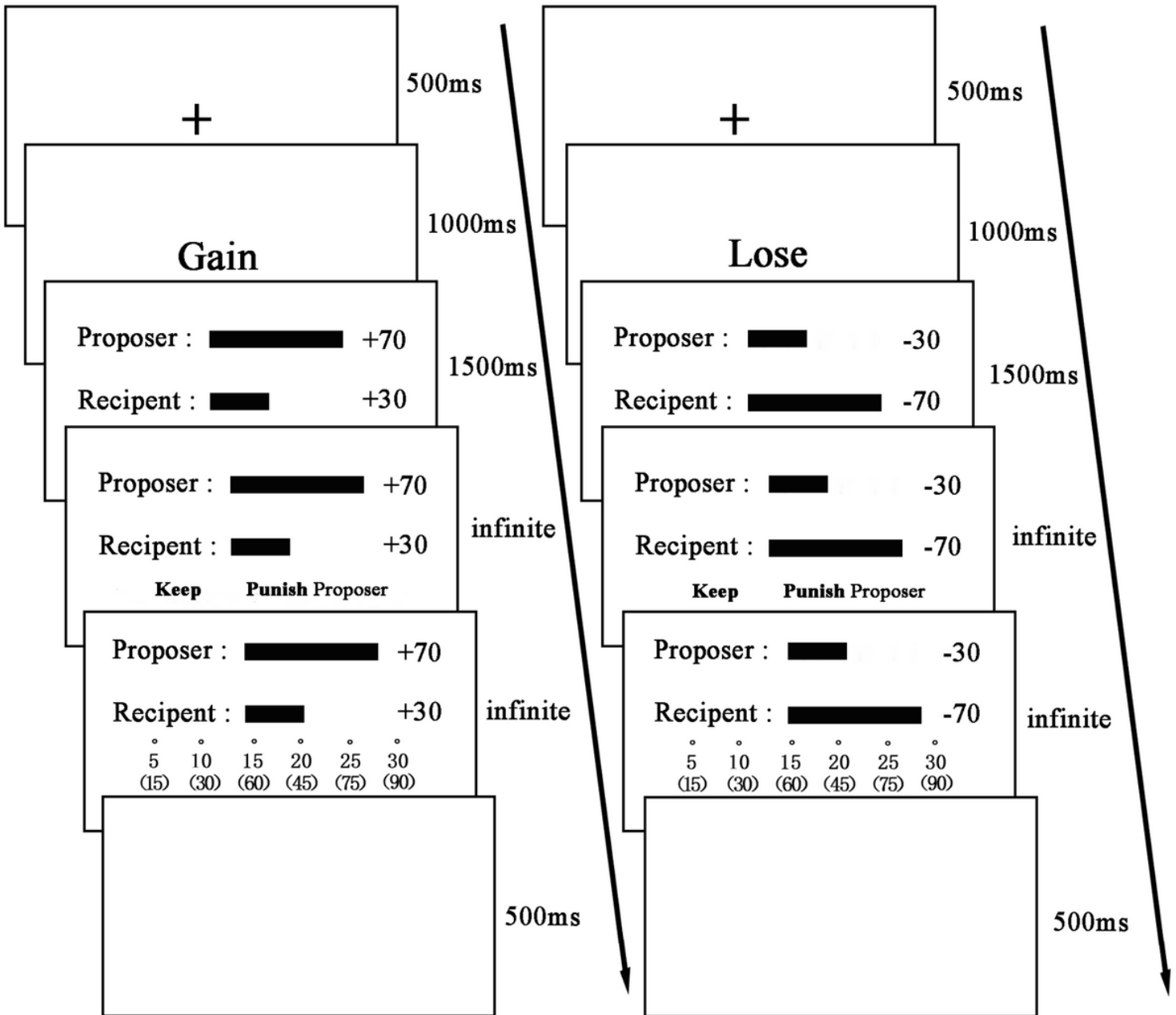
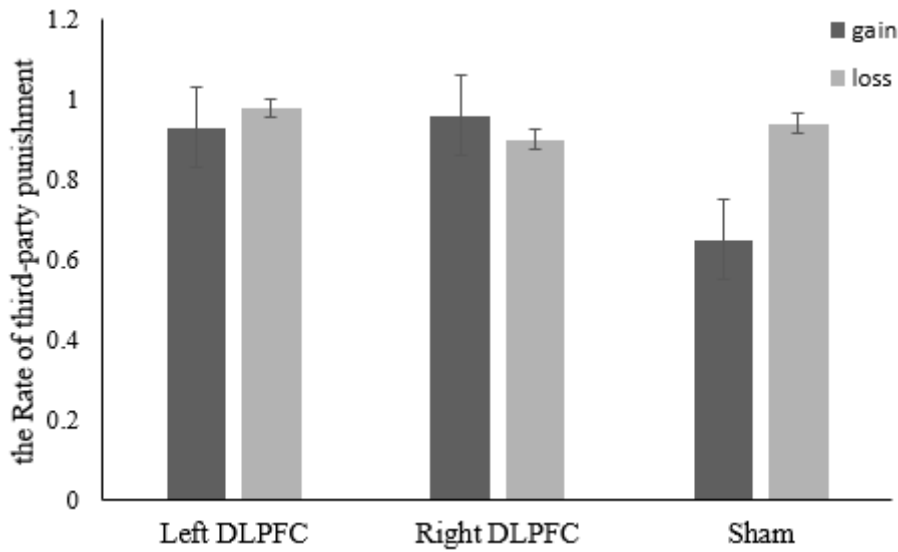


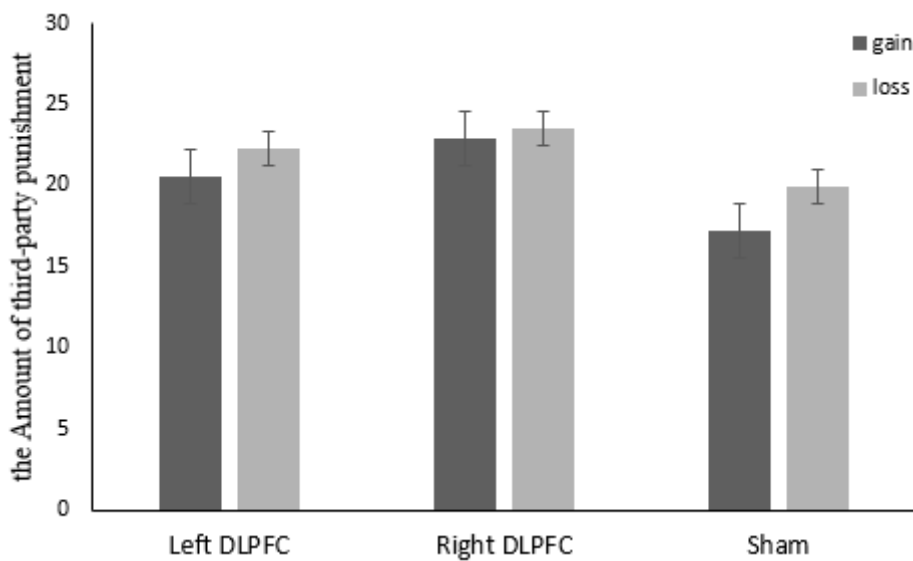
Figure 3

Flow chart of experimental trials in Study 2.



**Figure 4**

Different rTMS stimuli affect the decision-making rate of third-party punishment in the gain and loss context. Error bars = SEM.



**Figure 5**

Different rTMS stimuli affect the decision-making amount of third-party punishment in the gain and loss context. Error bars = SEM.