# Lateral prefrontal/orbitofrontal cortex has different roles in norm compliance in gain and loss domains: a transcranial direct current stimulation study

Yunlu Yin,[1,*] Hongbo Yu,[1,2,*] Zhongbin Su,[1] Yuan Zhang[1,3] and Xiaolin Zhou[1,4,5,6]

with and without sanction) and functional MRI, Spitzer et al. (2007) showed that when asked to allocate a certain amount of monetary

polarity on the target brain area depended on the central electrode. The current distribution under HD-tDCS has been partially validated by empirical data through a MRI-guided finite element model (Datta et al., 2009; Edwards et al., 2013), and recent studies showed that current density of HD-tDCS falls off with increasing cortical depth (Datta et al., 2009). The current intensity was 2.0 mA which created ~0.5 mA/cm$^2$ peak current density at the central electrode, and 0.125 mA/cm$^2$ peak current density at the return electrodes. Stimulation started 8 min before the task, and was delivered during the entire course of the task (~20 min) with an additional 30-s ramp-up at the beginning of stimulation and 30-s ramp-down at the end. The placement of electrodes was the same for the sham and the cathodal stimulation. However, for the sham stimulation, the initial 30 s ramp-up was immediately followed by the 30-s ramp-down, and there was no stimulation for the rest of the session (cf. Gandiga et al., 2006; Douglas et al., 2015). For both the cathodal and sham stimulation conditions, participants felt a little uncomfortable initially, but gradually the feelings associated with stimulation became negligible before the task started, according to our post-experiment interview.

Compared with the conventional bipolar tDCS, HD-tDCS has been shown to have better spatial focality and prolonged effect (Datta et al., 2009; Caparelli-Daquer et al., 2012; Kuo et al., 2013; Shen et al., 2016). Although HD-tDCS is associated with stronger

from the allocator. If the amount the allocator (i.e., the participant) allocated to the receiver was less than that minimum amount, a sanction may or may not be imposed on the allocator, depending on a prior decision by the computer (see below). To avoid learning effect, no feedback of earning/loss or sanction was provided. The participant was also told that a gain round and a loss round would be randomly chosen and realized after the experiment; this was to motivate the participant to treat each round equally and independently.

Each round began with the presentation of a white fixation cross against a black background, lasting for 4000 to 6000 ms with a step of 400 ms (Fig. 1). Then a cue of the total allocation amount (a picture of 20 yuan bill) was presented for 2000 ms, followed by a sentence indicating that punishment threat would be randomly decided by the computer for this trial. This sentence remained on the screen for 2000–5000 ms (with a step of 400 ms). Then the decision (Waive vs. Retain) together with a picture of computer were presented on the screen for 3000 ms. Specifically, 'Waive' means the computer decides that no sanction will be imposed on the current round, so the participant can allocate as she wishes without worrying about sanction. 'Retain' means the computer decides to keep the sanction threat on the current trial. In that case, if the participant's allocation was less than the minimum expectation given by the receiver, the participant would receive a sanction (although he/she did not know whether he/she was actually sanctioned in a given trial). Finally, after a 2000-to-4000-ms fixation, a distribution screen was presented. The participant was required to make the allocation within 10 s by pressing two buttons to adjust the allocation amount with a step of 2 yuan and a third button to confirm the allocation. The allocation was directed to the receiver so that in the gain context the positive points allocated to the receiver would be added to the receiver's account, while in the loss context, the negative points allocated to the receiver would be deducted from the partner's account. Button press was counterbalanced across participants. The initial amount on the side of the participant was either 0 or 20 yuan (0 or 20 yuan in the loss context) and was counterbalanced across conditions.

The allocation task consisted of a gain block and a loss block,

indicating that lPFC/lOFC may not play a direct role in mediating norm compliance in the loss-sharing situation.

To view the data from another perspective, we calculated the degree of threat-induced (or strategic) compliance for each domain and each treatment group by subtracting the degree of voluntary compliance from that of the compliance under sanction threat (Retain-Waive; cf. Ruff et al., 2013). This analysis is not independent of the above analysis for the data in Fig. 2A, but it allows us to make cross-study comparison (e.g., Ruff et al., 2013). It is clear from Fig. 2B that disruption of the lPFC/lOFC function reduced the threat-induced compliance in the gain domain, but not in the loss domain. To compare threat-induced compliance between the current study and that in Ruff et al. (2013) (termed "sanction-induced compliance" there), we carried out a two (context: Gain vs. Loss) by two (treatment: sham vs. cathodal) mixed-design ANOVA for threat-induced compliance. The two-way interaction between context and treatment was significant, $F_{1,57} = 6.08$, $P = 0.017$, partial $\eta^2 = 0.10$. Pairwise comparison showed that in the gain domain, cathodal tDCS significantly reduced threat-induced compliance, $t(57) = 3.16$, $P = 0.003$ after Bonferroni-correction for multiple comparison, replicating Ruff et al. (2013; see their Fig. 2A). By contrast, the tDCS effect was not significant for the loss domain, $t(57) = 0.34$, $P > 0.1$. Viewed in an alternative way, the difference in threat-induced compliance between gain and loss domains was significant only in the sham group, $t(28)$, $P = 0.018$, not in the cathodal group $t(29) = 0.98$, $P > 0.1$. This indicated that the flexibility in adjusting one's strategy across contexts relies causally on the function of lPFC/lOFC.

## Fairness perception is not affected by gain-loss context or tDCS condition

To test whether participants' perception of fairness norm was affected by gain–loss context and tDCS condition, we carried out a three-way ANOVA with time (before vs. after experiment), context (gain vs. loss), and tDCS treatment (sham vs. cathodal) as independent variable, and the perceived fairness ratings as dependent variable. Due to a technical error, the perceived fairness ratings from 10 participants in the cathodal group and nine participants in the sham

rejection rates in the loss context than in the gain context, suggesting that they were more willing to suffer personal cost to punish norm violators in the loss context. Using functional MRI, Yu et al. (2014) further demonstrate that rejecting unfair offers in the loss domain activate the dorsal striatum, an indication of rewarding and satisfactory experience (see also De Quervain et al., 2004; Crockett et al., 2013). It is thus clear from these studies that people have higher demand for fairness in the loss-sharing context. It is possible that in the current study, the participants were (implicitly or explic-

and such requirement is abolished in the loss context, probably because other motivations (e.g., enhanced fairness demand or harm/guilt aversion) become prominent in loss domain.

## Acknowledgement

This work was supported by the Natural Science Foundation of China (Grant 31630034) and the National Basic Research Program of China (973 Program: 2015CB856400). HY was supported by a Newton International Fellowship from the British Academy (NF160700).

## Conßict of interest

The authors declare no conßict of interest.

## Author contributions

YY, HY, YZ, and XZ contributed to the design of the study and

Ruff, C.C., Ugazio, G. & Fehr, E. (2013) Changing social norm compliance with noninvasive brain stimulation. *Science*, 342, 482–484.
Rushworth, M.F.S., Behrens, T.E.J., Rudebeck, P.H. & Walton, M.E. (2007) Contrasting roles for cingulate and orbitofrontal cortex in decisions and social behaviour. *Trends Cogn. Sci.*, 11, 168–176.
Rutledge, R.B., Smittenaar, P., Zeidman, P., Brown, H.R., Adams, R.A., Lindenberger, U., Dayan, P. & Dolan, R.J. (2016) Risk taking for potential reward decreases across the lifespan. *Curr. Biol.*, 26, 1634–1639.
Shen, B., Yin, Y., Wang, J., Zhou, X., McClure, S.M. & Li, J. (2016) High-definition tDCS alters impulsivity in a baseline-dependent manner. *Neuroimage*, 143, 343–352.
Sober, E. & Wilson, D.S. (1998) *Unto Others: the Evolution and Psychology of Unselfish Behavior*. Harvard University Press, Cambridge, MA.
Spitzer, M., Fischbacher, U., Herrnberger, B., Grön, G. & Fehr, E. (2007) The neural signature of social norm compliance. *Neuron*, 56, 185–196.
Strang, S., Gross, J., Schuhmann, T., Riedl, A., Weber, B. & Sack, A. (2015) Be nice if you have to – the neurobiological roots of strategic fairness. *Soc. Cogn. Affect. Neur.*, 10, 790–796.
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, 515–518.
Tversky, A. & Kahneman, D. (1981) The framing of decisions and the psychology of choice. *Science*, 211, 453–458.
Tversky, A. & Kahneman, D. (1991) Loss aversion in riskless choice: a reference-dependent model. *Quart. J. Econ.*, 106, 1039–1061.