NeuroImage 169 (2018) 323–333

Contents lists available at ScienceDirect

NeuroImage

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

Chronic nicotine exposure impairs uncertainty modulation on reinforcement learning in anterior cingulate cortex and serotonin system

Zhengde Wei [a,b,1], Long Han [a,1], Xiuying Zhong [c], Ying Liu [d], Rujing Zha [a], Ying Wang [a], Li-Zhuang Yang [a], Junjie Bu [a], Hongwen Song [a], Wenjuan Wang [e,f], Yifeng Zhou [a], Ping Gao [c], Xiaochu Zhang [a,g,h,i,*]

[a] Key Laboratory of Brain Function and Disease, Chinese Academy of Sciences, School of Life Sciences, University of Science & Technology of China, Hefei, Anhui 230027, China
[b] Shanghai Key Laboratory of Psychotic Disorders, Shanghai Mental Health Center, Shanghai Jiao Tong University School of Medicine, Shanghai 200030, China
[c] The CAS Key Laboratory of Innate Immunity and Chronic Disease, Innovation Center for Cell Biology, School of Life Sciences, University of Science and Technology of China, Hefei, China
[d] Provincial Hospital Affiliated to Anhui Medical University, Hefei, Anhui, 230001, China
[e] School of Science, Anhui Agricultural University, Hefei, Anhui, 230036, China
[f] School of Information Science and Technology, University of Science and Technology of China, Hefei, Anhui, 230026, China
[g] School of Humanities & Social Science, University of Science & Technology of China, Hefei, Anhui 230026, China
[h] Centers for Biomedical Engineering, University of Science & Technology of China, Hefei, Anhui 230027, China
[i] Anhui Mental Health Center, Hefei, Anhui, 230022, China

ARTICLE INFO

Keywords:
Reward prediction errors
Serotonin
Maladaptive decision-making
Addiction

ABSTRACT

Deficits in the computational processes of reinforcement learning have been suggested to underlie addiction. Additionally, environmental uncertainty, which is encoded in the anterior cingulate cortex (ACC), modulates reward prediction errors (RPEs) during reinforcement learning and exacerbates addiction. The present study tested whether and how the ACC would have an essential role in drug addiction by failing to use uncertainty to modulate the RPEs during reinforcement learning. In Experiment I, we found that the ACC/medial prefrontal cortex (MPFC) did not modulate RPE learning according to uncertainty in smokers. The effect of uncertainty × RPE in the ACC/MPFC was correlated with the learning rate of RPEs and the duration of nicotine use. Experiment II demonstrated that serotonin, but not dopamine, receptor mRNA expression significantly decreased in the ACC of the nicotine exposed compared to the control rats. Furthermore, there was a positive correlation between learning rate and serotonin receptor mRNA expression in the ACC. Therefore, all present results suggest that impairments in uncertainty modulation in the ACC disrupt reinforcement learning processes in chronic nicotine users and contribute to maladaptive decision-making. These findings support interventions for pathological decision-making in drug addiction that strongly focus on the serotonin system in ACC.

Introduction

Addiction can be operationally defined as continued maladaptive choices associated with urges to seek and take a drug, even with knowledge of the harmful consequences (DSM-V; American Psychiatric Association, 2013; Tanabe et al., 2013). Abnormal reinforcement learning processes may underlie addiction (Redish, 2004).

Assessment of environmental uncertainty is a fundamental component modulating the process of reinforcement learning (Schultz, 2011). Furthermore, environmental uncertainty induced by incomplete knowledge of a reward environment is important to optimal decision-making (Behrens et al., 2007). Previous studies found that environmental uncertainty modulates reward prediction errors (RPEs), which represents the difference between predicted and experienced outcomes, and are encoded in the anterior cingulate cortex (ACC) (Behrens et al., 2007). After ACC lesions in monkeys, only the most

* Corresponding author. Key Laboratory of Brain Function and Disease, Chinese Academy of Sciences, School of Life Sciences, University of Science & Technology of China, Hefei, Anhui 230027, China.
E-mail address: zxcustc@ustc.edu.cn (X. Zhang).
[1] These authors contributed equally to this work.

https://doi.org/10.1016/j.neuroimage.2017.11.048
Received 10 May 2017; Received in revised form 4 November 2017; Accepted 21 November 2017
Available online 6 December 2017

recent outcome influenced subsequent decisions, indicating that ACC dysfunction alters the influence of environmental uncertainty on learning (Kennerley et al., 2006).

There is evidence that environmental uncertainty influences drug consumption, suggesting that it exacerbates addiction (Zhang, 2014). In one review, Schultz (2011) proposed a hypothesis that addictive drugs likely influence environmental uncertainty signals, which distort the effects of RPEs and enhance the associability terms of learning rules. However, no direct evidence has been reported to support this hypothesis. Thus, the present study investigated whether and how the ACC would contribute to drug addiction by failing to use uncertainty to modulate RPEs during reinforcement learning.

Besides neuroimaging studies, the biochemical and molecular mechanisms of impairments in environmental uncertainty modulation in the ACC during drug addiction have not been elucidated. However, modulation of environmental uncertainty is profoundly influenced by dopamine (DA) (Niv et al., 2005; Linnet et al., 2012; Mikhael and Bogacz, 2016) and serotonin (5-HT) (Murphy et al., 2009; Cools et al., 2011; Rogers, 2011). A proposed model suggests that the DA levels are critical contributors to reward uncertainty (Friston, 2005; Tobler et al., 2005; Mikhael and Bogacz, 2016). Additionally, the 5-HT system innervates specific forebrain areas that are important to regulating learning processes (Clarke et al., 2004). Performance on the Iowa Gambling Task (IGT), which requires that decisions are made under uncertainty, is associated with the 5-HTTLPR genotype (Stoltenberg and Vandever, 2010). Furthermore, several studies have demonstrated associations between low 5-HT and drug addiction (Balasubramani et al., 2015; Muller and Homberg, 2015). However, current data have not elucidated whether impairment in environmental uncertainty modulation during drug addiction is associated with 5-HT or DA in the ACC.

We tested our hypotheses using two experiments. Experiment I tested human participants using a modified version of the IGT (Bechara et al., 1994),which is a classical decision-making task under uncertainty, and functional magnetic resonance imaging. This experiment investigated whether ACC dysfunctions in modulating RPE learning according to environmental uncertainty occurred in smokers. In Experiment II, we used a Rat Gambling Task (RGT) to test whether decreased DA or 5-HT receptor mRNA expression in the ACC was associated with reduced RPE learning in rats with chronic nicotine exposure.

Materials and methods

Experiment I

Participants

Internet advertisements were used to recruit potential participants. Thirty-eight smokers (SMO; all men; mean age: 24.5 ± 2.3 years; mean education: 16.2 ± 1.9 years; >10 cigarettes/d for at least one year) and 37 non-smokers (CON; all men; mean age: 23.7 ± 2.3 years; mean education: 16.7 ± 1.9 years; never smoked before) were included. All participants had normal or corrected-to-normal vision and were right-handed. Potential participants were excluded if they reported a diagnosis of a psychiatric disorder, use of addictive drugs (except nicotine in the smoker group), a prior head injury, or any contraindications for participating in a magnetic resonance imaging (MRI) study (e.g., non-removable metallic implants). After having smoked their last cigarette, smokers filled in consent forms and relative questionnaires, read the instruction file and underwent task practice under the supervision of an experimenter, which lasted for about 2 h. Thus, smoking cessation was required from the smokers were during this 2 h session. Severity of nicotine dependence was assessed using the Fagerström Test of Nicotine Dependence [FTND] (Heatherton et al., 1991). The experiment was approved by the Research Ethics Committee of the University of Science and Technology of China, and written informed consent was obtained from all participants, consistent with the Declaration of Helsinki. The methods were carried out in accordance with the approved guidelines.

Task and stimuli

During functional MRI (fMRI) acquisition, participants performed a modified version of the IGT (as described previously (Wang et al., 2017), Fig. 1). In brief, the participants were informed that the aim of the task was to win as many points as possible by selecting one card at a time from four decks; they would be paid 10¥/1000 points.

The task involved four decks of cards labeled A, B, C, and D. Selecting a card from deck A and B would secure 100 points, while selecting a card from decks C and D, 50 points. However, at times selecting a card from any of the four decks could also result in a loss. Every set of 10 cards from decks A or B secured 1000 points but resulted in the loss of 1250 points in unpredictable punishments (for every set of 10 cards, deck A contained 5 losses: 150, 200, 250, 300, 350 points; deck B contained 1 loss: 1250 points). On the other hand, every one set of 10 cards from decks C or D secured 500 points but deprived of 250 points in unpredictable punishments (for every set of 10 cards, deck C contained 5 losses: 25, 40, 50, 60, 75 points; deck D contained 1 loss: 250 points). Therefore, decks A and B were disadvantageous because of a net loss (−250 points/10 cards), while decks C and D were advantageous because of a net gain (+250 points/10 cards). The participants were instructed that some decks were advantageous while the others were disadvantageous; however, all participants were not aware of the relative frequency or magnitude of gains/punishments before the experiment.

Similar to our previous study (Wang et al., 2017), the task was extended to 180 trials from the original 100 trials in Bechara's study (1994), in order to facilitate rule learning. The net score was then obtained by calculating the difference between the advantageous and disadvantageous decks (C + D) - (A + B) in each block of 20 trials, and the total net score (TNS) was obtained by calculating the difference between the advantageous and disadvantageous decks (C + D) - (A + B) for 180 trials. The whole task was divided into three scan runs. Each scan run consisted of three blocks separated by a 30-s resting interval, and each task block consisted of 20 trials. Each trial was divided into two phases: (1) the choice phase (4 s), where the participants could see the four decks and make a choice, or the computer would make a random choice after 4 s (our participants missed only 1.03 trials on an average, and we excluded these trials from the analysis below), and (2) the outcome phase (1 s), where the participants observed the outcome on the screen.

Behavioral analyses—reinforcement learning model

We analyzed behavioral and neural data using a reinforcement learning model adapted from previous studies (Behrens et al., 2007; Park et al., 2010; Tanabe et al., 2013). RPEs were included in this model, following the suggestion by Preushoff and Bossaert (Preuschoff and Bossaerts, 2007). A RPE (δ_t) was defined as the difference between the real reward r_t and the predicted reward $\widehat{v}_t$ at trial t.

$$\delta_t = r_t - \widehat{v}_t \tag{1}$$

The model uses the RPEs to update prediction of reward using the following equations:

$$\widehat{v}_{t+1} = \widehat{v}_t + \alpha \cdot \delta_t \tag{2}$$

The equation indicates that where α is the learning rate for RPEs (Behrens et al., 2007). Let $\widehat{\pi}_{it}$ indicate the probability of selecting option i at trial t. Calculated values of $\widehat{\pi}_{it}$ were determined by the following equation:

$$\pi_{it} = \frac{\widehat{e^{v_{it}}}}{\sum_{j=1}^{n} \widehat{e^{v_{jt}}}} \tag{3}$$

Uncertainty ($Uncertainty_t$) in each trial was quantified with Shannon's entropy (Bach et al., 2011), using the following equation:

$$Uncertainty_t = -\sum_{i=1}^{4} \pi_{it} \log_2 \pi_{it} \tag{4}$$

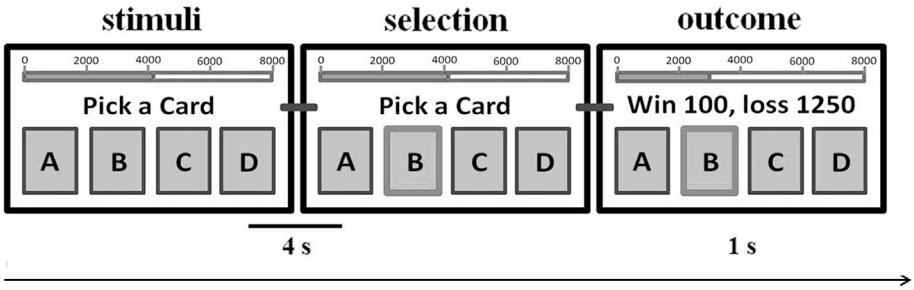

Fig. 1. Iowa Gambling Task design used in Experiment I. Each trial was divided into the two phases: (1) during the selection phase, 4 s were provided for choice pondering and selection, and if no choice was made within 4 s, the computer would make a random choice; (2) during the outcome phase, the outcome was presented on the screen for 1 s. After each trial, the next trial began immediately without inter-trial intervals (ITIs).

The learning rate was estimated separately by maximizing the log-likelihood for each participant:

$$\mathrm{MLL} = \max_{k,l} \sum_{t=1}^{180} \log \pi_{t,i_t} \tag{5}$$

with i_t, the deck selected at trial t, $t = 1, \ldots, 180$, $i_t \in \{1, 2, 3, 4\}$, and $\widehat{\pi}_{t,i_t}$, the probability for selecting deck i_t at trial t. The algorithms were processed in Matlab (R2008a) with statistical tools.

Parametric regression models (i.e., linear vs. quadratic) were tested to examine the relationship between the learning rate of RPEs and total net scores. We tested whether the learning rates of RPEs in smokers were lower than those for non-smokers, using a two-sample *t*-test. Group differences in the total net scores were analyzed. Further, we also performed new group comparisons of learning rates and total net scores, with age and education as covariates. We also assessed group differences in net scores over time, using a two (group) x nine (block) repeated-measures ANOVA, followed by Student-Newman-Keuls post hoc analyses.

MRI acquisition

Gradient echo-planar imaging data were acquired during the entire task, using a 3.0 T Trio scanner (Siemens Medical Solutions, Erlangen, Germany) with a circularly polarized head coil, at the Anhui Provincial Hospital. A T2*-weighted echo-planar imaging sequence (FOV = 240 mm, TE = 30 ms, TR = 2000 ms, flip angle = 85°, matrix = 64 × 64) with 33 axial slices (no gaps, voxel size: 3.75 × 3.75 × 3.7 mm^3) covering the whole brain was used to acquire the functional MR images. Before entering the scanner, participants were instructed to keep their heads still during all scans. Resting-state fMRI data (240 vol, 8 min) were acquired by asking the participants to keep their eyes closed and to remain at rest. Further, three functional scan runs occurred during the IGT, with each lasting 7 min. There was an interval of approximately 1 min between every two runs. High-resolution T1-weighted spin-echo imaging data were also acquired for anatomical overlays and three-dimensional gradient-echo imaging for stereotaxic transformations.

fMRI data analyses

All echo-planar imaging analyses were processed using Analysis of Functional Neuroimages (Cox, 1996) and MATLAB (version 7.6.0.324; Mathworks, Natick, MA). Functional data were realigned to the second volume. The realigned images were normalized to the Talairach coordinate. Raw data were corrected for temporal shifts between slices and for motion, spatially smoothed with a Gaussian kernel (full width at half maximum = 8 mm), and temporally normalized (for each voxel, the signal of each volume was divided by the temporally averaged signal).

To elucidate neural responses that correlated with ongoing RPEs under uncertainty, a parametric general linear model was used. Regressors of interest were RPEs and the interaction between RPEs and uncertainty. Regressors of no interest were stimulus onset, outcome onset, and uncertainty. RPEs, uncertainty, and interaction between RPEs and uncertainty regressors onset corresponded to the outcome phase. These regressors were convolved with a hemodynamic response function (HRF) and simultaneously regressed against the blood oxygenation level-dependent (BOLD) signal in each voxel. The regressors were not orthogonalized and there was no significant collinearity among the regressors. Six regressors for head motion were also included. Individual contrast images were analyzed for each regressor of interest's responses and second-level random effect analyses were conducted using one-sample *t*-tests to generate statistical z-score maps. Between-group whole brain comparisons were analyzed using two-sample *t*-tests. We set the family-wise error at a cluster-level threshold of $p < 0.01$ and the voxel-level threshold at $p < 0.0001$ (cluster size: 189 mm^3) for the whole brain analyses.

We defined the ACC/medial prefrontal cortex (MPFC) in which activity significantly correlated with interaction uncertainty × RPE and differed significantly between groups in the present study (see Result section) as the region of interest (ROI). Then, we used the program 3dmerge (https://afni.nimh.nih.gov/pub/dist/doc/program_help/3dmerge.html) to convert the ACC/MPFC into a ROI, and we took parameter estimates (z-score) in each participant from local average in a mask back-projected from the ROI. The differential responses (interaction uncertainty × RPE) in the ACC/MPFC were then used to predict individual learning rate of RPEs, severity of nicotine dependence (e.g., FTND score) and self-reported duration of nicotine use.

Functional connectivity analyses

To investigate functional connectivity of the ACC/MPFC during addiction, we used the "psychophysiological interaction" (PPI) (Friston et al., 1997). We extracted the entire experimental time series of each participant for the ACC/MPFC cluster, which significantly correlated with interaction uncertainty × RPE and differed significantly between groups. To create the PPI regressors, we multiplied these normalized time series data by condition vectors containing ones for 1 s during each feedback type (earn or lose outcome) and zeros under all other conditions in all the trials except the missing trials. These regressors were used as covariates in a separate regression analysis, which included regressors for earn and loss outcomes convolved with an HRF to explain feedback-related activity. The resulting parameter estimates (z-score) represented the degree to which activity in each voxel correlated with activity in the ACC/MPFC. Individual contrast images for connectivity during earn > loss feedback were calculated and analyzed using one- or two-sample *t*-tests ($p < 0.01$, family-wise error, corrected). To further examine whether the strength of functional connectivity was engaged in RPE-related processes, a correlation between the learning rate of RPEs and functional connectivity was examined, and a FDR-correction ($p < 0.05$) was employed for multiple comparisons.

We defined the left parahippocampus, left posterior cingulate cortex (PCC), and right PCC, which were significantly connected with the ACC/MPFC, with coupling connectivity with the ACC/MPFC that was significantly correlated with the learning rate of RPE as the ROI. Group differences in RPE-related functional connectivity in these ROIs were tested

using two-sample *t*-tests. Further, group differences in the relationship between learning rates and functional connectivity in these ROIs were tested using univariate ANOVAs. Then, we also performed new group comparisons above, with age and education as covariates.

To further test the intactness of RPE-related functional connectivity in smokers, we used a resting-state functional connectivity (RSFC) analysis. For the resting-state fMRI data, we regressed the motion data from the time series and performed bandpass temporal filtering (0.01–0.08 Hz) on the residual signals. Then, to further reduce nuisance signals, the average white matter and cerebral spinal fluid (CSF) signals were regressed out. The white matter mask was determined from the high-resolution structural images, using the FAST segmentation program (Functional MRI of the Brain software library; www.fmrib.ox.ac.uk). Each participant's CSF mask was manually drawn to fit the anatomical boundaries of a standardized Talairach brain atlas, which was transformed onto the image space of each participant and modified according to the cortical structures of the participant's brain by referencing the anatomical boundaries in the high-resolution structural image. These nuisance signals were used to account for fluctuations that were unlikely to be relevant to neuronal activity (Birn et al., 2006). Further, data scrubbing was performed, and any volume with a framewise-dependent value for the temporal derivative of a time course's root mean squared head motion variance exceeding 0.5 was excluded (Waheed et al., 2016). The percentage of average scrubbed time-points was 1.8% for all participants. We defined the left parahippocampus, left PCC, right PCC, and ACC/MPFC, which were our ROIs. The preprocessed resting-state fMRI time series were averaged within each ROI. Correlations between the averaged ROI time series were calculated and then transformed to Fisher z-values. Finally, the group differences in resting-state functional connectivity between the left parahippocampus, left PCC, right PCC, and ACC/MPFC were tested using two-sample *t*-tests. Then, we also performed new group comparisons above, with age and education as covariates.

Experiment II

Subjects

Subjects were male Sprague-Dawley rats (n = 27) weighing 240–260 g at the beginning of the experiments. The rats were housed in groups of four, in a temperature-controlled room (23 ± 2 °C) on a 12-h light/dark cycle. Food and drinking water were available ad libitum. Rats were acclimated to the housing conditions for at least one week and were subsequently handled for approximately 3 min/d for 5 d prior to experimental manipulations. Rats were randomly assigned to one of two matched groups, based on baseline body weights: a control group (n = 14) and a nicotine group (n = 13). The experiments were performed between 09:00 and 17:00. Animal housing, care, and application of experimental procedures were in accordance with the guidelines published in the European Communities Council Directive of November and were approved by the Institutional Animal Use and Care Committee of the University of Science and Technology of China. All efforts were made to minimize the discomfort of animals and to reduce the number of animals used.

Apparatus

Two identical operant chambers (adapted from five-choice serial reaction time chambers, AniLab Software & Instruments Co. Ltd. Ningbo, China) were used for the RGT. Four circular holes were available on a curved wall and could be dimly illuminated with a white light-emitting diode located behind them. A food tray on the opposite wall was connected to an external dispenser delivering water. All circular holes and the food tray had infrared beams crossing their entrances in order to detect responses. A clear vertical Plexiglas partition (28 cm × 0.5 cm × 30 cm) with a central opening (7 cm × 7 cm) was placed across the centre of the chamber, parallel to the food wall. This partition was equidistant (15 cm) to each nose-poke hole and to prevent thigmotaxic behavior. Each chamber was surrounded by a sound-attenuating cubicle, which was ventilated by a fan that provided low-level background noise. We used the SuperState software (Zhang, 2006) to control the inputs and outputs of the operant chambers, and to collect behavioral responses.

Nicotine treatments

Rats received subcutaneous vehicle or nicotine hydrogen tartrate dissolved in saline solution (NaCl 0.9%, 1 mg/ml, pH 7.4) at a dose of 0.6 mg/kg of nicotine in 1 ml/kg saline solution, once daily for 26 d. According to the recruitment criteria for smokers in the human study (smoking cigarettes continuously for at least 12 months) and the time of training in the RGT, we confirmed that the rats required subcutaneous nicotine for 26 d.

Rat Gambling Task

2-3 h after nicotine treatment, rats were trained or tested according to previously described methods (Rivalan et al., 2009). Water deprivation was maintained for 20 h/d such that body weights were maintained between 85 and 90% of the original weights, but food was freely available (Kelsey and Niraula, 2013). In order to support that learning the RGT is similar to the IGT, we designed a water reward version of the IGT (section I of Supplementary Information).

Training. Rats were initially trained to associate two consecutive nose pokes in the same hole with a reward (0.1 mL water). Rats could freely choose between the four holes during daily sessions until they completed 50 trials within one session (20 min cut-off). When a nose poke occurred, only that particular hole remained illuminated until the water was collected. The next trial began immediately, with all holes lighted. The training phase usually lasted 5–7 d and tests were performed the following day.

Test. The test consisted of a single 1 h session. Rats experienced all four task choice outcomes for the first time during the unique 1 h test session. In the RGT, rats had free access to the four nose poke holes (A-D), with each choice having different outcomes. Choice C or D resulted in immediate delivery of 0.05 mL water, whereas choice A or B resulted in immediate delivery of 0.1 mL water. Choices A and B were disadvantageous since they could be followed by unpredictable longer penalties consisting of time-outs during which no reward could be obtained (time-out of 222 and 444 s respectively) compared to the advantageous choices (C and D, 12 and 6 s). Penalties occurred at a low probability (¼) for B and C and a high probability (½) for A and D. During the penalty, the chosen hole remained illuminated to facilitate association between each choice and its consequences. A brief extinction of this light (1 s) signaled the end of the time-out. The four hole lights were then illuminated again to allow the rat to make a new choice. Trials not associated with a penalty had no time-out. The theoretical maximum gain was the same for choices (C and D) and was three times higher than for both choices (A and B; lower probabilities of a penalty are compensated by a longer duration of the penalty).

Quantitative real-time PCR (qRT-PCR)

One day after the RGT test, 5-HT receptor and DA receptor mRNA expression was assessed in the ACC, using qRT-PCR. Rats were euthanized by decapitation; the brain was rapidly removed from the head, and carefully cut into 2 mm coronal slices. The regions of the ACC of rats were quickly dissected on ice according to the coordinates from Paxinos and Watsons (Paxinos and Watsons, 2007), and stored at −80 °C until RNA extraction. The total RNA was extracted using TriZol agent (Invitrogen), following the manufacturer's protocol. Total RNA (1 μg) was used to synthesize cDNA with SuperScript III Reverse Transcriptase (Invitrogen). Additionally, qRT-PCR was performed using iQ SYBR Green Supermix and the iCycler Real-time System (Bio-Rad). The annealing temperature for each primer pair was optimized using temperature gradient PCR. The relative expression of individual transcripts was normalized to β-actin. The following primers were used: β-actin, upstream′-TTGCTGACAGGATGCAGAA-3′ and downstream 5′-ACCAATCCACACAGAGTACTT′; 5-HT_{1A} receptor′ upstream

5′-AGACCTCTCGAACCTGCCC-3′ and downstream 5′-AGACCTCTCGAACCTGCCC-3′; 5-HT_{2A} receptor, upstream 5′-AACGGTCCATCCACAGAG-3′ and downstream 5′-AACAGGAAGAACACGATGC-3′; 5-HT_{2C} receptor, upstream 5′-TTGGACTGAGGGACGAAAGC-3′ and downstream 5′-GGATGAAGAATGCCACGAAGG-3′; DA_1 receptor, upstream 5′-CGGGCTGCCAGCGGAGAG-3′ and downstream 5′-TGCCCAGGAGAGTGGACAGG-3′; DA_2 receptor, upstream 5′-CATTGTCTGGGTCCTGTCCT-3′ and downstream 5′-GACCAGCAGAGTGACGATGA-3′. All primers were purchased from Sango Biotech Company (Shanghai, China).

Data analyses

To analyze choice patterns throughout the RGT, the 60-min sessions were divided into six 10 min intervals. The percentage of advantageous choices = the sum of the two advantageous choices/total number of choices made * 100. All data were analyzed using repeated measures ANOVA with choice and/or session as the within-subjects factors, and group (nicotine vs. control) as the between-subjects factor. If the outcome of the main ANOVA yielded significant effects at the $p < 0.05$ level, Student-Newman-Keuls post hoc analyses were performed. Group differences in total earnings were analyzed. A confidence limit of $p < 0.05$ was considered statistically significant.

Relative target gene mRNA expression was determined using the ΔCt method. The resulting threshold cycle (Ct) values of the target gene were normalized to the Ct values for endogenous β-actin in the same samples: ΔCt = Ct (target gene) - Ct (β-actin). These values were further normalized to a control group sample: ΔΔCt = ΔCt (sample) – ΔCt (reference sample; ΔCt values of a random sample of the control group). Target gene mRNA expression relative to β-actin ($2^{-\Delta\Delta Ct}$) was then obtained. A two-sample *t*-test was used to assess group differences in gene mRNA expression. Correlations between the learning rate of RPEs and 5-HT or DA gene mRNA expression were calculated.

Experiment III assessed the effects of modified IGT using water as the reward in human participants (section I of Supplementary Information).

Results

Experiment I

Behavioral results

Parametric regression fit using a quadratic model (AIC: 508.124) was better than the fit using a linear model (AIC: 541.794) for analyzing the relationship between learning rates of RPEs and total net scores. Therefore, using a quadratic model, learning rates of RPEs were significantly associated with total net scores on the IGT (all participants: $F_{2,72} = 91.988$, $p < 0.001$ and $t_{72} = -6.622$, $p < 0.001$; CON: $F_{2,34} = 27.224$, $p < 0.001$ and $t_{34} = -4.276$, $p < 0.001$; SMO: $F_{2,35} = 49.604$, $p < 0.001$ and $t_{35} = -4.250$, $p < 0.001$; Fig. 2a–c). Smokers had lower learning rates of RPEs compared to non-smokers ($t_{73} = -1.707$, $p = 0.046$; Fig. 2d), also controlling the effects of age and education ($t_{71} = -1.688$, $p = 0.048$). Additionally, smokers had lower total net scores compared to non-smokers ($t_{73} = -3.270$, $p = 0.002$; Fig. 2e), also controlling the effects of age and education ($t_{71} = -2.892$, $p = 0.005$). A two (group) × nine (block) repeated-measures ANOVA, indicated significant main effects of group ($F_{1,657} = 33.223$, $p < 0.001$) and block ($F_{8,657} = 41.752$, $p < 0.001$), but no significant interaction ($F_{8,657} = 0.702$, $p = 0.690$). A post hoc analysis revealed that smokers performed significantly worse than non-smokers on five blocks (block_$3_{trial:41-60}$: $t_{73} = -2.708$, $p = 0.008$; block_$4_{trial:61-80}$: $t_{73} = -2.435$, $p = 0.017$; block_$7_{trial:121-140}$: $t_{73} = -2.361$, $p = 0.021$; block_$8_{trial:141-160}$: $t_{73} = -2.581$, $p = 0.012$; block_$9_{trial:161-180}$: $t_{73} = -2.234$, $p = 0.029$). The groups exhibited similar performance on

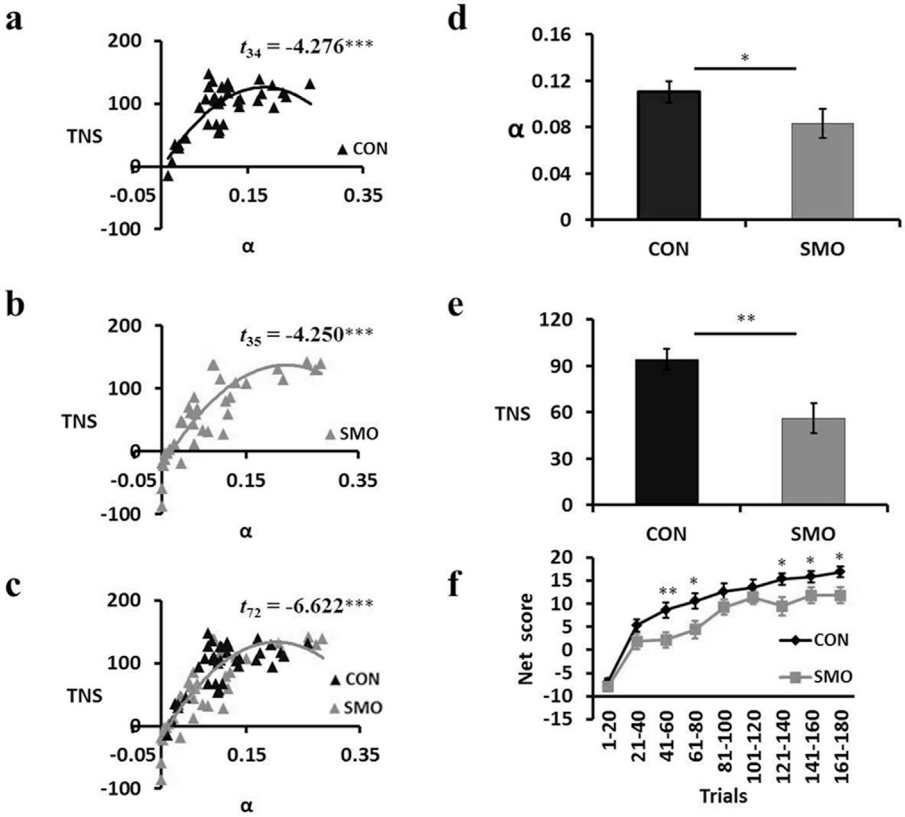

Fig. 2. Behavioral results showing that learning rates (i.e. α) were significantly correlated with total net scores (TNS) in non-smoking controls (CON) (a), smokers (SMO) (b), and all participants (c). The learning rates (d), TNS (e), and the net scores over time (f) in smokers were lower than in non-smoking controls. *$p < 0.05$, **$p < 0.01$, ***$p < 0.001$. For d, e and f, plotted data represent mean ± s. e.m. across participants.

the other four blocks (block_$1_{trial:1-20}$: $t_{73} = -0.817$, $p = 0.416$; block_$2_{trial:21-40}$: $t_{73} = -1.545$, $p = 0.127$; block_$5_{trial:81-100}$: $t_{73} = -1.225$, $p = 0.224$; block_6 $_{trial:101-120}$: $t_{73} = -1.139$, $p = 0.258$; Fig. 2f).

fMRI results

Neural substrates of RPEs. ***All participants.*** Individually generated RPEs were significantly correlated with BOLD signal in the bilateral striatum, ACC/MPFC, dorsolateral prefrontal cortex (DLPFC), PCC, and parahippocampus (Fig. S5).

Group differences. A whole brain group comparison revealed no differences between smokers and non-smokers. With ROI analysis, smokers had significantly weaker RPE-related activation in the bilateral striatum (section II of Supplementary Information; Fig. S3b).

Activation of uncertainty × RPE. ***All participants.*** Individually generated uncertainty × RPE were significantly correlated with BOLD signal in the bilateral ACC/MPFC, PCC, and temporal cortex (Fig. S6).

Group differences. A whole brain group comparison revealed significant differences in the ACC/MPFC (peak voxel coordinates: 1, −47, 20; corrected $p < 0.005$, voxel-level $p < 0.0001$) between smokers and non-smokers (Fig. 3a).

Additionally, we found that the interaction uncertainty × RPE effect size (z-score) in the ACC/MPFC significantly correlated with the learning rates of RPEs in all participants ($r = 0.350$, $p = 0.002$; Fig. 3b) and non-smokers ($r = 0.371$, $p = 0.024$; Fig. 3c), but not in smokers ($r = 0.247$, $p = 0.135$; Fig. 3d). Furthermore, longer duration of nicotine use was associated with weaker effect size in the ACC/MPFC ($r = -0.369$, $p = 0.023$; Fig. 3e). However, the FTND was not associated with the effect size in the ACC/MPFC ($r = 0.196$, $p = 0.239$).

Functional connectivity. ***All participants.*** A whole brain connectivity analysis demonstrated significant functional connectivity in the ACC/MPFC and the bilateral striatum, PCC, parahippocampus, and left dorsolateral prefrontal cortex during earn > loss outcomes (Fig. S7). Furthermore, learning rates were significantly correlated with functional connectivity between the left parahippocampus and the ACC/MPFC (all participants: $r = 0.340$, $p = 0.003$), between the left PCC and the ACC/MPFC (all participants: $r = 0.310$, $p = 0.007$), and between the right PCC and the ACC/MPFC (all participants: $r = 0.315$, $p = 0.006$) (Fig. 4). The correlations above survived after the correction for multiple comparisons (FDR-correction, $p < 0.05$).

Group differences. A whole brain group comparison of PPI revealed no differences between smokers and non-smokers. For the ROI analysis, no group differences were observed in RPE-related functional connectivity (left parahippocampus: $t_{73} = 1.211$, $p = 0.230$; left PCC: $t_{73} = 1.180$, $p = 0.242$; right PCC: $t_{73} = 0.938$, $p = 0.352$; Fig. 5), also controlling the effects of age and education (left parahippocampus: $t_{71} = 0.924$, $p = 0.359$; left PCC: $t_{71} = 0.894$, $p = 0.374$; right PCC: $t_{71} = 0.904$, $p = 0.369$). Additionally, no differences were observed between smokers and non-smokers for relationships between the learning rates and functional connectivity in the left parahippocampus and the ACC/MPFC ($F_{1,74} = 1.203$, $p = 0.276$), the left PCC and the ACC/MPFC ($F_{1,74} = 0.695$, $p = 0.407$), or the right PCC and the ACC/MPFC ($F_{1,74} = 0.017$, $p = 0.897$) (Fig. 5).

Furthermore, there were no differences between smokers and non-smokers in RSFC of the left parahippocampus and the ACC/MPFC ($t_{73} = 0.852$, $p = 0.397$), the left PCC and the ACC/MPFC ($t_{73} = -0.293$, $p = 0.771$), or the right PCC and the ACC/MPFC ($t_{73} = -0.171$, $p = 0.865$), also controlling the effects of age and education (left parahippocampus: $t_{71} = 0.262$, $p = 0.794$; left PCC: $t_{71} = -0.716$, $p = 0.476$; right PCC: $t_{71} = -0.550$, $p = 0.584$).

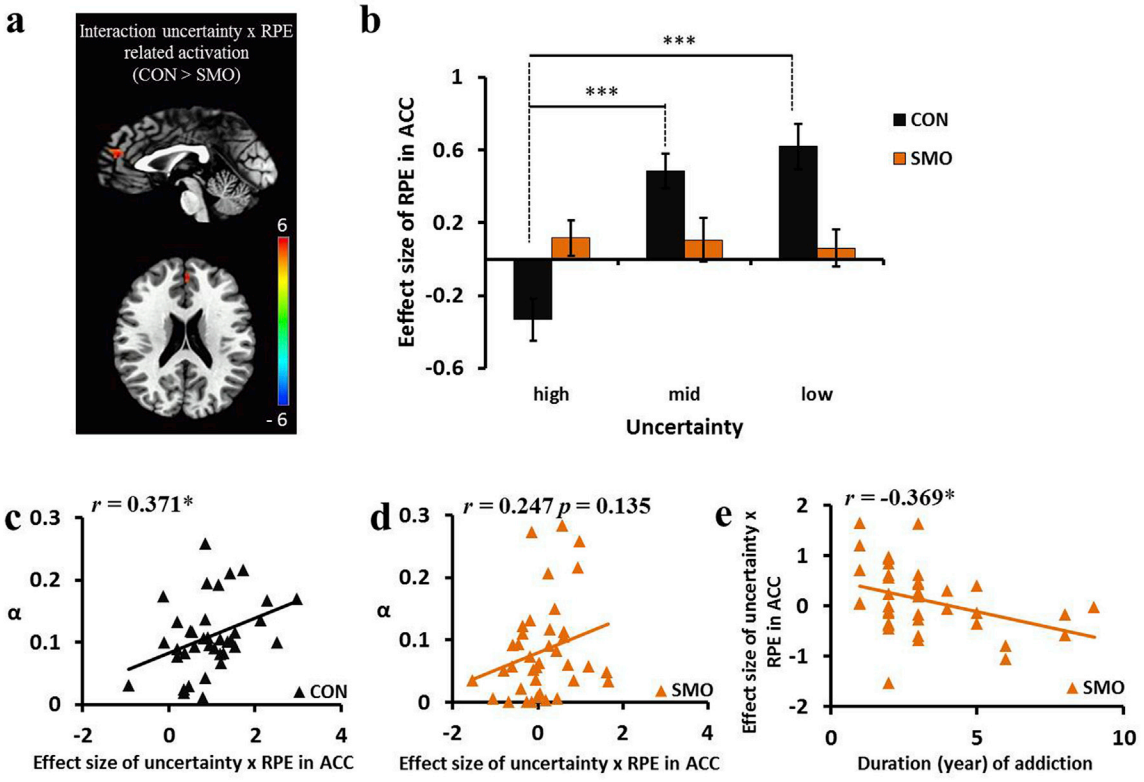

Fig. 3. The anterior cingulate cortex/medial prefrontal cortex (ACC/MPFC) did not modulate reward prediction errors (RPEs) learning in smokers. (a) Smokers (SMO) experienced weaker ACC/MPFC activation related to uncertainty × RPE compared to non-smoking controls (CON). Maps were thresholded at $p < 0.005$, corrected for multiple comparisons using cluster correction, voxel-level $p < 0.0001$. (b–d) Activation in the ACC/MPFC related to uncertainty × RPE was associated with learning rates in all participants and non-smokers, but not in smokers. (e) Activation in the ACC related to uncertainty × RPE was associated with duration of nicotine use in smokers. *$p < 0.05$, **$p < 0.01$, ***$p < 0.001$. For b, plotted data represent mean ± s. e.m. across participants.

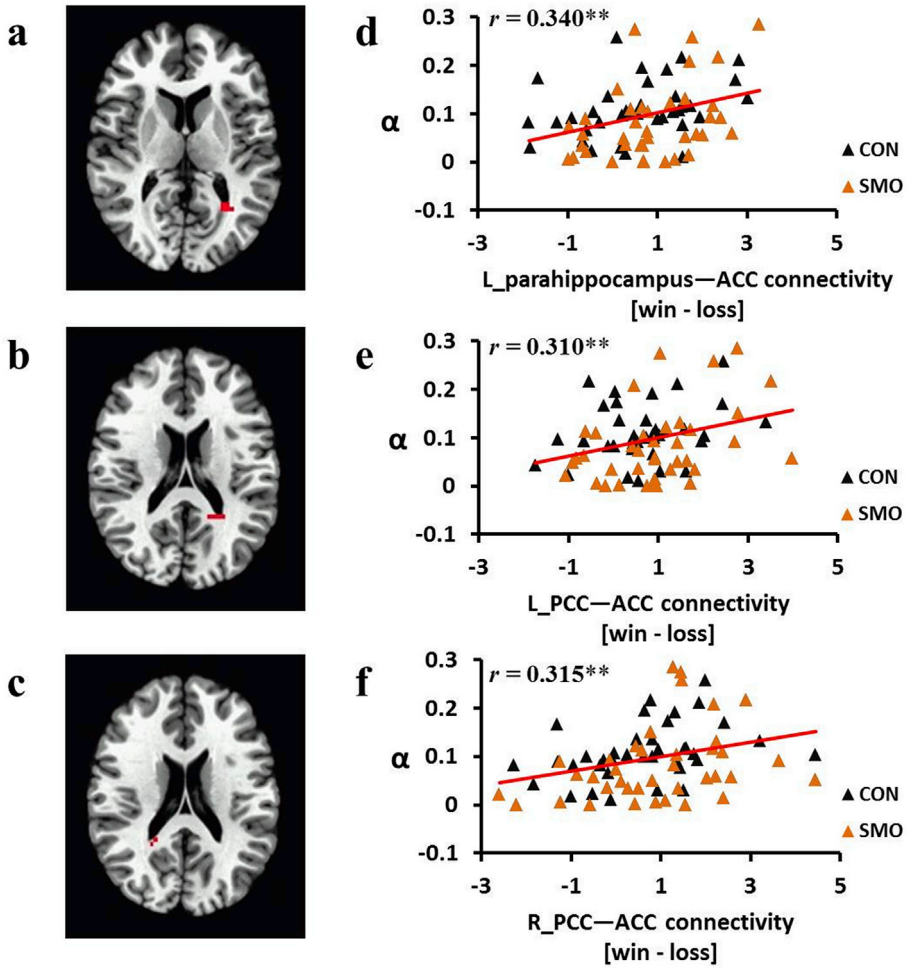

Fig. 4. Relationship between learning rates and connectivity of the left parahippocampus/left posterior cingulate cortex (PCC)/right PCC-anterior cingulate cortex/medial prefrontal cortex (ACC/MPFC). (a–c) Whole brain connectivity analysis indicated significant functional connectivity between the ACC/MPFC and the left parahippocampus, left PCC, and right PCC during earn > loss outcomes. (d–f) Learning rates were significantly correlated with functional connectivity between the left parahippocampus-ACC/MPFC, the left PCC-ACC/MPFC, and the right PCC-ACC/MPFC. $^{**}p < 0.01$. CON = non-smoking controls; SMO = smokers.

Experiment II

RGT

Consistent with previous results(Rivalan et al., 2011), choice patterns changed over time in the control group ($F_{5,78} = 5.19$, $p < 0.001$). As shown in Fig. 6a, the rats initially selected advantageous options and disadvantageous options equally, before developing a slight tendency for the more advantageous options. The nicotine group had smaller total earnings compared to the control group (nicotine: 4.83 ± 1.09; control: 6.24 ± 1.37; $F_{1,25} = 8.612$, $p = 0.007$). A two (group) × six (block) repeated-measures ANOVA revealed significant main effects of group ($F_{1,160} = 53.649, p < 0.001$) and block ($F_{5,160} = 43.264, p < 0.001$), however there was no significant interaction ($F_{5,160} = 0.970$, $p = 0.210$). A post hoc analysis demonstrated that the nicotine group performed significantly worse than the control group during five blocks (block_2$_{time:10–20min}$: $t_{25} = 2.558$, $p = 0.017$; block_3$_{time:20–30min}$: $t_{25} = 2.567$, $p = 0.017$; block_4$_{time:30–40min}$: $t_{25} = 3.109$, $p = 0.005$; block_5$_{time:40–50min}$: $t_{25} = 4.056$, $p < 0.001$; block_6$_{time:50–60min}$: $t_{25} = 5.085$, $p < 0.001$); however, this effect did not occur during one block (block_1$_{time:0–10min}$: $t_{25} = 1.471$, $p = 0.157$). There were no weight differences between the nicotine and control groups (nicotine: 414.31 ± 23.3; control: 422.07 ± 27.83; $F_{1,25} = 0.612, p = 0.441$) on the testing day. There was also no effect of motivation to retrieve rewards, as indicated by the mean latency to collect water during the test (nicotine: 2.19 ± 0.53; control: 2.16 ± 0.39; $F_{1,25} = 0.032$, $p = 0.856$). Furthermore, the nicotine group had lower learning rates of RPEs compared to the control group ($t_{25} = 3.968, p = 0.001$; Fig. 6b). We found that the learning rates were significantly correlated with total earnings of each rat ($r = 0.586$, $p = 0.01$).

qRT-PCR

Gene expressions of the 5-HT_{1A} and 5-HT_{2A} receptors decreased significantly in the nicotine groups compared to the control groups (5-HT_{1A}-R: $t_{25} = 2.419$, $p = 0.023$; 5-HT_{2A}-R: $t_{25} = 2.558$, $p = 0.017$; Fig. 6c–d). 5-HT_{2C} receptors did not significantly change ($t_{25} = -1.070$, $p = 0.295$). In contrast, expressions of the DA_1 and DA_2 receptors did not significantly change in the nicotine groups compared to the control groups (DA_1-R: $t_{25} = 0.736$, $p = 0.469$; DA_2-R: $t_{25} = 0.945$, $p = 0.354$; Fig. 6e–f). Further, we examined a potential correlation between learning rates and 5-HT mRNA expression. We found that the learning rates were significantly correlated with 5-HT receptor mRNA expression (5-HT_{1A}-R: $r = 0.406, p = 0.036$; 5-HT_{2A}-R: $r = 0.462, p = 0.015$; Fig. 6g–h) and were not significantly correlated with DA receptor mRNA expression (DA_1-R: $r = 0.056$, $p = 0.783$; DA_2-R: $r = 0.156$, $p = 0.437$).

Discussion

We found that both human smokers and rats with chronic nicotine exposure had poorer learning rates of RPEs and lower total net scores on the IGT, compared to control groups without nicotine exposure.

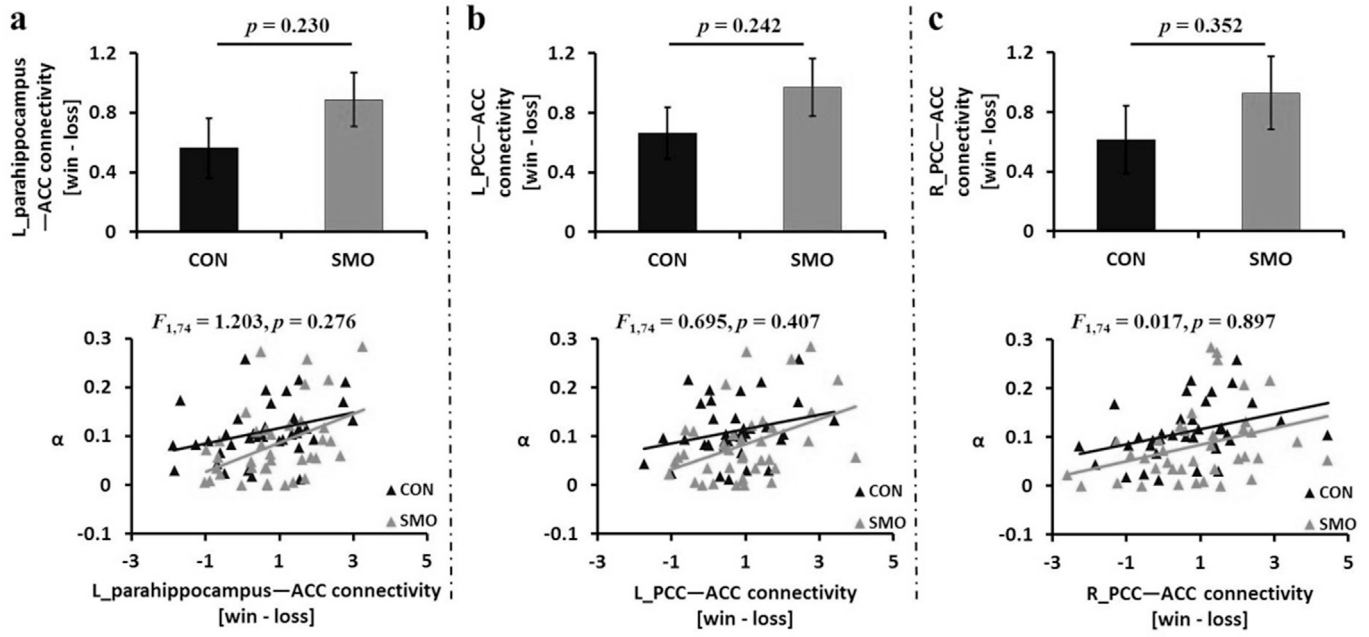

Fig. 5. (a) No group differences were observed in functional connectivity of the anterior cingulate cortex/medial prefrontal cortex (ACC/MPFC) and the left parahippocampus during earn > loss outcomes, or for the relationship between functional connectivity and learning rates. (b) No group differences were observed in functional connectivity of the ACC/MPFC and the left posterior cingulate cortex (PCC), or for the relationship between functional connectivity and learning rates. (c) No group differences were observed in functional connectivity of the ACC/MPFC and the right PCC, or for the relationship between functional connectivity and learning rates. CON = non-smoking controls; SMO = smokers. Plotted data represent mean ± s.e.m. across participants.

Furthermore, learning rates were significantly correlated with total net scores, and the ACC did not modulate RPEs according to various levels of uncertainty in smokers. Finally, we found that 5-HT receptor mRNA expression in the ACC decreased in rats exposed to nicotine group, compared to the control group, and learning rates were correlated with 5-HT receptor mRNA expression in the ACC.

A recent modeling study hypothesized that abnormal reinforcement learning may contribute to drug addiction; drug abuse is hypothesized to distort prediction error signals independently from true error rates (Redish, 2004). This model is supported by numerous studies that have demonstrated weak computation of RPEs during drug addiction (Chiu et al., 2008; Park et al., 2010; Rose et al., 2012, 2014; Carey et al., 2015; Parvaz et al., 2015). It is well-known that a negative mental state can worsen addiction (Kessler et al., 2003). This mental state is often caused by unpredictable and stressful events such as marital breakup, death of a loved one, and unemployment, which can increase environmental uncertainty of life. Generally, uncertainty exacerbates addiction (Zhang, 2014). Environmental uncertainty is considered as a potential vulnerability to addictive drugs (Schultz, 2011). From this perspective, drug-induced deficient uncertainty signals contribute to compulsive drug use via two separate mechanisms. First, enhanced uncertainty signal would boost the effects of positive prediction errors and blunt the effects of negative prediction errors, thereby promoting the associability of learning rule. Second, enhanced uncertainty signals may make it difficult to distinguish risk and value signals via orbitofrontal neurons and thus produce an overly optimistic value signal. The present study supports the first type of contribution. In our present study, we divided the uncertainty regarding the task into three levels according to the magnitude of uncertainty, then analyzed how the uncertainty modulates RPE processing. Non-smokers processed RPE according to the level of uncertainty (section III of Supplementary Information). Relative to non-smokers, the ACC did not use uncertainty to modulate RPE processes in smokers. We also showed that longer durations of nicotine use are associated with weaker uncertainty modulation in the ACC. Deficits in uncertainty modulation may result in abnormal weight given to RPEs, thereby contributing to inaccurate action value updating and pathological decision-making.

In the present study, we showed that variations in learning rates reflect uncertainty modulation in the ACC, which is consistent with previous research (Behrens et al., 2007). Thus, a potential neuromodulator involved in the processes of uncertainty modulation on RPEs in the ACC would be expected to be associated with learning rates. A pharmacological fMRI study on the 5-HT system indicated that 5-HT may contribute to reinforcement learning by directly modulating brain responses to stimuli (Tanaka et al., 2004, 2007). Additionally, 5-HT transporter expression affects information processing biases for uncertain cues in mice (McHugh et al., 2015). Experiment II of the present study found that 5-HT_{1A} and 5-HT_{2A} receptor gene expression in the ACC were associated with learning rates. Thus, we suggest that 5-HT is a neuromodulator of uncertainty on RPEs in the ACC.

Given the relationship between 5-HT and uncertainty modulation on RPEs in the ACC, decreases in 5-HT are expected to capture the impaired uncertainty modulation function during drug addiction. Previous studies have reported that deficits in 5-HT, especially in the MPFC, disturb adaptation to changes in the required action for a given cue by making participants more likely to maintain prelearned behaviors (Clarke et al., 2004). In contrast, antidepressants acting on 5-HT facilitate prediction error signaling (Graf et al., 2016). In Experiment II, 5-HT_{1A} and 5-HT_{2A} receptor gene expressions in the ACC, which were associated with learning rates of RPEs, were significantly decreased in the nicotine groups compared to the control groups. Therefore, the present study suggests that decreased 5-HT receptor mRNA expression in the ACC may be the biochemical and molecular mechanism of impairments in uncertainty modulation in chronic nicotine users.

RPEs are generated by phasic DA signaling in the midbrain, which projects to other areas within the mesostriatocortical network (Schultz, 1998; McClure et al., 2003; Liu et al., 2007; Cohen et al., 2012). However, the present study found no differences in mRNA expression for DA_1 or DA_2 receptors between rats with chronic nicotine exposure and control rats. These findings suggest that chronic nicotine exposure may not

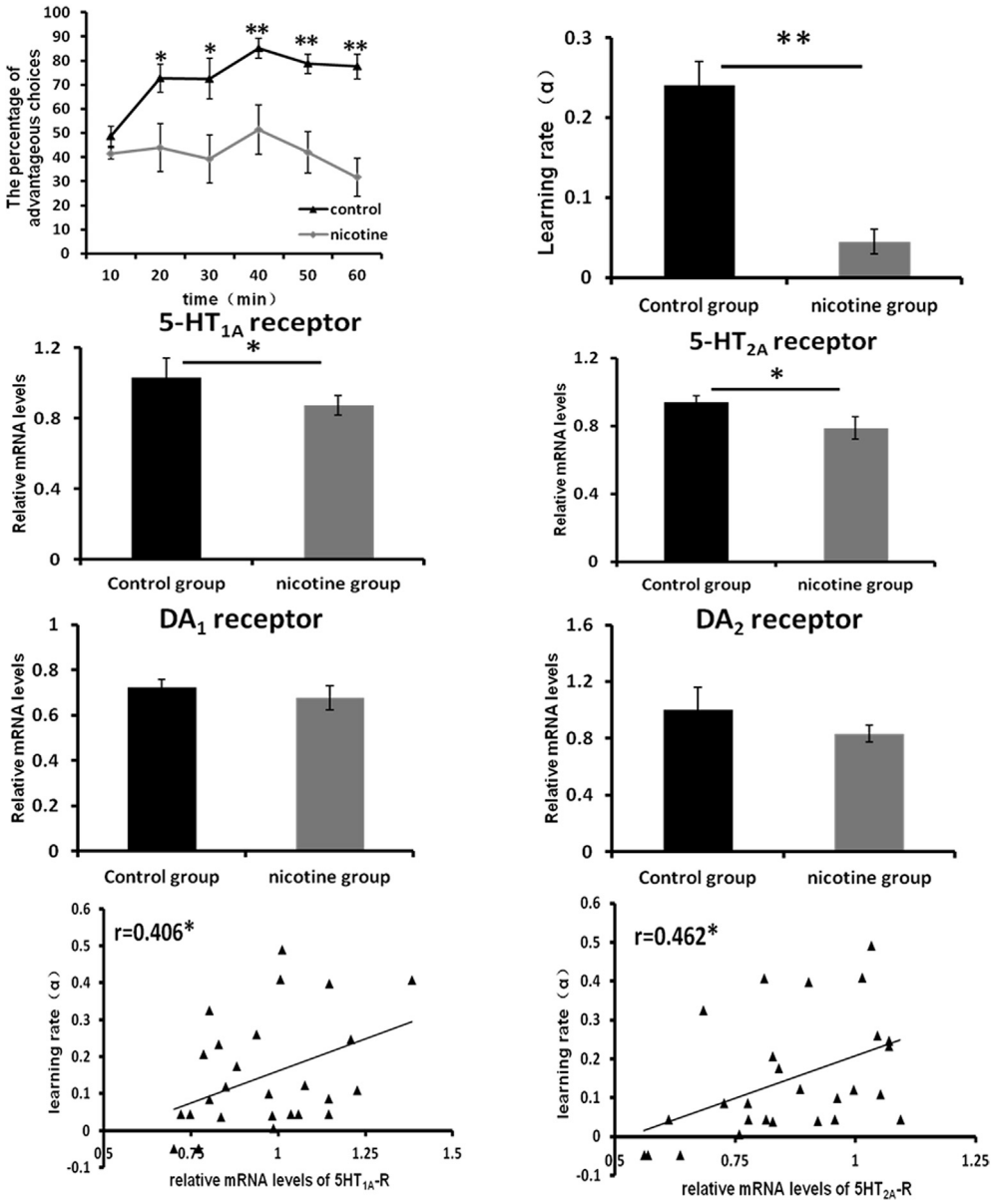

Fig. 6. Behavioral and quantitative real-time polymerize chain reaction results in rats. (a) The control group systematically preferred advantageous options as the experiment progressed, whereas the nicotine group did not develop this preference. (b) The nicotine group had lower learning rates of reward prediction errors (RPEs) compared to the control group. (c–f) Gene expression of serotonin (5-HT) receptor subtypes decreased significantly in the nicotine groups compared to the control groups; similar effects did not occur for dopamine (DA) receptor subtypes. (g–h) Genetic expression of the 5-HT_{1A} and 5-HT_{2A} receptor subtype was significantly correlated with learning rates of RPEs in all rats. Control group: n = 14; Nicotine group: n = 13. *$p < 0.05$, **$p < 0.01$. Plotted data represent mean ± s. e.m. across subjects.

influence DA functioning in the ACC by altering mRNA expression, and an alternate mechanism likely exists. In contrast, 5-HT has an influence opposite to DA in reinforcement processes (Daw et al., 2002), and can suppress DA activity in the ventral tegmental area and substantia nigra, as well as decrease DA release in the nucleus accumbens and striatum (Kapur and Remington, 1996; Daw et al., 2002). Nicotine decreases 5-HT receptor expression in the cortex by altering 5-HT neurotransmission (Kenny et al., 2001). DA prediction error signaling in the ACC may be enhanced in nicotine users by relieving 5-HT depression, while enabling DA_1 and DA_2 receptor expression to be unaltered in the ACC. DA_2 receptor expression appeared to be increased in the nicotine group in Fig. 6f although it is not significant. If the number of rats is increased in a future study, a significant difference may be found.

We found that the association between ACC-related functional connectivity and learning rates was not affected in smokers. Learning modulation is a complicated cognitive process involving a network that includes the bilateral PCC and left parahippocampus centered, which have critical connections with the ACC (see section V of Supplementary Information). Although PPIs and RSFC do not provide any information regarding directionality, our results suggest that RPE signal propagation within this network is accurate in smokers; however, RPE signals are abnormally caused by impaired uncertainty modulation in the ACC before propagating to other regions. Additionally, impairments in the ACC were reflected only in local functioning, and not in local structure (see VBM data in section IV of Supplementary Information). Thus, the ACC may be a novel target to treat maladaptive decision-making in chronic nicotine users by improving ACC functioning with techniques such as transcranial direct stimulation or real-time fMRI neurofeedback.

In the present study, no differences were observed in the net score during blocks 5 and 6. In previous studies (Fernie and Tunney, 2006; Harman, 2011; Xiao et al., 2011), the same phenomenon was shown. One explanation is based on the theory of ‘exploration-exploitation’(Daw

et al., 2006). When learning a certain rule to some degree, subjects shift from 'exploitation' to 'exploration'. Thus the net score increased slowly or even decreased slightly in latter blocks. Meanwhile, the net score of smokers increased steadily because they had not acquired the rule yet, resulting in no observed differences in the net score during blocks 5 and 6.

In our study, we found that disrupted uncertainty modulation in the ACC contributes to deficits in RPE processes and adaptive decision-making in chronic nicotine users. Our results suggested that the deficits were correlated with chronic nicotine exposure. However, adequate proof is required in future to demonstrate the causality between the deficits and drug addiction.

A potential limitation of the present study was that we did not record electrophysiological activity of 5-HT neurons in the ACC during the RGT, which would have enabled us to compare activity between different levels of uncertainty. We report indirect evidence that decreased 5-HT receptor mRNA expression is the biochemical mechanism of uncertainty modulation impairments in chronic nicotine users. This evidence includes the association between learning rates and the effect size of the interaction between uncertainty x RPE in the ACC of smokers, and the association between learning rates and 5-HT receptor mRNA expression in rats with chronic nicotine exposure. However, direct evidence is required and will be obtained in the future by recording electrophysiological activity. The second potential limitation was the differences in types of punishments in human and rat tasks. Human punishments consisted of actual losses, whereas rats received timeouts. Given that the reward (water) is instantly consumed in the RGT, rather than accumulated over time and then consumed, it is impossible to withdraw the reward from the animals once they have been won, i.e., for truly reproducing the sensation of loss. In order to model the concept of loss in the RGT, punishment is delivered as delays (timeouts). However, Experiment III (a similar task to the one used in rats, section I of Supplementary Information) results show that the processes in human and rat tasks may be similar. The third potential limitation was the combination of long-term effects and acute withdrawal effects. Both, smokers and rats with chronic nicotine exposure were in the withdrawal state, which may have influenced the results. However, our results showed that the effect size in ACC/MPFC was associated with duration of nicotine use, supporting the long-term effects of smoking. Stricter experimental control should be employed in further to separate these two kinds of effects. The fourth potential limitation was that we could not rule out the effect of craving. In our study, all smokers were required smoking cessation 2 h before the experiment in order to obtain the same craving baseline. The fifth potential limitation was that water deficient state resulted from water deprivation may be somewhat impact on rat's behavioral choice. Water deprivation is used to increase the rats' motivation to work for water delivery. When rats undergone mild water deprivation in our preliminary experiments, they needed very long periods of training, even no accomplished the training task. Besides, there was no effect of food deprivation on exploratory behavior(Pierre et al., 2001). Furthermore, the previous study showed that the employment of different food deprivation levels has been found to have no impact on decision-making(Rivalan et al., 2009).

In summary, our findings provide evidence that disrupted uncertainty modulation in the ACC, which are related to decreased 5-HT receptor mRNA expression in the ACC, contribute to deficits in RPE processes and adaptive decision-making in chronic nicotine users. Given that current therapeutic interventions focus on reinforcement modification, our neural findings support interventions for pathological decision-making during drug addiction that focus on the 5-HT system in the ACC. Our observations also extend the scope of neurobiological deficits underlying reinforcement learning during drug addiction.

Funding

This work was supported by grants from the National Natural Science Foundation of China (31230032, 31471071, 31771221, 31500917) and the Fundamental Research Funds for the Central Universities of China, and The National Key Basic Research Program (2016YFA0400900), and China Postdoctoral Science Foundation (2016M592051), and MOE-Mircrosoft Key Laboratory of USTC. The numerical calculations in this study were partly performed with the supercomputing system at the Supercomputing Centre of USTC.

Acknowledgements

Conflict of interest

The authors declare no conflicts of interest.

Authors' contributions

ZDW, LH, YW and XCZ conceived and designed the study. ZDW and LH obtained the findings. ZDW, LH, YL and YW were responsible for acquisition of data. ZDW, LH analyzed and interpreted the data. XYZ, RZ, LZY, JJB, HWS, WW, YFZ, PG provided administrative, technical, or material support. XCZ supervised the study. ZDW, LH, YFZ, and XCZ drafted the paper and all authors contributed to critical revision for intellectual content.

Appendix A. Supplementary data

Supplementary data related to this article can be found at https://doi.org/10.1016/j.neuroimage.2017.11.048.

References

American Psychiatric Association, 2013. Diagnostic and Statistical Manual of Mental Disorders (DSM V). American Psychiatric Association, Washington DC.
Bach, D.R., Hulme, O., Penny, W.D., Dolan, R.J., 2011. The known unknowns: neural representation of second-order uncertainty, and ambiguity. J. Neurosci. Official J. Soc. Neurosci. 31, 4811–4820.
Balasubramani, P.P., Chakravarthy, V.S., Ravindran, B., Moustafa, A.A., 2015. A network model of basal ganglia for understanding the roles of dopamine and serotonin in reward-punishment-risk based decision making. Front. Comput. Neurosci. 9, 76.
Bechara, A., Damasio, A.R., Damasio, H., Anderson, S.W., 1994. Insensitivity to future consequences following damage to human prefrontal cortex. Cognition 50, 7–15.
Behrens, T.E., Woolrich, M.W., Walton, M.E., Rushworth, M.F., 2007. Learning the value of information in an uncertain world. Nat. Neurosci. 10, 1214–1221.
Birn, R.M., Diamond, J.B., Smith, M.A., Bandettini, P.A., 2006. Separating respiratory-variation-related fluctuations from neuronal-activity-related fluctuations in fMRI. NeuroImage 31, 1536–1548.
Carey, S.E., Nestor, L., Jones, J., Garavan, H., Hester, R., 2015. Impaired learning from errors in cannabis users: dorsal anterior cingulate cortex and hippocampus hypoactivity. Drug Alcohol Dependence 155, 175–182.
Chiu, P.H., Lohrenz, T.M., Montague, P.R., 2008. Smokers' brains compute, but ignore, a fictive error signal in a sequential investment task. Nat. Neurosci. 11, 514–520.
Clarke, H.F., Dalley, J.W., Crofts, H.S., Robbins, T.W., Roberts, A.C., 2004. Cognitive inflexibility after prefrontal serotonin depletion. Science 304, 878–880.
Cohen, J.Y., Haesler, S., Vong, L., Lowell, B.B., Uchida, N., 2012. Neuron-type-specific signals for reward and punishment in the ventral tegmental area. Nature 482, 85–88.
Cools, R., Nakamura, K., Daw, N.D., 2011. Serotonin and dopamine: unifying affective, activational, and decision functions. Neuropsychopharmacol. Official Publ. Am. Coll. Neuropsychopharmacol. 36, 98–113.
Cox, R.W., 1996. AFNI: software for analysis and visualization of functional magnetic resonance neuroimages. Comput. Biomed. Res. 29 (3), 162–173.
Daw, N.D., Kakade, S., Dayan, P., 2002. Opponent interactions between serotonin and dopamine. Neural Netw. Official J. Int. Neural Netw. Soc. 15, 603–616.
Daw, N.D., O'Doherty, J.P., Dayan, P., Seymour, B., Dolan, R.J., 2006. Cortical substrates for exploratory decisions in humans. Nature 441, 876–879.
Fernie, G., Tunney, R.J., 2006. Some decks are better than others: the effect of reinforcer type and task instructions on learning in the Iowa Gambling Task. Brain Cogn. 60, 94–102.
Friston, K., 2005. A theory of cortical responses. Philos. Trans. R. Soc. Lond. B Biol. Sci. 360, 815–836.
Friston, K.J., Buechel, C., Fink, G.R., Morris, J., Rolls, E., Dolan, R.J., 1997. Psychophysiological and modulatory interactions in neuroimaging. NeuroImage 6, 218–229.
Graf, H., Metzger, C.D., Walter, M., Abler, B., 2016. Serotonergic antidepressants decrease hedonic signals but leave learning signals in the nucleus accumbens unaffected. Neuroreport 27, 18–22.
Harman, J.L., 2011. Individual differences in need for cognition and decision making in the Iowa Gambling Task. Personality Individ. Differ. 51, 112–116.

Heatherton, T.F., Kozlowski, L.T., Frecker, R.C., Fagerstrom, K.O., 1991. The fagerstrom test for nicotine dependence: a revision of the fagerstrom tolerance questionnaire. Br. J. Addict. 86, 1119–1127.
Kapur, S., Remington, G., 1996. Serotonin-dopamine interaction and its relevance to schizophrenia. Am. J. Psychiatry 153, 466–476.
Kelsey, J.E., Niraula, A., 2013. Effects of acute and sub-chronic nicotine on impulsive choice in rats in a probabilistic delay-discounting task. Psychopharmacology 227, 385–392.
Kennerley, S.W., Walton, M.E., Behrens, T.E., Buckley, M.J., Rushworth, M.F., 2006. Optimal decision making and the anterior cingulate cortex. Nat. Neurosci. 9, 940–947.
Kenny, P.J., File, S.E., Rattray, M., 2001. Nicotine regulates 5-HT(1A) receptor gene expression in the cerebral cortex and dorsal hippocampus. Eur. J. Neurosci. 13, 1267–1271.
Kessler, R.C., Berglund, P., Demler, O., Jin, R., Koretz, D., Merikangas, K.R., Rush, A.J., Walters, E.E., Wang, P.S., 2003. The epidemiology of major depressive disorder: results from the National Comorbidity Survey Replication (NCS-R). JAMA 289, 3095–3105.
Linnet, J., Mouridsen, K., Peterson, E., Moller, A., Doudet, D.J., Gjedde, A., 2012. Striatal dopamine release codes uncertainty in pathological gambling. Psychiatry Res. 204, 55–60.
Liu, X., Powell, D.K., Wang, H., Gold, B.T., Corbly, C.R., Joseph, J.E., 2007. Functional dissociation in frontal and striatal areas for processing of positive and negative reward information. J. Neurosci. Official J. Soc. Neurosci. 27, 4587–4597.
McClure, S.M., Berns, G.S., Montague, P.R., 2003. Temporal prediction errors in a passive learning task activate human striatum. Neuron 38, 339–346.
McHugh, S.B., Barkus, C., Lima, J., Glover, L.R., Sharp, T., Bannerman, D.M., 2015. SERT and uncertainty: serotonin transporter expression influences information processing biases for ambiguous aversive cues in mice. Genes Brain Behav. 14, 330–336.
Mikhael, J.G., Bogacz, R., 2016. Learning reward uncertainty in the basal ganglia. PLoS Comput. Biol. 12, e1005062.
Muller, C.P., Homberg, J.R., 2015. The role of serotonin in drug use and addiction. Behav. Brain Res. 277, 146–192.
Murphy, S.E., Longhitano, C., Ayres, R.E., Cowen, P.J., Harmer, C.J., Rogers, R.D., 2009. The role of serotonin in nonnormative risky choice: the effects of tryptophan supplements on the “reflection effect” in healthy adult volunteers. J. Cogn. Neurosci. 21, 1709–1719.
Niv, Y., Duff, M.O., Dayan, P., 2005. Dopamine, uncertainty and TD learning. Behav. Brain Funct. BBF 1, 6.
Park, S.Q., Kahnt, T., Beck, A., Cohen, M.X., Dolan, R.J., Wrase, J., Heinz, A., 2010. Prefrontal cortex fails to learn from reward prediction errors in alcohol dependence. J. Neurosci. 30, 7749–7753.
Parvaz, M.A., Konova, A.B., Proudfit, G.H., Dunning, J.P., Malaker, P., Moeller, S.J., Maloney, T., Alia-Klein, N., Goldstein, R.Z., 2015. Impaired neural response to negative prediction errors in cocaine addiction. J. Neurosci. Official J. Soc. Neurosci. 35, 1872–1879.
Paxinos, G., Watsons, C., 2007. The Rat Brain in Stereotaxic Coordinates, sixth ed. Elsevier Press, New York.
Pierre, P.J., Skjoldager, P., Bennett, A.J., Renner, M.J., 2001. A behavioral characterization of the effects of food deprivation on food and nonfood object interaction: an investigation of the information-gathering functions of exploratory behavior. Physiology Behav. 72, 189–197.
Preuschoff, K., Bossaerts, P., 2007. Adding prediction risk to the theory of reward learning. Ann. N. Y. Acad. Sci. 1104, 135–146.
Redish, A.D., 2004. Addiction as a computational process gone awry. Science 306, 1944–1947.
Rivalan, M., Ahmed, S.H., Dellu-Hagedorn, F., 2009. Risk-prone individuals prefer the wrong options on a rat version of the Iowa Gambling Task. Biol. Psychiatry 66, 743–749.
Rivalan, M., Coutureau, E., Fitoussi, A., Dellu-Hagedorn, F., 2011. Inter-individual decision-making differences in the effects of cingulate, orbitofrontal, and prelimbic cortex lesions in a rat gambling task. Front. Behav. Neurosci. 5, 22.
Rogers, R.D., 2011. The roles of dopamine and serotonin in decision making: evidence from pharmacological experiments in humans. Neuropsychopharmacology 36, 114–132.
Rose, E.J., Ross, T.J., Salmeron, B.J., Lee, M., Shakleya, D.M., Huestis, M., Stein, E.A., 2012. Chronic exposure to nicotine is associated with reduced reward-related activity in the striatum but not the midbrain. Biol. Psychiatry 71, 206–213.
Rose, E.J., Salmeron, B.J., Ross, T.J., Waltz, J., Schweitzer, J.B., McClure, S.M., Stein, E.A., 2014. Temporal difference error prediction signal dysregulation in cocaine dependence. Neuropsychopharmacol. Official Publ. Am. Coll. Neuropsychopharmacol. 39, 1732–1742.
Schultz, W., 1998. Predictive reward signal of dopamine neurons. J. Neurophysiology 80, 1–27.
Schultz, W., 2011. Potential vulnerabilities of neuronal reward, risk, and decision mechanisms to addictive drugs. Neuron 69, 603–617.
Stoltenberg, S.F., Vandever, J.M., 2010. Gender moderates the association between 5-HTTLPR and decision-making under ambiguity but not under risk. Neuropharmacology 58, 423–428.
Tanabe, J., Reynolds, J., Krmpotich, T., Claus, E., Thompson, L.L., Du, Y.P., Banich, M.T., 2013. Reduced neural tracking of prediction error in substance-dependent individuals. Am. J. Psychiatry 170, 1356–1363.
Tanaka, S.C., Doya, K., Okada, G., Ueda, K., Okamoto, Y., Yamawaki, S., 2004. Prediction of immediate and future rewards differentially recruits cortico-basal ganglia loops. Nat. Neurosci. 7, 887–893.
Tanaka, S.C., Schweighofer, N., Asahi, S., Shishida, K., Okamoto, Y., Yamawaki, S., Doya, K., 2007. Serotonin differentially regulates short- and long-term prediction of rewards in the ventral and dorsal striatum. PLoS One 2, e1333.
Tobler, P.N., Fiorillo, C.D., Schultz, W., 2005. Adaptive coding of reward value by dopamine neurons. Science 307, 1642–1645.
Waheed, S.H., Mirbagheri, S., Agarwal, S., Kamali, A., Yahyavi-Firouz-Abadi, N., Chaudhry, A., DiGianvittorio, M., Gujar, S.K., Pillai, J.J., Sair, H.I., 2016. Reporting of resting-state functional magnetic resonance imaging preprocessing methodologies. Brain Connect. 6, 663–668.
Wang, Y., Ma, N., He, X., Li, N., Wei, Z., Yang, L., Zha, R., Han, L., Li, X., Zhang, D., Liu, Y., Zhang, X., 2017. Neural substrates of updating the prediction through prediction error during decision making. NeuroImage 157, 1–12.
Xiao, L., Bechara, A., Palmer, P.H., Trinidad, D.R., Wei, Y., Jia, Y., Johnson, C.A., 2011. Parent-child engagement in decision making and the development of adolescent affective decision capacity and binge drinking. Pers. Individ. Dif. 51, 285–292.
Zhang, F., 2006. SuperState: a computer program for the control of operant behavioral experimentation. J. Neurosci. Methods 155, 194–201.
Zhang, Z.Y.a.R., 2014. Rational addictive behavior under uncertainty. Yorken 14/12.