

## Original Article

**Cite this article:** Deng G, Ai H, Qin L, Xu J, Feng C, Xu P (2023). Dissociated modulations of intranasal vasopressin on prosocial learning between reward-seeking and punishment-avoidance. *Psychological Medicine* **53**, 5415–5427. <https://doi.org/10.1017/S0033291722002483>

Received: 22 December 2021

Revised: 8 June 2022

Accepted: 18 July 2022

First published online: 19 August 2022

**Key words:**

arginine vasopressin (AVP); brain oscillations; computational modeling; event-related potential (ERP); prosocial learning; punishment-avoidance; reward-seeking

**Authors for correspondence:**


Pengfei Xu,

E-mail: [pxu@bnu.edu.cn](mailto:pxu@bnu.edu.cn);

Chunliang Feng,

E-mail: [chunliang.feng@m.scnu.edu.cn](mailto:chunliang.feng@m.scnu.edu.cn)

# Dissociated modulations of intranasal vasopressin on prosocial learning between reward-seeking and punishment-avoidance

Guangzhi Deng<sup>1</sup> , Hui Ai<sup>2</sup>, Lili Qin<sup>2</sup>, Jie Xu<sup>2</sup>, Chunliang Feng<sup>3,4</sup> and Pengfei Xu<sup>1,5</sup>

<sup>1</sup>Beijing Key Laboratory of Applied Experimental Psychology, National Demonstration Center for Experimental Psychology Education (BNU), Faculty of Psychology, Beijing Normal University, Beijing, China; <sup>2</sup>Shenzhen Key Laboratory of Affective and Social Neuroscience, Magnetic Resonance Imaging Center, Center for Brain Disorders and Cognitive Sciences, Shenzhen University, Shenzhen, China; <sup>3</sup>Key Laboratory of Brain, Cognition and Education Sciences, Ministry of Education (South China Normal University), Guangzhou, China; <sup>4</sup>School of Psychology, Center for Studies of Psychological Application, Guangdong Key Laboratory of Mental Health and Cognitive Science, South China Normal University, Guangzhou, China and <sup>5</sup>Center for Emotion and Brain, Shenzhen Institute of Neuroscience, Shenzhen, China

**Abstract**

**Background.** As an integral ingredient of human sociality, prosocial behavior requires learning what acts can benefit or harm others. However, it remains unknown how individuals adjust prosocial learning to avoid punishment or to pursue reward. Given that arginine vasopressin (AVP) is a neuropeptide that has been involved in modulating various social behaviors in mammals, it could be a crucial neurochemical facilitator that supports prosocial learning.

**Methods.** In 50 placebo controls and 54 participants with AVP administration, we examined the modulation of AVP on the prosocial learning characterized by reward and punishment framework, as well as its underlying neurocomputational mechanisms combining computational modeling, event-related potentials and oscillations.

**Results.** We found a self-bias that individuals learn to avoid punishment asymmetrically more severely than reward-seeking. Importantly, AVP increased behavioral performances and learning rates when making decisions to avoid losses for others and to obtain gains for self. These behavioral effects were underpinned by larger responses of stimulus-preceding negativity (SPN) to anticipation, as well as higher punishment-related feedback-related negativity (FRN) for prosocial learning and reward-related P300 for proself benefits, while FRN and P300 neural processes were integrated into theta (4–7 Hz) oscillation at the outcome evaluation stage.

**Conclusions.** These results suggest that AVP context-dependently up-regulates altruism for concerning others' losses and reward-seeking for self-oriented benefits. Our findings provide insight into the selectively modulatory roles of AVP in prosocial behaviors depending on learning contexts between proself reward-seeking and prosocial punishment-avoidance.

**Introduction**

Prosocial behavior refers to the behavior that intends to benefit others in social interactions (Fehr & Fischbacher, 2003; Fiske & Taylor, 2013; Moskowitz, 2005). To behave prosocially, humans need to process social information and learn about the impacts that their actions can have on others (Fehr & Fischbacher, 2003; Fiske & Taylor, 2013; Lockwood, Apps, Valton, Viding, & Roiser, 2016). Reinforcement learning theory provides a powerful framework for understanding how humans and other species form action–outcome associations (Sutton & Barto, 2018). Recent evidence has shown that prosocial behaviors can be described in terms of reinforcement learning when people learn to benefit themselves (self-oriented learning) and others (prosocial learning) (Lockwood et al., 2016), suggesting there is a self-bias that humans learn faster from feedback to reward themselves than others (Lockwood et al., 2016; Martins, Lockwood, Cutler, Moran, & Paloyelis, 2022).

Learning is one of the most crucial abilities of our brain to adapt to social life (Alberts, 1994; van den Berg, Molleman, & Weissing, 2015). Individuals learn to make decisions that maximize personal utilities or obtain necessary supplies. Yet, it has long been recognized that many real-world decisions are made in a social context, i.e. choices involving not only personal goals, but also potential benefits for others. Recent studies have shown that people can behave prosocially or egoistically by learning the consequences of their decisions for others and themselves from reward feedback (Liao, Huang, & Luo, 2021; Lockwood et al., 2016). Reward and punishment represent crucial elements of reinforcement learning – positive and negative feedbacks have dissociable effects on learning (Galea, Mallia, Rothwell, & Diedrichsen, 2015) and may bias human independent estimates of the information content

(Pulcu & Browning, 2017). Dissociable effects of positive and negative feedback also exhibit context-dependent modulations on decision-making. For example, people tend to display an optimistic bias toward self-relevant beliefs, updating their beliefs to a greater extent following positive than negative feedback (Sharot & Garrett, 2016), but also prioritize learning for concerning the suffering of others and avoiding harming others from punishment feedback (Crockett, Kurth-Nelson, Siegel, Dayan, & Dolan, 2014; Lockwood, Klein-Flügge, Abdurahman, & Crockett, 2020). However, it remains largely unclear on the role of the self/other-orientation in prosocial learning, especially learning from punishment. Acknowledging more about how we learn prosocial behavior in different social contexts may help us understand the atypical behaviors across psychiatric conditions such as antisocial behavior (Lock, 2008) and autism spectrum disorders (Apps, Rushworth, & Chang, 2016; Lockwood *et al.*, 2016).

It is critical to understand the neurochemical systems and neurocomputational mechanisms of prosocial learning. As an evolutionarily conserved neuropeptide, arginine vasopressin (AVP) modulates various social behaviors in mammals (Caldwell, 2017; Winslow, Hastings, Carter, Harbaugh, & Insel, 1993). Animal studies have illustrated the role of AVP in social memory (Albers, 2015; Caldwell, 2017), social communication/recognition (Song, Larkin, Malley, & Albers, 2016; Song *et al.*, 2014), aggression (Caldwell & Albers, 2004; Gobrogge, Liu, Young, & Wang, 2009), and pair bonding (Liu, Curtis, & Wang, 2001; Pitkow *et al.*, 2001). Not only aggressive but also prosocial behaviors can be modulated by AVP in specific social contexts. Human genetic studies have revealed the link between AVP system and complex human social behaviors, such that polymorphisms of the human AVP receptor gene (AVPR1A) have been associated with reciprocity and trust (Nishina, Takagishi, Takahashi, Sakagami, & Inoue-Murayama, 2019) as well as altruistic behavior (Avinun *et al.*, 2011; Knafo *et al.*, 2008; Wang *et al.*, 2016). Importantly, the intranasal administration of AVP has been widely applied to humans to reveal the causal role of AVP in human social cognition and is considered as an effective means to directly affect central processes through the blood-brain barrier (Born *et al.*, 2002; Dhuria, Hanson, & Frey, 2010). For instance, intranasal AVP regulates the auditory attention, perception, and memory of emotional and social cues (Dodt *et al.*, 1994; Uzefovsky, Shalev, Israel, Knafo, & Ebstein, 2012; Zink *et al.*, 2011), as well as risky decision making, cooperation, and prosocial behaviors (Feng, Qin, Luo, & Xu, 2020; Feng *et al.*, 2015; Neto *et al.*, 2020; Patel *et al.*, 2015; Rilling *et al.*, 2014). Therefore, AVP is a strong molecular candidate of prosocial learning by modulating on the underlying neurocomputational mechanisms when we learn to act prosocially under specific social contexts.

To examine learning adaptation of the self/other-oriented bias in reward-seeking and punishment-avoidance and modulated effect of intranasal AVP on specific prosocial learning, we designed a probabilistic reversal learning task in which participants learned to benefit self/others, or to avoid punishment for self/others separately. Specifically, the neurocomputational mechanisms underlying the effect of AVP on prosocial learning were examined by using computational modeling as well as recording event-related potentials (ERPs) and brain oscillations. We hypothesized that self/other-oriented bias would be specific in reward/punishment-related prosocial learning and AVP might be a crucial modulatory that supports prosocial learning.

## Materials and methods

### Participants

One hundred and four healthy participants were recruited in the current study (age: 18–26; 54 males; two left-handed). For the effect size ( $f=0.30$ ), type I error rate of 0.05, and statistical power of 0.8, G-Power 3.1 yielded a required minimum sample size of 58 participants for two [drug administration: placebo (PBO) *v.* AVP] between-subject factor and interactions with other factors in a repeated-measure design (Faul, Erdfelder, Lang, & Buchner, 2007). Participants were recruited via an online recruiting system and received monetary compensations. All potential participants completed a medical history questionnaire. Participants were not recruited if they reported any clinical disorder, drug/medication/alcohol abuse, or had recently participated in any other drug studies, or majored in economics/psychology. Participants were kept away from caffeine and alcohol on the day of experiment and from drink (except for water) and food for 2 h before the drug administration. The study was carried out according to the 1964 Helsinki Declaration and its later amendments and was approved by the local Ethics Committee. Written informed consent was obtained from each participant before the experiment. For electroencephalography (EEG) analyses, data from six participants were excluded because of incomplete EEG data, and data from four participants were discarded due to none trial available for any condition after denoising or due to left-handedness for the stimulus-preceding negativity (SPN) analysis.

### Administration of AVP and PBO

Drug administration of the current study was randomized, double-blind, and PBO-controlled. Participants were randomly assigned to the PBO or the AVP group. The PBO group self-administered 20 IU of PBO ( $n=50$ ; 24 females) intranasally and the AVP group 20 IU of vasopressin ( $n=54$ ; 27 females). The effective time of 20 IU AVP on social processes is about 80 min (Born *et al.*, 2002; Thompson, George, Walton, Orr, & Benson, 2006). In the experiment, an experimenter inspected the drug administration; however, both the experimenters and the participants were blind to the drug administration. Participants were asked to place the nasal applicator in one nostril and to press the lever until they felt a mist of spray in the nostril, then to breathe in deeply through the nose. Subsequently, participants were instructed to repeat this process in the other nostril. Each application involved both nostrils. In each application, the drug was applied three times in total with a 30 s delay. Participants proceeded to the main experiment approximately 20 min after drug treatment (Thompson *et al.*, 2006).

### Task procedure

The experiment consisted of two probabilistic reversal learning tasks (Fig. 1a), reward learning task (RLT) and punishment learning task (PLT). Each session included two runs, and participants made choices either for self or the other participant (informed that was the next participant) in each run (two runs are pseudo-random in each session). Therefore, there were four conditions in total, including making decisions for self in RLT session (SR), making decisions for others in RLT session (OR), making decisions for self in PLT session (SP), and making decisions for others in PLT session (OP). Participants were instructed to complete two

learning tasks. At the beginning of each trial, participants were instructed to make decisions for self or others. In each trial, after a fixation of 750–1250 ms, two visual stimuli/options were simultaneously presented to participants and asked to choose one option with the corresponding mouse click. In RLT session, one option was designated as the optimal option that associated with a high probability (70%) to obtain monetary reward (winning 5 cents) and a low probability (30%) to get a null reward (0 cent). The other option was linked to a low probability (40%) to obtain a reward and high probability (60%) to get a null reward. In contrast to RLT session, participants were informed that they were given 400 cents as initial funding in PLT. One option was associated with high probability (70%) to not to be punished (0 cent) and a low probability (30%) to be punished (losing 5 cents). The other option was associated with low probability (40%) to not to be punished and high probability (60%) to be punished. Once participants had chosen the optimal option on four consecutive occasions, the contingencies would reverse with a probability of 25% on each successive trial. Once the reinforcement contingencies reversed, the option with high rewards or low punishments (winning 5 cents in RLT frame or losing 0 cent in PLT frame) became frequently punished (winning 0 cent in RLT frame or losing 5 cents in PLT frame) and vice versa. Participants then needed to choose the other option – the one with high rewards or low punishments after reversed. To avoid participants using explicit strategies, such as counting the number of trials to reversal, they were not informed the details of how reversals were triggered by the computer but just be informed that reversals occurred randomly throughout the experiment. Participants were asked to obtain rewards or avoid punishments as more as possible, which were related to their payments.

### EEG data collection and preprocessing

EEG data were recorded continuously from 64 scalp sites using electrodes mounted on an elastic cap (Compumedics, Texas, USA), with an online reference to the left mastoid. All inter-electrode impedances were maintained below 5 k $\Omega$ . The EEG and electrooculography were filtered using a 0.05–100 Hz band-pass and continuously sampled at 500 Hz in each channel for off-line analysis. EEGs were re-referenced to the algebraic average of left mastoid and right mastoid. Eye blinks and muscle artifacts were cleaned using independent components analysis from the EEGLAB toolbox (Delorme & Makeig, 2004). Trials contaminated with artifacts exceeding  $\pm 100$   $\mu$ V were excluded from averaging.

### Data analysis

#### Behavioral measure

To quantify the performance of participants in the tasks, the accuracy was analyzed using repeated measures analysis of variance (ANOVA) with Context (RLT *v.* PLT) and Target (Self *v.* Other) as within-subject factors, and with Drug (PBO *v.* AVP) as a between-subject factor.

#### Computational model

To flexibly estimate participants' choices in response to changes in reward and punishment contingencies, we use positive–negative model (P-N), a different extension of the Rescorla–Wagner model (Rescorla & Wagner, 1972), to capture dissociable learning

effects from positive and negative outcomes separately:

$$V_{c,t} = \begin{cases} V_{c,t-1} + \eta^{pos}(O_{t-1} - V_{c,t-1}), & \text{if } O_{t-1} > 0 \\ V_{c,t-1} + \eta^{neg}(O_{t-1} - V_{c,t-1}), & \text{if } O_{t-1} < 0 \end{cases} \quad (1)$$

where  $\eta^{pos}$  is the reward learning rate (0 in negative feedback trials), and  $\eta^{neg}$  is the learning rate for negative feedback (0 in positive feedback trials);  $O$  is the received outcome; the value  $V$  at each trial  $t$  for the chosen option  $c$  ( $V_{c,t}$ ) is updated with the actual prediction error ( $O_{t-1} - V_{c,t-1}$ ).

We fitted models using the hBayesDM package (Ahn, Haines, & Zhang, 2017). Parameter estimation was performed with hierarchical Bayesian analysis using Stan language in R (Carpenter et al., 2017; Team, 2016). Markov chain Monte Carlo sampling was used to perform posterior inference and we compared and selected the optimal model by using LOOIC. We compared three computational models: Fictitious update model, a model which assumes that participants simultaneously update the value of the chosen and unchosen options; Experience-weighted attraction model, a model which captures the attribution of significance to past experience over and above new information as an individual progress through the task; and Positive–Negative model, which hypothesizes that individuals may update the estimation of the values by learning from positive and negative outcomes separately. See online Supplementary information for model comparisons. To identify the optimal learning parameters for each model, we simulated choice data for each learning rate with random noise. Then we inputted the simulated data to each model for fitting, exploring the parameter recovery and identifying the optimal learning rate (Crawley et al., 2020; Wilson & Collins, 2019). Next, we used the estimated model parameters from the winning model to simulate choices. For the following analyses, we excluded the data of two participants whose accuracy was lower than 45%, given that the data were outliers which the model could not fit precisely (Frank, Seeberger, & O'Reilly, 2004).

### EEG data analysis

To examine the neural mechanism of prosocial learning and AVP modulation, we were interested in the motivation and prediction-related SPN (Brunia & Damen, 1988; Hackley, Valle-Inclán, Masaki, & Hebert, 2014; Masaki, Yamazaki, & Hackley, 2010; Moris, Luque, & Rodríguez-Fornells, 2013), feedback-related negativity (FRN) that associated with expectation and learning processing (Gehring, Goss, Coles, Meyer, & Donchin, 1993; Holroyd & Coles, 2002; Miltner, Braun, & Coles, 1997; Yeung, Holroyd, & Cohen, 2005), and outcome evaluation-related P300 (Nieuwenhuis, Aston-Jones, & Cohen, 2005; Osinsky, Mussel, & Hewig, 2012). Given that the frontal theta is associated with the update of dynamic prediction error and cognitive control, while delta reflects the prediction of future behavioral adjustments (Cavanagh & Frank, 2014; Cohen, Elger, & Ranganath, 2007; Hauser et al., 2014), we also expected to observe theta (Bernat, Nelson, Steele, Gehring, & Patrick, 2011; Hauser et al., 2014) and delta oscillations (Bernat et al., 2011; Cavanagh, 2015) at outcome evaluation stage (see online Supplementary Fig. S1).

For ERP analyses, we were interested in the FRN and P300 as well as slow waves, such as the SPN. The original EEG data were low-pass filtered at 20 Hz for the SPN analysis, but band-pass filtered with cutoffs of 0.1 and 30 Hz to remove low-frequency waves from the EEG for the FRN and P300 analyses (Brunia, van Boxtel, & Böcker, 2012; Zheng, Li, Wang, Wu, & Liu,



2015). The filtered EEG data were then segmented into epochs that were time-locked to the feedback onset. For the SPN, epochs were extracted from  $-2500$  to  $500$  ms, with the activity from  $-2500$  to  $-2300$  ms serving as the baseline (Hackley et al., 2014; Masaki, Takeuchi, Gehring, Takasawa, & Yamazaki, 2006). We selected this baseline at the start of anticipation because SPN was a slow and negative wave that progressively developed prior to the feedback presentation, assuming that the baseline interval did not contain the signal of SPN. For the FRN and P300, epochs were extracted from  $-500$  to  $1000$  ms around each feedback onset for further analyses. Afterward, epochs were extracted from  $-200$  to  $1000$  ms, with the activity from  $-200$  to  $0$  ms serving as the baseline for the analyses of FRN and P300 (Zheng et al., 2015), assuming that neural activity in this period is unaffected by the feedback presentation. For illustration, SPN waveforms were filtered with a low-pass cutoff at  $7$  Hz ( $24$  dB/octave).

Based on the grand-average waveforms and topographic maps, amplitude of SPN from  $-200$  to  $0$  ms (i.e. the  $200$  ms window immediately prior to the feedback onset) was extracted as the mean voltage at bilateral electrode sites (F5/6, and FC5/6). Two participants were excluded in the SPN analysis because of no trial available in one condition after denoising and two were excluded due to left-handedness. The data were analyzed by using a repeated measure ANOVA, with Context (RLT *v.* PLT), Target (Self *v.* Other), Hemisphere (Left *v.* Right), and Site (F5/6 *v.* FC5/6) as within-subject factors, and with Drug (PBO *v.* AVP) as the between-subject factor. Based on previous studies, we used a peak-to-peak method to measure the FRN (Holroyd, Nieuwenhuis, Yeung, & Cohen, 2003; Osinsky et al., 2012; Osinsky, Walter, & Hewig, 2014). See online Supplementary information for details. To isolate the FRN from the confusion of positive feedback (Holroyd, Krigolson, Baker, Lee, & Gibson, 2009; Walsh & Anderson, 2012; Zheng et al., 2015), we created peak-to-peak FRN difference waves (negative feedback minus positive feedback under each condition) separately for the positive and negative feedback contexts in RLT as well as PLT (Pfabigan, Alexopoulos, Bauer, & Sailer, 2011). The peak-to-peak FRN difference waves were computed under SR, OR, SP and OP conditions. We measured the peak-to-peak FRN difference waves at FCz, a location that used to analyze the FRN in reinforcement learning and correlated with the update of dynamic prediction error (Hauser et al., 2014), and where the difference waves were maximal across the entire sample. Similarly, the P300 amplitude was calculated as the mean voltage difference wave at CPz (Cavanagh, 2015) given a posterior distribution of the P300 component in the period  $320$ – $420$  ms after feedback onset. The  $2$  (Context)  $\times$   $2$  (Target)  $\times$   $2$  (Drug) ANOVA was used to examine the differences in FRN and P300, respectively.

Next, we focused on oscillations at delta band ( $<4$  Hz) and theta band ( $4$ – $7$  Hz) in the outcome evaluation (see online Supplementary information for details). Time–frequency distributions of the EEG time course were obtained using a windowed Fourier transform with a fixed  $200$  ms Hanning window for theta signal acquisition and with a fixed  $500$  ms window for delta signal acquisition. For each epoch, thus, there was a complex time–frequency spectral estimate at each point of the time–frequency plane, extending from  $-500$  to  $1000$  ms (in  $2$  ms intervals) in the time domain, and from  $1$  to  $30$  Hz (in  $1$  Hz intervals) in the frequency domain. The resulting spectrogram represents the signal power as a joint function of time and frequency at each time–frequency point. As the center of fixed  $200$  ms

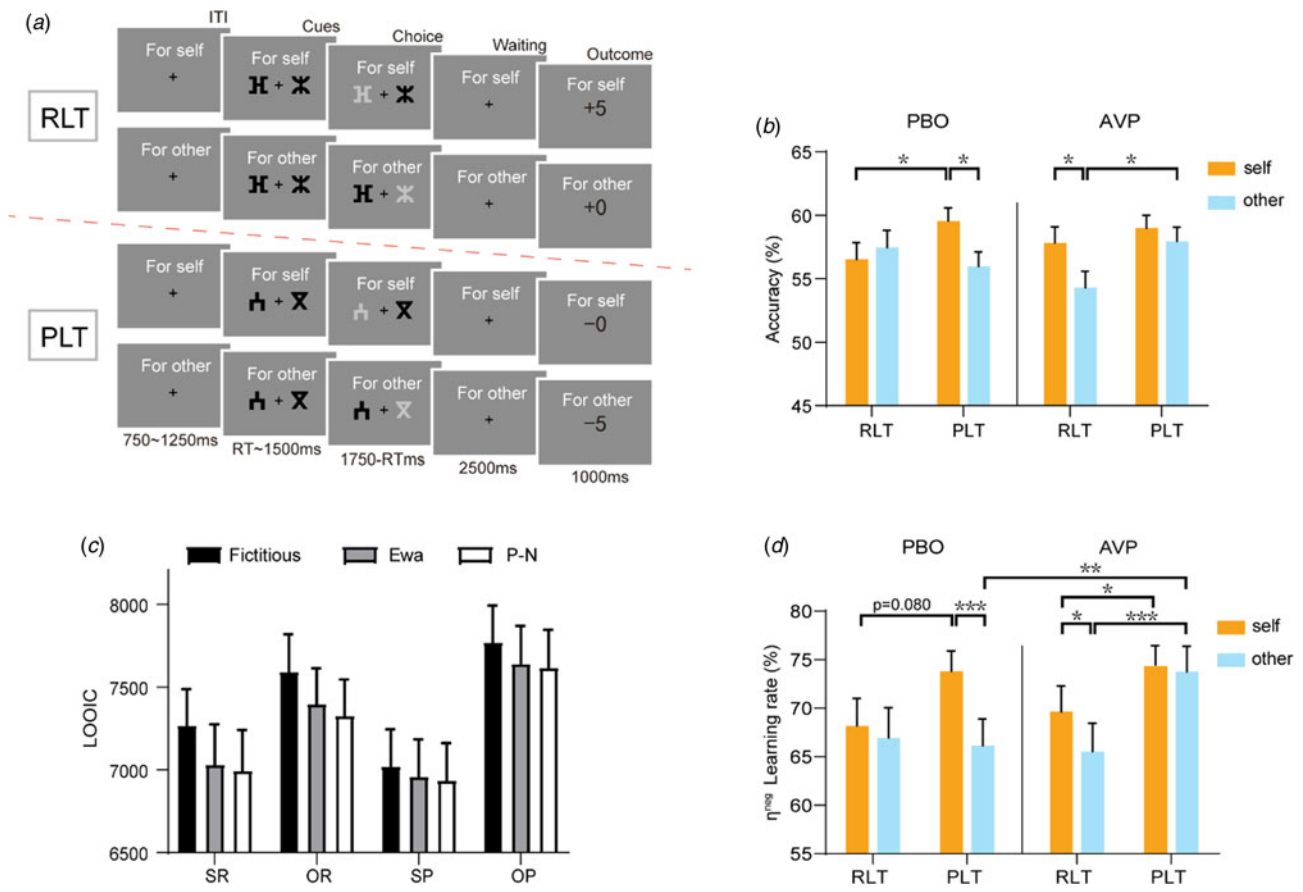
Hanning window moves among the time range from  $-100$  to  $0$  ms, the complex time–frequency spectral estimate of the time–frequency plane would be contaminated by the signals after feedback onset (Hu & Zhang, 2019). Therefore, the spectrogram was baseline-corrected (with the reference interval from  $-300$  to  $-200$  ms relative to feedback onset) at each frequency using the subtraction approach (Cavanagh, 2015). The mean of theta activity ( $4$ – $7$  Hz) was extracted in the  $100$ – $300$  ms interval following feedback onset at FCz, because topographic distributions of power exhibited a fronto-central peak that was maximal around FCz. The mean of delta ( $<4$  Hz) activity was extracted in the  $320$ – $420$  ms interval at Cz. Power differences between negative and positive feedback as well as differences of frequency activities were then compared by using the  $2$  (Frequency band)  $\times$   $2$  (Context)  $\times$   $2$  (Target)  $\times$   $2$  (Drug) repeated measures ANOVA. To explore the relationship between behavioral adjustments and brain oscillations, we correlated the average delta activity with reaction time (RT) (Cavanagh, 2015) and estimated the moderating effect of AVP on the relation between delta activity and RT (online Supplementary Fig. S1).

Moreover, we conducted two moderation models at the anticipation stage and the outcome evaluation stage, estimating the moderating role of AVP in the relation between neuroelectrophysiological signals and psychological processes. For all statistical tests, Greenhouse–Geisser epsilon correction was applied for non-sphericity when appropriate (Jennings & Wood, 1976). The partial eta-squared ( $\eta_p^2$ ) was reported as a measure of effect size. The Bonferroni procedure was used to corrected for multiple comparisons in the post hoc analyses.

## Results

### *Behavioral differences in learning to avoid punishment and AVP modulation on self-related reward-seeking and other-regarded punishment-avoidance*

The  $2 \times 2 \times 2$  ANOVA of accuracy showed a significant main effect of Context that the accuracy was higher in punishment learning than reward learning ( $F_{(1,102)} = 4.396$ ,  $p = 0.038$ ,  $\eta_p^2 = 0.041$ ), a significant main effect of Target that the accuracy was higher when participants learned for themselves than others ( $F_{(1,102)} = 5.046$ ,  $p = 0.027$ ,  $\eta_p^2 = 0.047$ ), and a three-way significant interaction of Context  $\times$  Target  $\times$  Drug ( $F_{(1,102)} = 5.231$ ,  $p = 0.024$ ,  $\eta_p^2 = 0.049$ ; Fig. 1b). However, the main effect of Drug ( $F_{(1,102)} = 0.011$ ,  $p = 0.916$ ,  $\eta_p^2 = 0.000$ ) and the interaction effect of Context  $\times$  Drug ( $F_{(1,102)} = 1.204$ ,  $p = 0.275$ ,  $\eta_p^2 = 0.012$ ) and effect of Target  $\times$  Drug ( $F_{(1,102)} = 0.362$ ,  $p = 0.549$ ,  $\eta_p^2 = 0.004$ ) were not significant. Simple effect analyses of the three-way significant interaction showed that the interaction between effect of Context and Target was significant in PBO group ( $F_{(1,102)} = 4.19$ ,  $p = 0.043$ ,  $\eta_p^2 = 0.099$ ) but not in AVP group ( $F_{(1,102)} = 1.37$ ,  $p = 0.245$ ,  $\eta_p^2 = 0.02$ ). In addition, the interaction effect between Context and Drug in Target of other condition was significant ( $F_{(1,102)} = 5.511$ ,  $p = 0.021$ ,  $\eta_p^2 = 0.051$ ). Specifically, in the PBO group, the accuracy of SP was significantly higher than SR ( $F_{(1,102)} = 3.946$ ,  $p = 0.050$ ,  $\eta_p^2 = 0.037$ ), and accuracy of SP was significantly higher than OP ( $F_{(1,102)} = 5.311$ ,  $p = 0.023$ ,  $\eta_p^2 = 0.049$ ). However, in the AVP group, the accuracy of OP was significantly higher compared to OR ( $F_{(1,102)} = 5.765$ ,  $p = 0.018$ ,  $\eta_p^2 = 0.053$ ), and accuracy of SR was significantly higher relative to OR ( $F_{(1,102)} = 4.983$ ,  $p = 0.028$ ,  $\eta_p^2 = 0.047$ ). These results suggested that AVP modulates the adaption of individuals' self/other-



**Fig. 1.** Experimental design and behavioral results. (a) Experimental procedure. After a fixation of 750–1250 ms, participants were given a response window with a maximum time of 1500 ms following the cue onset and they needed to make a choice. A 2500 ms waiting window (a fixation in the center) occurred after the response window. Subsequently, the outcome was presented for 1000 ms. Each participant was asked to learn for self and for others separately in RLT and PLT sessions. RLT represents Reward learning task and PLT Punishment learning task. (b) Behavioral accuracy across the four conditions (SR, SP, OR, OP) of the two groups (PBO v. AVP). Error bar indicates the standard error of the mean. \* $p < 0.05$ , \*\* $p < 0.01$ , \*\*\* $p < 0.001$ . (c) Comparisons of Fictitious, EWA, and P-N models by using LOOIC under each condition. Fictitious represents the fictitious update model; EWA represents the experience-weighted attraction model; P-N represents the positive–negative model. (d) Negative learning rate  $\eta^{neg}$  of P-N model across the four conditions (SR, SP, OR, OP) by two groups (PBO vs. AVP). SR, making decisions for self in RLT session; SP, making decisions for self in PLT session; OR, making decisions for others in RLT session; OP making decisions for others in PLT session.

oriented bias depending on the specific frames. Specifically, AVP promotes individuals’ self-bias in reward learning as compared to learn for others, and enhances prosocial performance in punishment learning as compared to learn for self.

**Computational evidence for self-bias on punishment learning and dissociable modulations of AVP in prosocial learning**

Bayesian model comparison showed that the positive–negative (P-N) model was superior to the other two models under all four conditions (Fig. 1c). Subsequently, two estimated P-N model parameters learning rate  $\eta^{pos}$  and  $\eta^{neg}$  were analyzed using the 2 (Learning rate)  $\times$  2 (Context)  $\times$  2 (Target)  $\times$  2 (Drug) ANOVA, where the four-way interaction was significant ( $F_{(1,100)} = 78.122, p = 0.000, \eta_p^2 = 0.439$ ). Interestingly, we found a three-way interaction of learning rate for negative feedback  $\eta^{neg}$  was significant ( $F_{(1,100)} = 6.999, p = 0.009, \eta_p^2 = 0.065$ ; Fig. 1d), which was consistent with the results of accuracy. We also found a significant main effect of Context with a better performance on punishment learning ( $F_{(1,100)} = 7.556, p = 0.007, \eta_p^2 = 0.070$ ) and significant main effect of Target ( $F_{(1,100)} = 13.532, p = 0.000, \eta_p^2 = 0.119$ ) for  $\eta^{neg}$  parameter, while the

main effect of Drug ( $F_{(1,100)} = 0.243, p = 0.623, \eta_p^2 = 0.002$ ) was not significant. The two-way interaction effect of Context  $\times$  Drug was significant ( $F_{(1,100)} = 5.882, p = 0.017, \eta_p^2 = 0.056$ ), while the interaction effect of Target  $\times$  Drug was not significant ( $F_{(1,100)} = 1.727, p = 0.192, \eta_p^2 = 0.017$ ). Simple effect analyses illustrated that the AVP group performed better than PBO in punishment learning ( $F_{(1,100)} = 13.654, p = 0.000, \eta_p^2 = 0.120$ ), rather than reward learning ( $F_{(1,100)} = 13.654, p = 0.821, \eta_p^2 = 0.120$ ).

Simple effect analyses of three-way interaction of  $\eta^{neg}$  parameter showed that in the PBO group, the negative feedback learning rate  $\eta^{neg}$  in SP was marginal significantly higher compared to SR ( $F_{(1,100)} = 3.134, p = 0.080, \eta_p^2 = 0.030$ ), and  $\eta^{neg}$  in SP was significantly higher than OP ( $F_{(1,100)} = 21.714, p = 0.000, \eta_p^2 = 0.178$ ). However, in the AVP group,  $\eta^{neg}$  in OP was significantly higher relative to OR ( $F_{(1,100)} = 12.480, p = 0.001, \eta_p^2 = 0.111$ ) and  $\eta^{neg}$  in SP was significantly higher compared to SR ( $F_{(1,100)} = 4.194, p = 0.043, \eta_p^2 = 0.040$ ). Moreover,  $\eta^{neg}$  in SR was significantly higher than OR ( $F_{(1,100)} = 3.960, p = 0.049, \eta_p^2 = 0.038$ ). Additionally,  $\eta^{neg}$  in OP was higher in AVP than under PBO condition ( $F_{(1,100)} = 7.872, p = 0.006, \eta_p^2 = 0.073$ ). These results supported the behavioral results and suggested that AVP may modulate the abilities of individuals to capture the information

of negative feedback trials, particularly for prosociality on punishment learning and for self-orientation on reward learning.

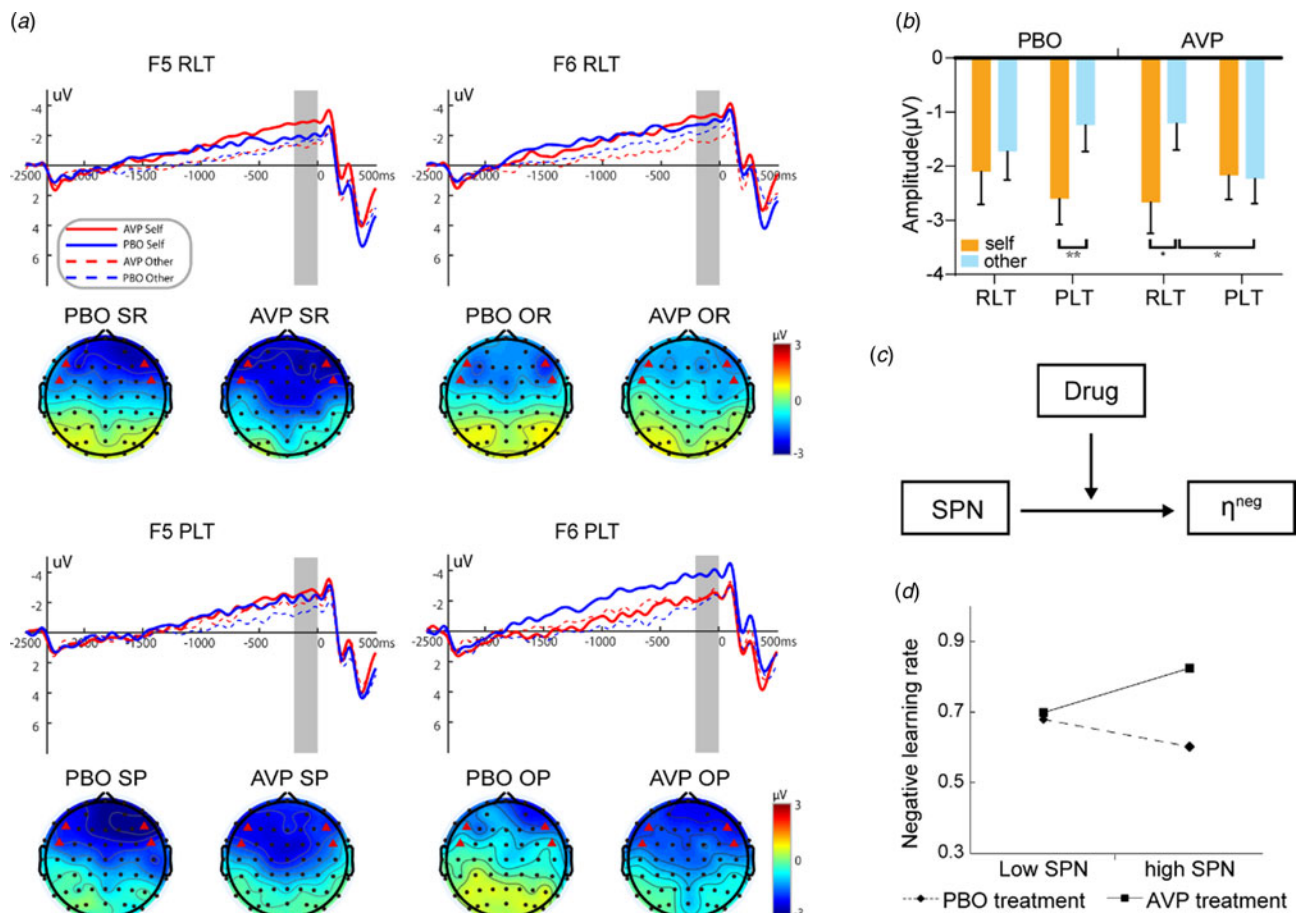
#### Identifying self-bias on punishment learning and dissociable modulations of AVP in prosocial learning using ERPs

**SPN at the stage of anticipation.** The SPN develops gradually as a relative negativity after the choice and reaches its maximum immediately prior to the feedback onset (Fig. 2a). The topography of the SPN appearing as a plateau-shaped tends to be larger in the frontal areas. The 2 (Context)  $\times$  2 (Target)  $\times$  2 (Drug) ANOVA of the SPN data revealed a significant three-way interaction effect ( $F_{(1,93)} = 5.651, p = 0.019, \eta_p^2 = 0.057$ ). Specifically, in the PBO group, the SPN amplitude in SP was significantly higher than OP ( $F_{(1,93)} = 7.31, p = 0.008, \eta_p^2 = 0.073$ ; Fig. 2b). While in AVP, the SPN amplitude in OP was significantly larger relative to OR ( $F_{(1,93)} = 4.154, p = 0.044, \eta_p^2 = 0.043$ ), and SPN in SR was significantly higher than in OR ( $F_{(1,93)} = 6.092, p = 0.015, \eta_p^2 = 0.061$ ). It should be mentioned that the PBO group showed a symmetrical distribution in the frontal areas ( $F_{(1,93)} = 4.315, p = 0.041, \eta_p^2 = 0.044$ ), which was in line with previous findings (Brunia, Hackley, van Boxtel, Kotani, & Ohgami, 2011). However, the main effect of Drug was not significant ( $F_{(1,93)} = 0.059,$

$p = 0.809, \eta_p^2 = 0.001$ ). At the anticipation stage, individuals prepare the brain for the upcoming feedback, reflecting SPN underlying modulations of AVP on anticipation in self-related reward-seeking and other-regarded punishment-avoidance behaviors.

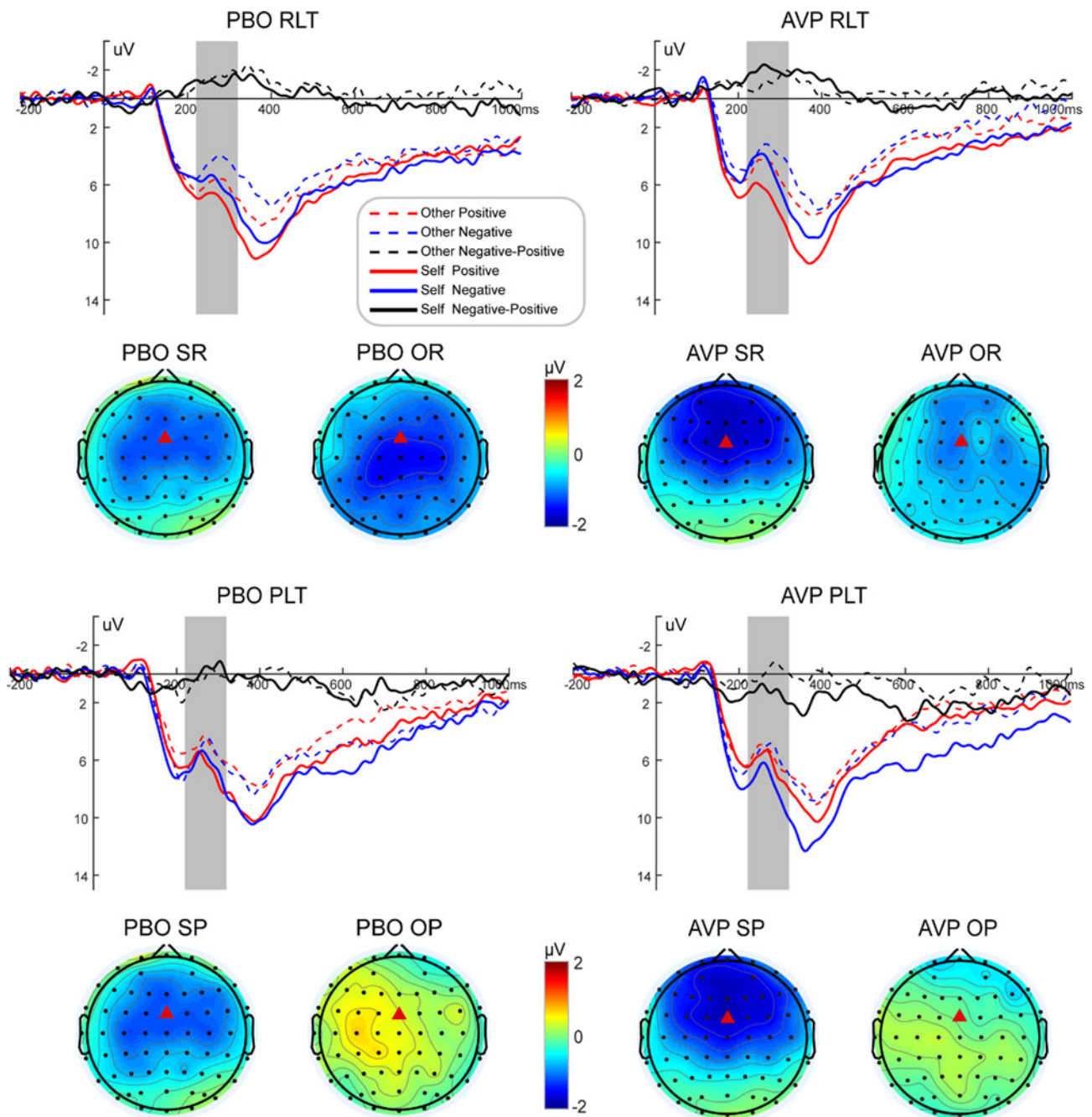
We established a moderation model (Fig. 2c) to estimate whether drug treatment would moderate the association between SPN amplitudes and  $\eta^{neg}$ . The results revealed a significant moderation under OP condition. Our model showed that there was a significant main effect of Drug on  $\eta^{neg}$  ( $b = 0.094, p = 0.009$ ) rather than SPN amplitudes ( $b = 0.002, p = 0.741$ ), and more importantly, the effect of SPN on  $\eta^{neg}$  was significantly moderated by Drug ( $b = 0.032, p = 0.005$ ). Simple slope tests revealed that higher level SPN amplitudes were associated with higher levels of negative learning rate in the AVP group ( $b_{simple} = 0.017, p = 0.015$ ; Fig. 2d), while this correlation was not significant in PBO ( $b_{simple} = -0.015, p = 0.093$ ). These results showed a moderating role of AVP in relation between anticipation and prosocial punishment learning.

**Dissociable neural processing between FRN and P300 at stage of outcome evaluation.** The difference waveform of FRN at FCz as a function of feedback type (positive v. negative) exhibited a negative deflection over the fronto-central regions during the feedback evaluation (Fig. 3), while the P300 at CPz showed a positive



**Fig. 2.** SPN. (a) SPN at electrodes F5 and F6 for four conditions. The shaded areas depict the SPN time window ( $-200$  to  $0$  ms), and the topographic maps are shown below. Red triangles display the electrodes selected for analysis. (b) The amplitude of SPN under SR, SP, OR, OP conditions by two groups (PBO v. AVP). SR, making decisions for self in RLT session; SP, making decisions for self in PLT session; OR, making decisions for others in RLT session; OP making decisions for others in PLT session. RLT represents Reward learning task and PLT Punishment learning task. Error bar indicates the standard error of the mean.  $*p < 0.05, **p < 0.01$ . (c) The proposed moderation model. SPN, SPN amplitude under condition of learning to avoid punishment for others; Drug, Drug treatment of PBO v. AVP;  $\eta^{neg}$ , negative learning rate. (d)  $\eta^{neg}$  among participants as a function of Drug and SPN amplitudes. Functions are graphed for two levels of the drug administration (PBO v. AVP) under condition of learning to avoid punishment for others.

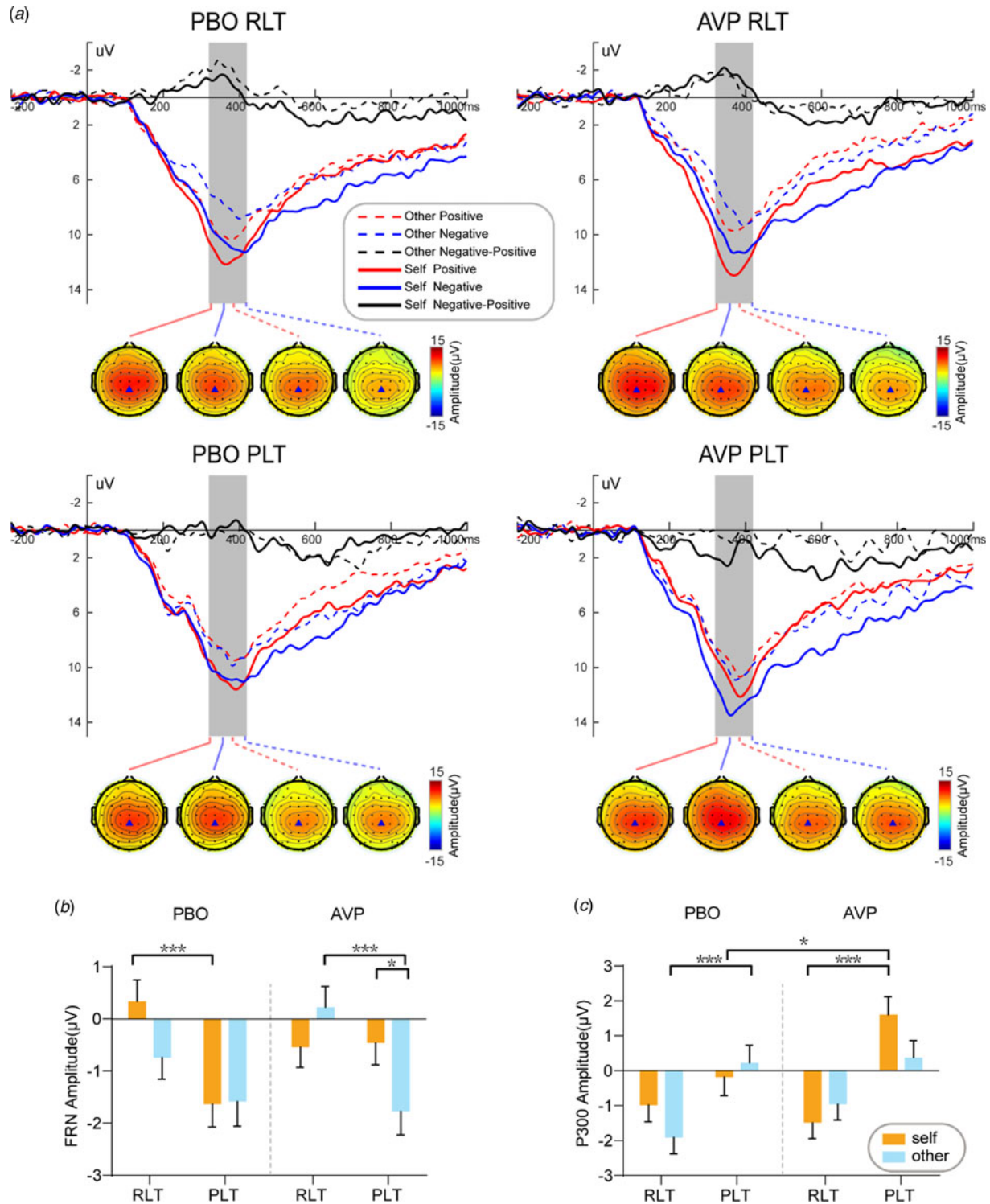




**Fig. 3.** FRN. ERPs of FRN for PBO and AVP groups at the stage of outcome evaluation under SR, OR, SP, and OP conditions at FCz. The differences between negative and positive feedback waveforms are also shown, where shaded areas indicate the FRN time window. Scalp map (220–320 ms) depicts the topography for the difference waveforms, where red triangles show the FCz electrode selected for analysis. SR, making decisions for self in RLT session; SP, making decisions for self in PLT session; OR, making decisions for others in RLT session; OP making decisions for others in PLT session. RLT represents Reward learning task and PLT Punishment learning task.

potential over the centroparietal regions (Fig. 4a). The  $2 \times 2 \times 2$  ANOVA of the peak-to-peak FRN difference wave reflected a three-way interaction ( $F_{(1,96)} = 7.811, p = 0.006, \eta_p^2 = 0.075$ ; Fig. 4b). Interestingly, in PBO, the negative amplitude of peak-to-peak FRN difference wave under SP condition was significantly larger than SR ( $F_{(1,96)} = 11.858, p = 0.001, \eta_p^2 = 0.11$ ), whereas in the AVP group, the difference amplitude in OP was significantly more negative compared to OR ( $F_{(1,96)} = 11.863, p = 0.001, \eta_p^2 = 0.11$ ). In addition, the difference amplitude of the AVP group under

condition OP was significantly larger than SP ( $F_{(1,93)} = 5.582, p = 0.020, \eta_p^2 = 0.055$ ). The  $2 \times 2 \times 2$  ANOVA of the P300 difference wave revealed a three-way interaction effect among Context, Target, and Drug ( $F_{(1,96)} = 5.332, p = 0.023, \eta_p^2 = 0.053$ ; Fig. 4c). The simple effect analysis illustrated that in the PBO group, the amplitude of P300 difference wave in response to OR was significantly larger than OP ( $F_{(1,96)} = 8.734, p = 0.004, \eta_p^2 = 0.083$ ), while difference wave in SR was significantly larger than SP in the AVP group ( $F_{(1,96)} = 25.261, p = 0.000, \eta_p^2 = 0.208$ ).



**Fig. 4.** P300. (a) ERPs of P300 for PBO and AVP at the later stage of feedback evaluation in SR, OR, SP, and OP at CPz when facing positive feedback and negative feedback, where shaded areas depict the P300 time window. Scalp maps (320–420 ms) show the topography for the P300, where blue triangles indicate the CPz electrode selected for analysis. (b) Peak-to-peak FRN difference waveform (negative minus positive feedback). Error bar indicates the standard error of the mean. \* $p < 0.05$ , \*\* $p < 0.01$ , \*\*\* $p < 0.001$ . (c) Peak-to-peak P300 difference waveform (negative minus positive feedback). SR, making decisions for self in RLT session; SP, making decisions for self in PLT session; OR, making decisions for others in RLT session; OP, making decisions for others in PLT session. RLT represents Reward learning task and PLT Punishment learning task.

Amplitude of the P300 difference wave in OP was higher in AVP relative to in PBO ( $F_{(1,96)} = 5.932$ ,  $p = 0.017$ ,  $\eta_p^2 = 0.058$ ). There were no significant main effect of Drug in neither FRN ( $F_{(1,96)} = 0.633$ ,  $p = 0.428$ ,  $\eta_p^2 = 0.007$ ) nor P300 ( $F_{(1,96)} = 2.467$ ,

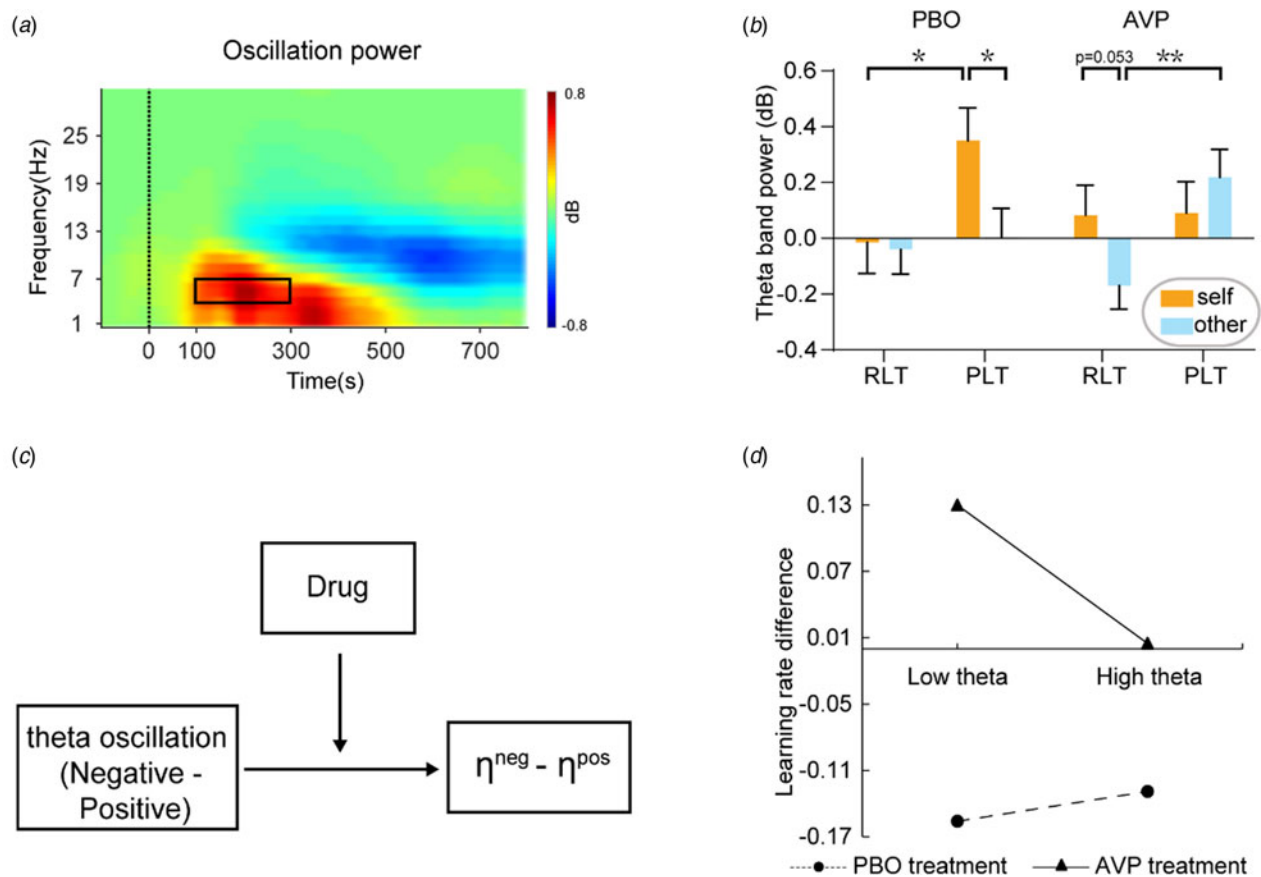
$p = 0.120$ ,  $\eta_p^2 = 0.025$ ). Under the interactive learning tasks, these results suggested that AVP enhanced difference wave of FRN which responded to other-related feedback in punishment situation and increased P300 that acted to self-related feedback in reward situation.



**Theta oscillation at stage of outcome evaluation.** To examine brain oscillations underlying the modulation of AVP on prosocial learning, 2 (Frequency band) × 2 (Context) × 2 (Target) × 2 (Drug) repeated measures ANOVA was applied to analyze differences of frequency activities under negative–positive feedback condition at FCz (Fig. 5). The results showed that a main effect of Context ( $F_{(1,96)} = 5.974, p = 0.004, \eta_p^2 = 0.084$ ) reflecting a larger response evoked on punishment learning. There was a significant four-way interaction effect ( $F_{(1,96)} = 5.974, p = 0.016, \eta_p^2 = 0.059$ ). *Post hoc* analyses reflected that the significant effect was focused on the three-way interaction effect of Context by Target by Drug on theta band ( $F_{(1,96)} = 4.559, p = 0.035, \eta_p^2 = 0.045$ ; Fig. 5b), rather than delta band ( $F_{(1,96)} = 3.081, p = 0.082, \eta_p^2 = 0.031$ ). Simple effect analyses showed that in PBO, the difference power from theta band in SP was significantly larger than that in SR ( $F_{(1,96)} = 4.703, p = 0.033, \eta_p^2 = 0.047$ ), and the power in SP was significantly higher than that in OP ( $F_{(1,96)} = 4.050, p = 0.047, \eta_p^2 = 0.040$ ). However, in the AVP group, theta band power in OP was significantly larger than OR ( $F_{(1,96)} = 8.054, p = 0.006, \eta_p^2 = 0.077$ ), and the power in SR was marginal significantly higher than OR ( $F_{(1,102)} = 3.854,$

$p = 0.053, \eta_p^2 = 0.039$ ). There was no significant main effect of Drug in FRN ( $F_{(1,96)} = 0.003, p = 0.960, \eta_p^2 = 0.000$ ). These results revealed that theta oscillation underlying the modulation of AVP on self-oriented reward-seeking and prosocial punishment-avoidant behaviors.

Moreover, our data revealed a significant moderating effect at the outcome evaluation stage (Fig. 5c), estimating the moderation of drug treatment on the association between theta oscillation difference (negative – positive) and difference between negative and positive learning rate parameters ( $\eta^{neg} - \eta^{pos}$ ) under OP condition. Our model showed that there was a significant main effect of Drug on  $\eta^{neg} - \eta^{pos}$  ( $b = 0.210, p = 0.000$ ) rather than theta oscillation difference ( $b = -0.035, p = 0.161$ ), and more importantly, the effect of theta oscillation difference on  $\eta^{neg} - \eta^{pos}$  was significantly moderated by Drug ( $b = -0.107, p = 0.036$ ). Simple slope tests revealed that lower level theta oscillation was associated with higher levels of  $\eta^{neg} - \eta^{pos}$  in the AVP group ( $b_{simple} = -0.088, p = 0.024$ ; Fig. 5d), while this correlation was not significant in PBO ( $b_{simple} = 0.019, p = 0.562$ ). These results revealed the moderating mechanism of AVP in prosocial punishment learning.



**Fig. 5.** Theta-band activity. (a) The mean oscillation activities of feedback evaluation across all conditions at FCz. Time–frequency plots the mean oscillation activities over time (x-axis; 0 is the onset of feedback stimulus) and frequency (y-axis). The square selects the data in a 100–300 ms window from the theta band (4–7 Hz). (b) Difference power between negative and positive feedback at theta band (4–7 Hz) under SR, OR, SP, and OP conditions by two groups (PBO v. AVP). Error bar indicates the standard error of the mean. \* $p < 0.05$ , \*\* $p < 0.01$ . SR, making decisions for self in RLT session; SP, making decisions for self in PLT session; OR, making decisions for others in RLT session; OP making decisions for others in PLT session. RLT represents Reward learning task and PLT Punishment learning task. (c) The proposed moderation model. Theta oscillation, theta oscillation difference (negative feedback minus positive feedback) under condition of learning to avoid punishment for others; Drug, drug treatment of PBO v. AVP;  $\eta^{neg} - \eta^{pos}$ , difference between negative and positive learning rates under prosocial punishment learning condition. (d)  $\eta^{neg} - \eta^{pos}$  among participants as a function of Drug and theta oscillation difference. Functions are graphed for two levels of the drug administration (PBO v. AVP) under condition of learning to avoid punishment for others.

## Discussion

In the present study, we used a combination of behavioral manipulation, computational modeling, and EEG to examine the diverse adaptations when people learn to benefit or avoid harms for themselves (self-oriented learning) and for others (prosocial learning) and the modulatory role of AVP in this adaption, from external behavioral performance, to internal psychological processes, and underlying neural dynamics. Our behavioral findings showed that the self-bias was specified for avoiding punishment and AVP increased learning performance for self in the reward-seeking and for others in the punishment-avoidance. Using computational modeling and electrophysiological measurements, we found the self-bias and the modulation of AVP was specific for negative feedback learning, underpinned by increased brain responses in anticipation (i.e. SPN) and in outcome evaluation (i.e. FRN and P300, as well as frontal theta oscillations). At the stage of outcome evaluation, AVP system improved prosocial learning by adjusting punishment-related early FRN neural process and acting on reward-related late P300 neural processes to enhance prosocial learning, while two diverse time-series responses were processing at theta band. At the anticipation stage, increased SPN in AVP than PBO suggests that AVP system directly modulated self/other-oriented bias to expedite learning for self-oriented reward-seeking and other-regarded punishment-avoidance behaviors. Together, our study shows the neurocomputational mechanisms of how we adapt to obtain reward or avoid punishment in self-oriented and prosocial learning, where AVP plays a context-dependent modulatory role.

In the PBO group, individuals behaved differently between for self and other in the learning task, suggesting a self-bias in punishment learning. Previous studies focusing on reward learning have also shown a better performance for self than for others in prosocial learning (Lockwood *et al.*, 2016; Martins *et al.*, 2022). Consistently, we found a self-bias of learning rate when participants learned from positive feedback in reward learning (online Supplementary Fig. S2), which was not shown in accuracy. Using prosocial learning paradigm characterized by reward and punishment, we found the self-oriented learning effect when avoiding punishment. However, there are also studies showing influences of social dilemma on people's social preferences, especially other-regarding concerns and altruism (Liu *et al.*, 2020; van Dijk & Wilke, 2000). Contrary to other social dilemma paradigms which normally consider a tradeoff between economic benefits and the feelings of others, decisions were made in self-action reference across reward and punishment in our task. Consideration without those tradeoffs may be one explanation for the absence of altruism in the PBO group. Therefore, we could measure the interaction effect between self/other-oriented biases and reward/punishment biases in social learning by our learning framework.

Our results showed that intranasal AVP up-regulated altruism concerning others' losses and reward-seeking for self-oriented benefits. Consistent with recent studies which has shown the involvement of AVP in prosocial behavior (Nishina *et al.*, 2019; Wang *et al.*, 2016), and social cooperative behaviors (Feng *et al.*, 2015), our results showed that AVP enhances altruism, particularly in protecting others from monetary losing. On the other hand, vasopressin can promote individuals to maximize personal utilities in adaption to the environment (Brunnlieb *et al.*, 2016; Patel *et al.*, 2015). This hypothesis is also supported by our results that AVP improves individuals' performances toward prosocial

benefits in the reward-seeking. Thus, AVP induced both prosocial and prosocial behaviors, depends on reward/punishment contexts. Although the nonsignificant difference toward learning performances and neural responses between the PBO and AVP groups, we found a three-way interaction which suggested the effects of AVP was conditional.

Different learning rates in the negative rather than positive feedback learning, which was modulated by the AVP system, suggest that learning information from negative feedback may be a crucial aspect when participants are making decisions on social learning. People are more sensitive to negative information than positive information, which has been shown as the negativity bias in attention (Rozin & Royzman, 2001) and as loss aversion in decision making (Tversky & Kahneman, 1991). Previous reinforcement-learning models suggest that negative outcomes make greater contributions to the overall feedback evaluation (Cavanagh, Frank, Klein, & Allen, 2010; Pearce & Hall, 1980). In social setting, negative events weigh more heavily than positive ones (Alves, Koch, & Unkelbach, 2017; Shin & Niv, 2021). It has also been shown that people take the consequence of their actions into account when it will have an impact on others, in particular learning to avoid harming others (Crockett *et al.*, 2014; Lockwood *et al.*, 2020). Therefore, negative information processing could be an important aspect in prosocial learning and more evolutionarily natural in the modulatory role of the AVP system.

The dissociated responses of FRN and P300 to specific learning framework suggest that distinct time-series of neural processes underlying self/other-oriented bias. FRN and P300 are the critical ERP components in outcome evaluation. Previous studies suggest that the FRN reflected a fast evaluation of outcome valence, with a larger differential effect between loss and win (Gehring *et al.*, 1993; Yeung *et al.*, 2005) and evaluation of consistency between expectation and actual outcomes (Holroyd & Coles, 2002). In contrast, P300 is related to reward processing and sensitive to a later, top-down controlled process of outcome evaluation (Cavanagh, 2015; Nieuwenhuis *et al.*, 2005; Pfabigan *et al.*, 2011). Under social context, FRN and P300 respond to different outcomes for oneself and others (Hu, Xu, & Mai, 2017; Qi, Wu, Raiha, & Liu, 2018). For instance, FRN is sensitive to self-benefit context while P300 responds to prosocial context in a gambling task (Qi *et al.*, 2018). Consistently, we observed a larger FRN when making decision for themselves in aversive situation, while a larger P300 response for others benefits in the PBO group. Interestingly, AVP up-regulated prosocial punishment-avoidance behaviors with a larger FRN response in aversive situation, while modulated the reward-seeking behaviors with a larger P300 response for self-related benefit. Together, these results suggested that the AVP system dissociated improves prosocial learning by adjusting punishment-related FRN neural process and acting on reward-related P300 neural processes.

Our results also shed light on brain oscillation mechanisms for modulations of AVP on prosocial learning at the outcome evaluation stage. Consistent with behavioral and computational model measures, the theta difference activity also supported the self-bias and dissociated modulations of AVP. Previous studies suggest that midfrontal theta band activity was predictive of cognitive control (Cavanagh & Frank, 2014; Cohen, 2011) and was indicative of altruistic behavioral responses (Rodrigues, Ulrich, & Hewig, 2015). Theta activity has also been shown to respond to social interactions (Rodrigues *et al.*, 2015; Tandler & Wagner, 2015). Therefore, theta activity, prosocial in reward-seeking and prosocial

in punishment-avoidance learning, indicates the underpinning control mechanisms of the interaction between self/other-oriented bias and feedback valence.

Lastly, SPN findings suggest the self-bias and AVP modulation at the anticipation stage of prosocial learning. In our study, SPN showed a self-bias when participants trying to avoid punishments. Larger amplitudes of SPN at the anticipation stage when individuals making decision for self-related reward-seeking and other-regarded punishment-avoidance, suggest that participants with AVP administration had a biased expectation to imminent outcome. The SPN, a slow and negative wave that progressively develops prior to the motivational stimuli (Brunia & Damen, 1988), has been considered to reflect outcome predictions and expectation of response reinforcement (Masaki et al., 2010). The right-hemisphere predominance in PBO was fairly compatible with previous findings, possibly reflecting contributions from the ventral attention system (Brunia et al., 2011; Zheng et al., 2015). SPN was interpreted as preparatory activity aimed at speeds up brain processes after the relevant stimulus, preparation of the brain for the upcoming event or action, and as an index of anticipatory attention (Brunia et al., 2011, 2012). Therefore, our SPN findings illuminate a self-bias in preparation for forthcoming aversive stimulus in social learning, while the AVP system directly modulates self/other-oriented bias at the anticipation stage to expedite self-oriented reward-seeking and other-oriented punishment-avoidance behaviors in prosocial learning.

Overall, our findings suggest that intranasal vasopressin modulates self/other-oriented bias by up-regulating self-related reward-seeking and other-regarded punishment-avoidance behaviors in prosocial learning. AVP modulates learning and processing of negative feedback at both stages of anticipation and outcome evaluation. These modulations of AVP systems are underpinned by punishment-related FRN for prosocial learning and reward-related P300 for proself learning, as well as theta band oscillations at the outcome evaluation stage and SPN at anticipation stage. Our work sheds light on the mechanisms of our prosocial behaviors and has important implications in the atypical social behaviors of psychiatry disorders.

**Supplementary material.** The supplementary material for this article can be found at <https://doi.org/10.1017/S0033291722002483>.

**Acknowledgements.** This study was supported by the National Natural Science Foundation of China (31871137, 31900757, 31920103009, and 32020103008), the Major Project of National Social Science Foundation (20&ZD153), Young Elite Scientists Sponsorship Program by China Association for Science and Technology (YESS20180158), Guangdong International Scientific Collaboration Project (2019A050510048), Natural Science Foundation of Guangdong Province (2020A1515011394 and 2021A1515010746), Shenzhen-Hong Kong Institute of Brain Science-Shenzhen Fundamental Research Institutions (2019SHIBS0003), and Shenzhen Science and Technology Research Funding Program (JCYJ20180507183500566, JCYJ20180306173253533 and JCYJ20190808121415365).

**Author contributions.** J. X., C. F., and P. X. designed research; L. Q. performed research; G. D. analyzed data; and G. D., H. A., C. F., and P. X. wrote the paper.

**Conflict of interest.** The authors declare that they have no competing financial interests.

## References

- Ahn, W. Y., Haines, N., & Zhang, L. (2017). Revealing neurocomputational mechanisms of reinforcement learning and decision-making with the hBayesDM package. *Computational Psychiatry*, 1, 24–57. [https://doi.org/10.1162/cpsy\\_a\\_00002](https://doi.org/10.1162/cpsy_a_00002).
- Albers, H. E. (2015). Species, sex and individual differences in the vasotocin/vasopressin system: Relationship to neurochemical signaling in the social behavior neural network. *Frontiers in Neuroendocrinology*, 36, 49–71. <https://doi.org/10.1016/j.yfrne.2014.07.001>.
- Albers, J. R. (1994). Learning as adaptation of the infant. *Acta Paediatrica*, Supplement, 397, 77–85. <https://doi.org/10.1111/j.1651-2227.1994.tb13269.x>.
- Alves, H., Koch, A., & Unkelbach, C. (2017). The ‘common good’ phenomenon: Why similarities are positive and differences are negative. *Journal of Experimental Psychology: General*, 146(4), 512–528. <https://doi.org/10.1037/XGE0000276>.
- Apps, M. A., Rushworth, M. F., & Chang, S. W. (2016). The anterior cingulate gyrus and social cognition: Tracking the motivation of others. *Neuron*, 90(4), 692–707. <https://doi.org/10.1016/j.neuron.2016.04.018>.
- Avinun, R., Israel, S., Shalev, I., Gritsenko, I., Bornstein, G., Ebstein, R. P., & Knafo, A. (2011). AVPR1A variant associated with preschoolers’ lower altruistic behavior. *PLoS ONE*, 6(9), e25274. <https://doi.org/10.1371/journal.pone.0025274>.
- Bernat, E. M., Nelson, L. D., Steele, V. R., Gehring, W. J., & Patrick, C. J. (2011). Externalizing psychopathology and gain-loss feedback in a simulated gambling task: Dissociable components of brain response revealed by time–frequency analysis. *Journal of Abnormal Psychology*, 120(2), 352–364. <https://doi.org/10.1037/a0022124>.
- Born, J., Lange, T., Kern, W., McGregor, G. P., Bickel, U., & Fehm, H. L. (2002). Sniffing neuropeptides: A transnasal approach to the human brain. *Nature Neuroscience*, 5(6), 514–516. <https://doi.org/10.1038/mn849>.
- Brunia, C. H., & Damen, E. J. (1988). Distribution of slow brain potentials related to motor preparation and stimulus anticipation in a time estimation task. *Electroencephalography and Clinical Neurophysiology*, 69(3), 234–243. [https://doi.org/10.1016/0013-4694\(88\)90132-0](https://doi.org/10.1016/0013-4694(88)90132-0).
- Brunia, C. H., Hackley, S. A., van Boxtel, G. J., Kotani, Y., & Ohgami, Y. (2011). Waiting to perceive: Reward or punishment?. *Clinical Neurophysiology*, 122(5), 858–868. <https://doi.org/10.1016/j.clinph.2010.12.039>.
- Brunia, C. H., van Boxtel, G. J., & Böcker, K. B. (2012). Negative slow waves as indices of anticipation: The Bereitschaftspotential, the contingent negative variation, and the stimulus-preceding negativity. In S. J. Luck & E. S. Kappenman (Eds.), *The Oxford handbook of event-related potential components* (pp. 189–207). New York: Oxford University Press. <https://doi.org/10.1093/oxfordhdb/9780195374148.013.0108>.
- Brunnlieb, C., Nave, G., Camerer, C. F., Schosser, S., Vogt, B., Münte, T. F., & Heldmann, M. (2016). Vasopressin increases human risky cooperative behavior. *Proceedings of the National Academy of Sciences of the United States of America*, 113(8), 2051–2056. <https://doi.org/10.1073/pnas.1518825113>.
- Caldwell, H. K. (2017). Oxytocin and vasopressin: Powerful regulators of social behavior. *The Neuroscientist: A Review Journal Bringing Neurobiology, Neurology and Psychiatry*, 23(5), 517–528. <https://doi.org/10.1177/1073858417708284>.
- Caldwell, H. K., & Albers, H. E. (2004). Effect of photoperiod on vasopressin-induced aggression in Syrian hamsters. *Hormones and Behavior*, 46(4), 444–449. <https://doi.org/10.1016/j.yhbeh.2004.04.006>.
- Carpenter, B., Gelman, A., Hoffman, M. D., Lee, D., Goodrich, B., Betancourt, M., ... Riddell, A. (2017). Stan: A probabilistic programming language. *Journal of Statistical Software*, 76(1), 1–32. <https://doi.org/10.18637/jss.v076.i01>.
- Cavanagh, J. F. (2015). Cortical delta activity reflects reward prediction error and related behavioral adjustments, but at different times. *NeuroImage*, 110, 205–216. <https://doi.org/10.1016/j.neuroimage.2015.02.007>.
- Cavanagh, J. F., & Frank, M. J. (2014). Frontal theta as a mechanism for cognitive control. *Trends in Cognitive Sciences*, 18(8), 414–421. <https://doi.org/10.1016/j.tics.2014.04.012>.
- Cavanagh, J. F., Frank, M. J., Klein, T. J., & Allen, J. J. (2010). Frontal theta links prediction errors to behavioral adaptation in reinforcement learning. *NeuroImage*, 49(4), 3198–3209. <https://doi.org/10.1016/j.neuroimage.2009.11.080>.
- Cohen, M. X. (2011). Error-related medial frontal theta activity predicts cingulate-related structural connectivity. *NeuroImage*, 55(3), 1373–1383. <https://doi.org/10.1016/j.neuroimage.2010.12.072>.



- Cohen, M. X., Elger, C. E., & Ranganath, C. (2007). Reward expectation modulates feedback-related negativity and EEG spectra. *NeuroImage*, 35(2), 968–978. <https://doi.org/10.1016/j.neuroimage.2006.11.056>.
- Crawley, D., Zhang, L., Jones, E. J., Ahmad, J., Oakley, B., & San Jose Caceres, A., ... EU-AIMS LEAP group. (2020). Modeling flexible behavior in childhood to adulthood shows age-dependent learning mechanisms and less optimal learning in autism in each age group. *PLoS Biology*, 18(10), e3000908. <https://doi.org/10.1371/journal.pbio.3000908>.
- Crockett, M. J., Kurth-Nelson, Z., Siegel, J. Z., Dayan, P., & Dolan, R. J. (2014). Harm to others outweighs harm to self in moral decision making. *Proceedings of the National Academy of Sciences of the United States of America*, 111(48), 17320–17325. <https://doi.org/10.1073/pnas.1408988111>.
- Delorme, A., & Makeig, S. (2004). EEGLAB: An open source toolbox for analysis of single-trial EEG dynamics including independent component analysis. *Journal of Neuroscience Methods*, 134(1), 9–21. <https://doi.org/10.1016/j.jneumeth.2003.10.009>.
- Dhuria, S. V., Hanson, L. R., & Frey, W. H., 2nd (2010). Intranasal delivery to the central nervous system: Mechanisms and experimental considerations. *Journal of Pharmaceutical Sciences*, 99(4), 1654–1673. <https://doi.org/10.1002/jps.21924>.
- Dotd, C., Pietrowsky, R., Sewing, A., Zabel, A., Fehm, H. L., & Born, J. (1994). Effects of vasopressin on event-related potential indicators of cognitive stimulus processing in young and old humans. *Journals of Gerontology*, 49(4), M183–M188. <https://doi.org/10.1093/geronj/49.4.M183>.
- Faul, F., Erdfelder, E., Lang, A. G., & Buchner, A. (2007). G\* Power 3: A flexible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behavior Research Methods*, 39(2), 175–191. <https://doi.org/10.3758/bf03193146>.
- Fehr, E., & Fischbacher, U. (2003). The nature of human altruism. *Nature*, 425(6960), 785–791. <https://doi.org/10.1038/nature02043>.
- Feng, C., Hackett, P. D., DeMarco, A. C., Chen, X., Stair, S., Haroon, E., ... Rilling, J. K. (2015). Oxytocin and vasopressin effects on the neural response to social cooperation are modulated by sex in humans. *Brain Imaging and Behavior*, 9(4), 754–764. <https://doi.org/10.1007/s11682-014-9333-9>.
- Feng, C., Qin, L., Luo, Y., & Xu, P. (2020). Intranasal vasopressin expedites dishonesty in women. *Hormones and Behavior*, 126, 104843. <https://doi.org/10.1016/j.yhbeh.2020.104843>.
- Fiske, S. T., & Taylor, S. E. (2013). *Social cognition: From brains to culture*. Los Angeles: Sage Publications.
- Frank, M. J., Seeberger, L. C., & O'Reilly, R. C. (2004). By carrot or by stick: Cognitive reinforcement learning in Parkinsonism. *Science*, 306(5703), 1940–1943. <https://doi.org/10.1126/science.1102941>.
- Galea, J. M., Mallia, E., Rothwell, J., & Diedrichsen, J. (2015). The dissociable effects of punishment and reward on motor learning. *Nature Neuroscience*, 18(4), 597–602. <https://doi.org/10.1038/nn.3956>.
- Gehring, W. J., Goss, B., Coles, M. G., Meyer, D. E., & Donchin, E. (1993). A neural system for error detection and compensation. *Psychological Science*, 4(6), 385–390. <https://doi.org/10.1111/j.1467-9280.1993.tb00586.x>.
- Gobrogge, K. L., Liu, Y., Young, L. J., & Wang, Z. (2009). Anterior hypothalamic vasopressin regulates pair-bonding and drug-induced aggression in a monogamous rodent. *Proceedings of the National Academy of Sciences of the United States of America*, 106(45), 19144–19149. <https://doi.org/10.1073/pnas.0908620106>.
- Hackley, S. A., Valle-Inclán, F., Masaki, H., & Hebert, K. (2014). Stimulus-preceding negativity (SPN) and attention to rewards. In G. R. Mangun (Ed.), *Cognitive electrophysiology of attention: Signals of the mind* (pp. 216–225). San Diego: Elsevier. <https://doi.org/10.1016/B978-0-12-398451-7.00017-8>.
- Hauser, T. U., Iannaccone, R., Stämpfli, P., Drechsler, R., Brandeis, D., Walitza, S., & Brem, S. (2014). The feedback-related negativity (FRN) revisited: New insights into the localization, meaning and network organization. *NeuroImage*, 84, 159–168. <https://doi.org/10.1016/j.neuroimage.2013.08.028>.
- Holroyd, C. B., & Coles, M. G. (2002). The neural basis of human error processing: Reinforcement learning, dopamine, and the error-related negativity. *Psychological Review*, 109(4), 679–709. <https://doi.org/10.1037/0033-295X.109.4.679>.
- Holroyd, C. B., Krigolson, O. E., Baker, R., Lee, S., & Gibson, J. (2009). When is an error not a prediction error? An electrophysiological investigation. *Cognitive, Affective and Behavioral Neuroscience*, 9(1), 59–70. <https://doi.org/10.3758/CABN.9.1.59>.
- Holroyd, C. B., Nieuwenhuis, S., Yeung, N., & Cohen, J. D. (2003). Errors in reward prediction are reflected in the event-related brain potential. *NeuroReport*, 14(18), 2481–2484. <https://doi.org/10.1097/00001756-200312190-00037>.
- Hu, L., & Zhang, Z. (2019). *EEG signal processing and feature extraction*. Singapore: Springer Singapore.
- Hu, X., Xu, Z., & Mai, X. (2017). Social value orientation modulates the processing of outcome evaluation involving others. *Social Cognitive and Affective Neuroscience*, 12(11), 1730–1739. <https://doi.org/10.1093/scan/nsx102>.
- Jennings, J. R., & Wood, C. C. (1976). The  $\epsilon$ -adjustment procedure for repeated-measures analyses of variance. *Psychophysiology*, 13(3), 277–278. <https://doi.org/10.1111/j.1469-8986.1976.tb00116.x>.
- Knafo, A., Israel, S., Darvasi, A., Bachner-Melman, R., Uzevovsky, F., Cohen, L., ... Ebstein, R. P. (2008). Individual differences in allocation of funds in the dictator game associated with length of the arginine vasopressin 1a receptor RS3 promoter region and correlation between RS3 length and hippocampal mRNA. *Genes, Brain and Behavior*, 7(3), 266–275. <https://doi.org/10.1111/j.1601-183X.2007.00341.x>.
- Liao, Z., Huang, L., & Luo, S. (2021). Intranasal oxytocin decreases self-oriented learning. *Psychopharmacology*, 238(2), 461–474. <https://doi.org/10.1007/s00213-020-05694-7>.
- Liu, J., Gu, R., Liao, C., Lu, J., Fang, Y., Xu, P., ... Cui, F. (2020). The neural mechanism of the social framing effect: Evidence from fMRI and tDCS studies. *Journal of Neuroscience*, 40(18), 3646–3656. <https://doi.org/10.1523/JNEUROSCI.1385-19.2020>.
- Liu, Y., Curtis, J. T., & Wang, Z. (2001). Vasopressin in the lateral septum regulates pair bond formation in male prairie voles (*Microtus ochrogaster*). *Behavioral Neuroscience*, 115(4), 910–919. <https://doi.org/10.1037/0735-7044.115.4.910>.
- Lock, M. P. (2008). Treatment of antisocial personality disorder. *The British Journal of Psychiatry*, 193(5), 426. <https://doi.org/10.1192/bjp.193.5.426>.
- Lockwood, P. L., Apps, M. A., Valton, V., Viding, E., & Roiser, J. P. (2016). Neurocomputational mechanisms of prosocial learning and links to empathy. *Proceedings of the National Academy of Sciences of the United States of America*, 113(35), 9763–9768. <https://doi.org/10.1073/pnas.1603198113>.
- Lockwood, P. L., Klein-Flügge, M. C., Abdurahman, A., & Crockett, M. J. (2020). Model-free decision making is prioritized when learning to avoid harming others. *Proceedings of the National Academy of Sciences of the United States of America*, 117(44), 27719–27730. <https://doi.org/10.1073/pnas.2010890117>.
- Martins, D., Lockwood, P., Cutler, J., Moran, R., & Paloyelis, Y. (2022). Oxytocin modulates neurocomputational mechanisms underlying prosocial reinforcement learning. *Progress in Neurobiology*, 213, 102253. <https://doi.org/10.1016/j.pneurobio.2022.102253>.
- Masaki, H., Takeuchi, S., Gehring, W. J., Takasawa, N., & Yamazaki, K. (2006). Affective-motivational influences on feedback-related ERPs in a gambling task. *Brain Research*, 1105(1), 110–121. <https://doi.org/10.1016/j.brainres.2006.01.022>.
- Masaki, H., Yamazaki, K., & Hackley, S. A. (2010). Stimulus-preceding negativity is modulated by action–outcome contingency. *NeuroReport*, 21(4), 277–281. <https://doi.org/10.1097/WNR.0b013e3283360bc3>.
- Miltner, W. H., Braun, C. H., & Coles, M. G. (1997). Event-related brain potentials following incorrect feedback in a time-estimation task: Evidence for a 'generic' neural system for error detection. *Journal of Cognitive Neuroscience*, 9(6), 788–798. <https://doi.org/10.1162/jocn.1997.9.6.788>.
- Morís, J., Luque, D., & Rodríguez-Fornells, A. (2013). Learning-induced modulations of the stimulus-preceding negativity. *Psychophysiology*, 50(9), 931–939. <https://doi.org/10.1111/psyp.12073>.
- Moskowitz, G. B. (2005). *Social cognition: Understanding self and others*. New York: Guilford Press.
- Neto, M. L., Antunes, M., Lopes, M., Ferreira, D., Rilling, J., & Prata, D. (2020). Oxytocin and vasopressin modulation of prisoner's dilemma strategies.

- Journal of Psychopharmacology*, 34(8), 891–900. <https://doi.org/10.1177/0269881120913145>.
- Nieuwenhuis, S., Aston-Jones, G., & Cohen, J. D. (2005). Decision making, the P3, and the locus coeruleus-norepinephrine system. *Psychological Bulletin*, 131(4), 510–532. <https://doi.org/10.1037/0033-2909.131.4.510>.
- Nishina, K., Takagishi, H., Takahashi, H., Sakagami, M., & Inoue-Murayama, M. (2019). Association of polymorphism of arginine-vasopressin receptor 1A (AVPR1A) gene with trust and reciprocity. *Frontiers in Human Neuroscience*, 13, 230. <https://doi.org/10.3389/fnhum.2019.00230>.
- Osinsky, R., Mussel, P., & Hewig, J. (2012). Feedback-related potentials are sensitive to sequential order of decision outcomes in a gambling task. *Psychophysiology*, 49(12), 1579–1589. <https://doi.org/10.1111/j.1469-8986.2012.01473.x>.
- Osinsky, R., Walter, H., & Hewig, J. (2014). What is and what could have been: An ERP study on counterfactual comparisons. *Psychophysiology*, 51(8), 773–781. <https://doi.org/10.1111/psyp.12221>.
- Patel, N., Grillon, C., Pavletic, N., Rosen, D., Pine, D. S., & Ernst, M. (2015). Oxytocin and vasopressin modulate risk-taking. *Physiology and Behavior*, 139, 254–260. <https://doi.org/10.1016/j.physbeh.2014.11.018>.
- Pearce, J. M., & Hall, G. (1980). A model for Pavlovian learning: Variations in the effectiveness of conditioned but not of unconditioned stimuli. *Psychological Review*, 87(6), 532–552. <https://doi.org/10.1037/0033-295X.87.6.532>.
- Pfabigan, D. M., Alexopoulos, J., Bauer, H., & Sailer, U. (2011). Manipulation of feedback expectancy and valence induces negative and positive reward prediction error signals manifest in event-related brain potentials. *Psychophysiology*, 48(5), 656–664. <https://doi.org/10.1111/j.1469-8986.2010.01136.x>.
- Pitkow, L. J., Sharer, C. A., Ren, X., Insel, T. R., Terwilliger, E. F., & Young, L. J. (2001). Facilitation of affiliation and pair-bond formation by vasopressin receptor gene transfer into the ventral forebrain of a monogamous vole. *Journal of Neuroscience*, 21(18), 7392–7396. <https://doi.org/10.1523/jneurosci.21-18-07392.2001>.
- Pulcu, E., & Browning, M. (2017). Affective bias as a rational response to the statistics of rewards and punishments. *eLife*, 6, e27879. <https://doi.org/10.7554/eLife.27879>.
- Qi, Y., Wu, H., Raiha, S., & Liu, X. (2018). Social value orientation modulates context-based social comparison preference in the outcome evaluation: An ERP study. *Neuropsychologia*, 112, 135–144. <https://doi.org/10.1016/j.neuropsychologia.2018.02.028>.
- Rescorla, R. A., & Wagner, A. R. (1972). A theory of Pavlovian conditioning: Variations in the effectiveness of reinforcement and nonreinforcement. In A. H. Black & W. F. Prokasy (Eds.), *Classical conditioning II: Current research and theory* (pp. 64–99). New York: Appleton-Century-Crofts.
- Rilling, J. K., DeMarco, A. C., Hackett, P. D., Chen, X., Gautam, P., Stair, S., ... Pagnoni, G. (2014). Sex differences in the neural and behavioral response to intranasal oxytocin and vasopressin during human social interaction. *Psychoneuroendocrinology*, 39, 237–248. <https://doi.org/10.1016/j.psyneuen.2013.09.022>.
- Rodrigues, J., Ulrich, N., & Hewig, J. (2015). A neural signature of fairness in altruism: A game of theta?. *Social Neuroscience*, 10(2), 192–205. <https://doi.org/10.1080/17470919.2014.977401>.
- Rozin, P., & Royzman, E. B. (2001). Negativity bias, negativity dominance, and contagion. *Personality and Social Psychology Review*, 5(4), 296–320. [https://doi.org/10.1207/S15327957PSPR0504\\_2](https://doi.org/10.1207/S15327957PSPR0504_2).
- Sharot, T., & Garrett, N. (2016). Forming beliefs: Why valence matters. *Trends in Cognitive Sciences*, 20(1), 25–33. <https://doi.org/10.1016/j.tics.2015.11.002>.
- Shin, Y. S., & Niv, Y. (2021). Biased evaluations emerge from inferring hidden causes. *Nature Human Behaviour*, 5(9), 1180–1189. <https://doi.org/10.1038/s41562-021-01065-0>.
- Song, Z., Larkin, T. E., Malley, M. O., & Albers, H. E. (2016). Oxytocin (OT) and arginine-vasopressin (AVP) act on OT receptors and not AVP V1a receptors to enhance social recognition in adult Syrian hamsters (*Mesocricetus auratus*). *Hormones and Behavior*, 81, 20–27. <https://doi.org/10.1016/j.yhbeh.2016.02.004>.
- Song, Z., McCann, K. E., McNeill, J. K., 4th, Larkin, T. E., 2nd, Huhman, K. L., & Albers, H. E. (2014). Oxytocin induces social communication by activating arginine-vasopressin V1a receptors and not oxytocin receptors. *Psychoneuroendocrinology*, 50, 14–19. <https://doi.org/10.1016/j.psyneuen.2014.08.005>.
- Sutton, R. S., & Barto, A. G. (2018). *Reinforcement learning: An introduction*. London: MIT Press.
- Team, S. D. (2016). RStan: The R interface to Stan. R Package Version, 2(1), 522.
- Tendler, A., & Wagner, S. (2015). Different types of theta rhythmicity are induced by social and fearful stimuli in a network associated with social memory. *eLife*, 4, e03614. <https://doi.org/10.7554/eLife.03614>.
- Thompson, R. R., George, K., Walton, J. C., Orr, S. P., & Benson, J. (2006). Sex-specific influences of vasopressin on human social communication. *Proceedings of the National Academy of Sciences of the United States of America*, 103(20), 7889–7894. <https://doi.org/10.1073/pnas.0600406103>.
- Tversky, A., & Kahneman, D. (1991). Loss aversion in riskless choice: A reference-dependent model. *The Quarterly Journal of Economics*, 106(4), 1039–1061. <https://doi.org/10.2307/2937956>.
- Uzefovsky, F., Shalev, I., Israel, S., Knafo, A., & Ebstein, R. P. (2012). Vasopressin selectively impairs emotion recognition in men. *Psychoneuroendocrinology*, 37(4), 576–580. <https://doi.org/10.1016/j.psyneuen.2011.07.018>.
- van den Berg, P., Molleman, L., & Weissing, F. J. (2015). Focus on the success of others leads to selfish behavior. *Proceedings of the National Academy of Sciences of the United States of America*, 112(9), 2912–2917. <https://doi.org/10.1073/pnas.1417203112>.
- van Dijk, E., & Wilke, H. (2000). Decision-induced focusing in social dilemmas: Give-some, keep-some, take-some, and leave-some dilemmas. *Journal of Personality and Social Psychology*, 78(1), 92–104. <https://doi.org/10.1037/0022-3514.78.1.92>.
- Walsh, M. M., & Anderson, J. R. (2012). Learning from experience: Event-related potential correlates of reward processing, neural adaptation, and behavioral choice. *Neuroscience and Biobehavioral Reviews*, 36(8), 1870–1884. <https://doi.org/10.1016/j.neubiorev.2012.05.008>.
- Wang, J., Qin, W., Liu, F., Liu, B., Zhou, Y., Jiang, T., & Yu, C. (2016). Sex-specific mediation effect of the right fusiform face area volume on the association between variants in repeat length of AVPR1A RS3 and altruistic behavior in healthy adults. *Human Brain Mapping*, 37(7), 2700–2709. <https://doi.org/10.1002/hbm.23203>.
- Wilson, R. C., & Collins, A. G. (2019). Ten simple rules for the computational modeling of behavioral data. *eLife*, 8, e49547. <https://doi.org/10.7554/eLife.49547>.
- Winslow, J. T., Hastings, N., Carter, C. S., Harbaugh, C. R., & Insel, T. R. (1993). A role for central vasopressin in pair bonding in monogamous prairie voles. *Nature*, 365(6446), 545–548. <https://doi.org/10.1038/365545a0>.
- Yeung, N., Holroyd, C. B., & Cohen, J. D. (2005). ERP correlates of feedback and reward processing in the presence and absence of response choice. *Cerebral Cortex*, 15(5), 535–544. <https://doi.org/10.1093/cercor/bhh153>.
- Zheng, Y., Li, Q., Wang, K., Wu, H., & Liu, X. (2015). Contextual valence modulates the neural dynamics of risk processing. *Psychophysiology*, 52(7), 895–904. <https://doi.org/10.1111/psyp.12415>.
- Zink, C. F., Kempf, L., Hakimi, S., Rainey, C. A., Stein, J. L., & Meyer-Lindenberg, A. (2011). Vasopressin modulates social recognition-related activity in the left temporoparietal junction in humans. *Translational Psychiatry*, 1(4), e3. <https://doi.org/10.1038/tp.2011.2>.