Supplementary materials for

Learning whom to cooperate with: neurocomputational mechanisms for choosing cooperative partners

Tao Jin[†], Shen Zhang[†], Patricia Lockwood, Iris Vilares, Haiyan Wu, Chao Liu[*], Yina Ma[*]

[*]Chao Liu, State Key Laboratory of Cognitive Neuroscience and Learning, Beijing Normal University, No. 19 Xinjiekouwai Street, Beijing, 100875, China

E-mail: liuchao@bnu.edu.cn

[*]Yina Ma, State Key Laboratory of Cognitive Neuroscience and Learning, Beijing Normal University, No. 19 Xinjiekouwai Street, Beijing, 100875, China

E-mail: yma@bnu.edu.cn

† These authors contributed equally to this work

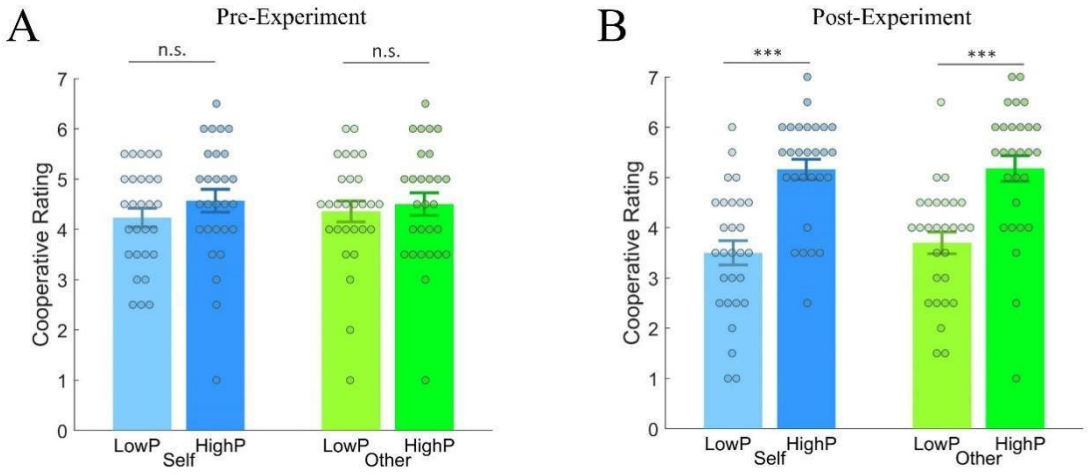

Fig S1. ***Cooperative ratings before (A) and after the experiment (B).*** An ANOVA on the cooperative rating, with time (pre- *vs.* post-experiment), agent (self *vs.* other), and the propensity of partners' cooperation (low *vs*. high cooperative), showed a significant main effect of the propensity of partners' cooperation ($F(1, 27) = 35.82$, $p < 0.001$, partial $\eta^2 = 0.57$) and a significant interaction between time and the propensity of partners' cooperation ($F(1, 27) = 21.84$, $p < 0.001$, partial $\eta^2 = 0.45$). **(A)** Before the experiment, the ratings toward partners with high and low cooperative propensities were similar (self: $p = 0.134$; other: $p = 0.518$). **(B)** After repeated interactions, participants rated partners with high cooperativeness significantly higher than those of low cooperativeness (self: $p < 0.001$; other: $p < 0.001$). (Data are shown as the mean ± SEM with overlaid dot plots. n.s. not significant; ***, $p < 0.001$).

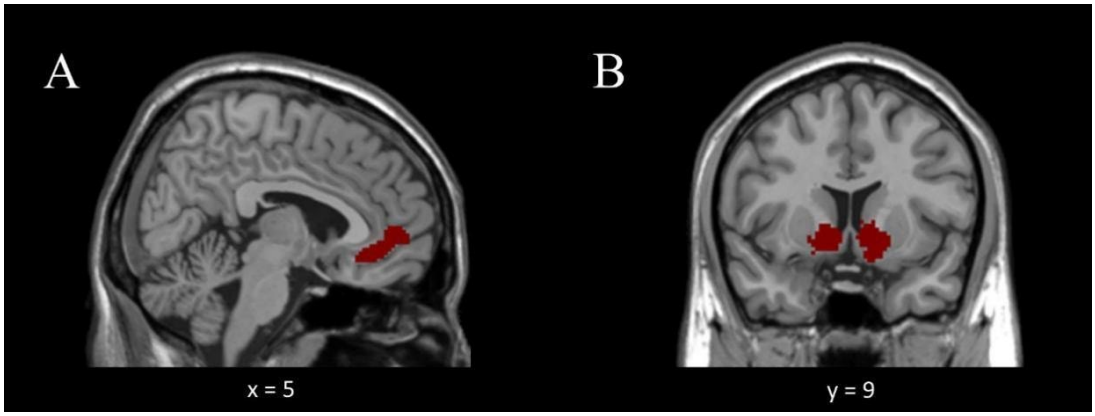

Fig S2. ***The regions of interest.*** **(A)** The pregenual anterior cingulate cortex (pgACC) was defined based on a previous related study (Lau et al., 2020). **(B)** The bilateral striatum was defined from term-based meta-analysis of "prediction error" in Neurosynth with the number of voxels more than 55 to exclude several voxels irrelevant to striatum (Yarkoni et al., 2011).

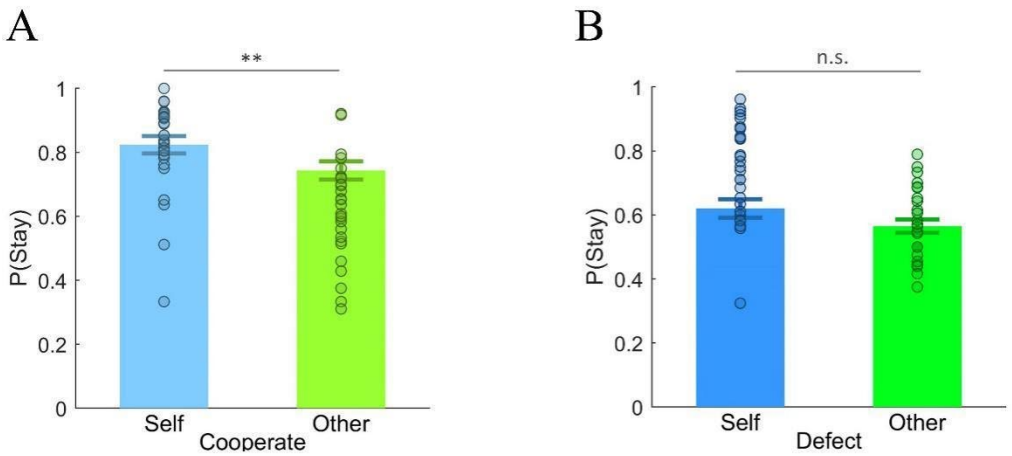

Fig S3. (A) When the chosen partner cooperated in the last trial, participants were more likely to stay with the previously chosen partner in the self condition than in the other condition (p = 0.005). **(B)** After the chosen partner defected in the last trial, there was no significant difference in the proportions of staying with the chosen partner in both conditions (p = 0.077). (Data are shown as the mean ± SEM with overlaid dot plots. n.s., non-significant; **, $p < 0.01$).

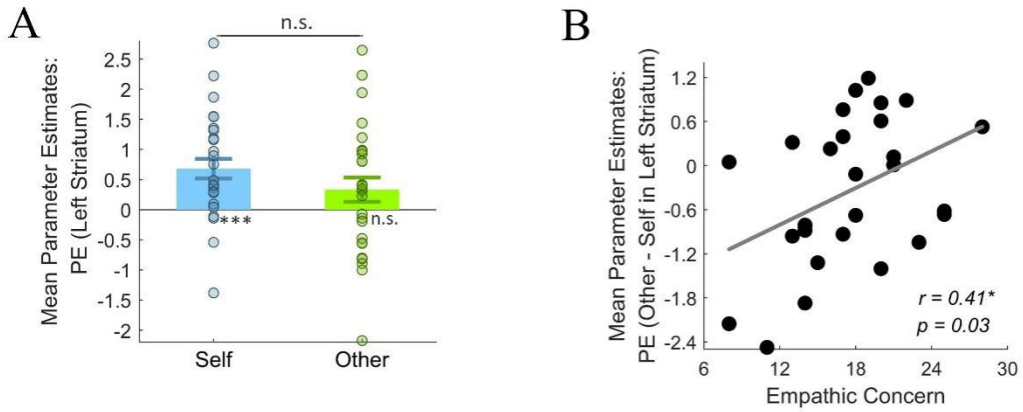

Fig S4. (A) The Region of Interest (ROI) analysis showed that the left striatum tracked prediction errors exclusively in the self condition, but not in the other condition. **(B)** Correlation analysis revealed the left striatal activity related to prediction errors of another person (*vs.* self) was positively associated with empathic concern. People with lower empathic concern showed lower left striatal activity to other-regarding PE (*vs.* self-regarding PE). (n.s, non-significant; * $p < 0.05$; *** $p < 0.001$.)

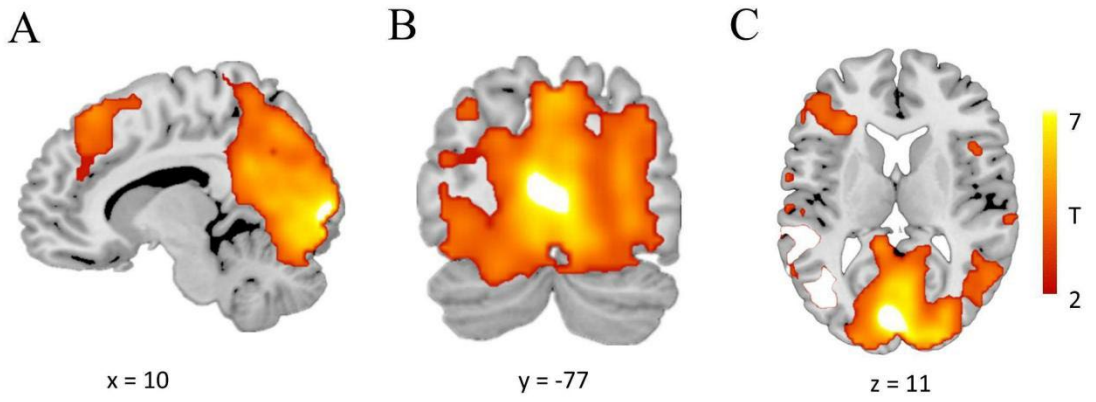

Fig S5. Using the standard cluster-forming threshold of $p < 0.001$ led to a large cluster of 147,090 voxels spanning multiple brain regions, including occipital, parietal, frontal and temporal lobes.

Table S1. ***Model-based fMRI analysis: brain areas associated with prediction errors***

Brain Regions	Side	MNI Coordination of Local Maxima (mm)			Local Maxima T/F	Cluster Size (voxel)	p-value
		x	y	z			
Self Condition: Cooperative PE							
pgACC	L/R	0	42	6	6.68	602	p<0.001
Striatum	L	-12	9	-3	6.30	61	p=0.047
Precentral gyrus	L	-6	-33	54	5.91	641	*p*<0.001
Cuneus/Middle Occipital Gyrus	R	18	-93	12	5.47	159	*p*<0.001
Extra-Nuclear/Putamen	L	-27	-15	0	5.34	193	*p*<0.001
Occipital Lobe	L	-15	-96	9	4.84	62	*p*=0.045
Other Conditionl: Cooperative PE							
pgACC	R	9	30	9	4.69	128	*p*=0.002

R, right; L, left; PE, prediction error. ACC, anterior cingulate cortex. Activations are significant in all above regions at $p<0.05$ FWE-whole brain corrected at the cluster level after voxel-wise thresholding at $p<0.001$

Table S2. ***MVPA results: activity patterns discriminating self- vs. other-related feedback in the whole-brain searchlight analysis***

Brain Regions	Hemisphere	MNI Coordination of Local Maxima (mm)			Local Maxima	Cluster size (voxel)	*p*-value
		x	y	z	T		
temporal parietal junction	R	45	-30	21	5.58	543	$p = 0.001$
striatum/insula	L	-21	6	6	5.14	401	$p = 0.004$
superior parietal lobe/precuneus	L	-18	-51	57	5.35	689	$p = 0.001$
posterior cingulate cortex/ corpus callosum	L	-6	-42	6	5.04	225	$p = 0.029$
middle cingulate cortex /superior frontal gyrus	L/R	0	21	45	5.04	526	$p = 0.001$
superior temporal gyrus	L	-48	-21	6	4.49	196	$p = 0.042$

R, right; L, left. Activations are significant in all above regions at $p<0.05$ FWE-whole brain corrected at the cluster level after voxel-wise thresholding at $p<0.001$.

Reference

Lau, T., Gershman, S. J., Cikara, M. (2020). Social structure learning in human anterior insula, *eLife*, 9, 1–17.

Yarkoni, T., Poldrack, R. A., Nichols, T. E., Van Essen, D. C., Wager, T. D. (2011). Large-scale automated synthesis of human functional neuroimaging data, *Nat. Methods*, 8, 665–670.