

Reward Prediction Error Signal Enhanced by Striatum–Amygdala Interaction Explains the Acceleration of Probabilistic Reward Learning by Emotion

Noriya Watanabe,^{1,2,3,4} Masamichi Sakagami,² and Masahiko Haruno^{1,4,5}

¹Center for Information and Neural Networks, National Institute of Information and Communications Technology, Suita, Osaka 565-0871, Japan,

²Tamagawa University Brain Science Institute, Machida, Tokyo 194-8610, Japan, ³Japan Society for Promotion of Science, and ⁴Graduate School of Frontier Biosciences, Osaka University, Osaka, Suita, Osaka 565-0871, Japan, and ⁵PRESTO, Japan Science and Technology Agency, Kawaguchi, Saitama 332-0012, Japan

Learning does not only depend on rationality, because real-life learning cannot be isolated from emotion or social factors. Therefore, it is intriguing to determine how emotion changes learning, and to identify which neural substrates underlie this interaction. Here, we show that the task-independent presentation of an emotional face before a reward-predicting cue increases the speed of cue–reward association learning in human subjects compared with trials in which a neutral face is presented. This phenomenon was attributable to an increase in the learning rate, which regulates reward prediction errors. Parallel to these behavioral findings, functional magnetic resonance imaging demonstrated that presentation of an emotional face enhanced reward prediction error (RPE) signal in the ventral striatum. In addition, we also found a functional link between this enhanced RPE signal and increased activity in the amygdala following presentation of an emotional face. Thus, this study revealed an acceleration of cue–reward association learning by emotion, and underscored a role of striatum–amygdala interactions in the modulation of the reward prediction errors by emotion.

Introduction

Students often notice that learning is affected by feelings of tension from teachers or in the classroom. Learning is not always rational, and is often affected by emotions and other social factors. Such observations pose an intriguing, but mostly unstudied, question about the mechanisms by which emotions and social factors affect learning.

Probabilistic reward learning has been studied intensively in humans and animals as a basic model of learning (Dayan, 2008), and can be understood within the framework of reinforcement learning models (Sutton and Barto, 1998), in which an agent aims to maximize their “reward.” The agent learns the “value,” which is the expected reward in a given context, by minimizing the “reward prediction error” (RPE), which is the difference between the expected reward value and the actual reward. The firing rates of midbrain dopamine neurons correlate with reward values and RPE signals (Schultz et al., 1997), and several human studies have identified the involvement of the ventral and dorsal striatum,

which both receive massive dopamine projections (McClure et al., 2003; O’Doherty et al., 2003; Haruno and Kawato, 2006).

In the present study, we conducted behavioral and functional magnetic resonance imaging (fMRI) experiments using a probabilistic reward learning task, and examined whether an emotional stimulus changes a previously established learning process. More specifically, we tested whether the presentation of a task-independent emotional face before a reward-predicting cue would change learning and reward value and RPE representations in the brain.

Materials and Methods

The participants in this study were undergraduate students who did not declare any history of psychiatric or neurological disorders. This study was approved by the ethics committee of the National Institute of Information and Communications Technology and Tamagawa University. All participants provided written informed consent before any experiments.

Experimental design

Behavioral task. Twenty participants (10 male, 10 female; mean age 20.0 ± 0.7 years) learned probabilistic associations between visual cues and rewards (65% or 35%) by trial and error (Fig. 1A). A task-independent fearful or neutral face [two expressions \times 20 stimuli taken from NimStim (<http://www.macbrain.org/>)] was presented on a screen (1000 ms) in a pseudo-random order before the visual cue (250 ms). Then, the participant pressed a ¥100 (large) or ¥1 (small) button within 1750 ms, depending on their expectation of which reward was more likely to follow the cue. The order of cue presentation and the position of the two buttons (left or right) were randomized across trials. A red dot was displayed under the chosen option, and the reward outcome was shown within a blue square. Although this task required the participants

Received July 17, 2012; revised Dec. 14, 2012; accepted Jan. 7, 2013.

Author contributions: M.H. designed research; N.W. performed research; N.W. and M.H. analyzed data; N.W., M.S., and M.H. wrote the paper.

This study was supported by PRESTO from the Japan Science and Technology Agency, and KAKENHI (Grant 22300139) to M.H. We are grateful to Drs. Peter Dyan and Okihide Hikosaka for helpful discussion, and Dr. Haji for technical assistance with fMRI.

The authors declare no competing financial interests.

Correspondence should be addressed to Masahiko Haruno, Nanobiology Building 7F, Osaka University 1-3 Yamadaoka, Suita, Osaka 565-0871, Japan. E-mail: mharuno@nict.go.jp.

DOI:10.1523/JNEUROSCI.3400-12.2013

Copyright © 2013 the authors 0270-6474/13/334487-07\$15.00/0

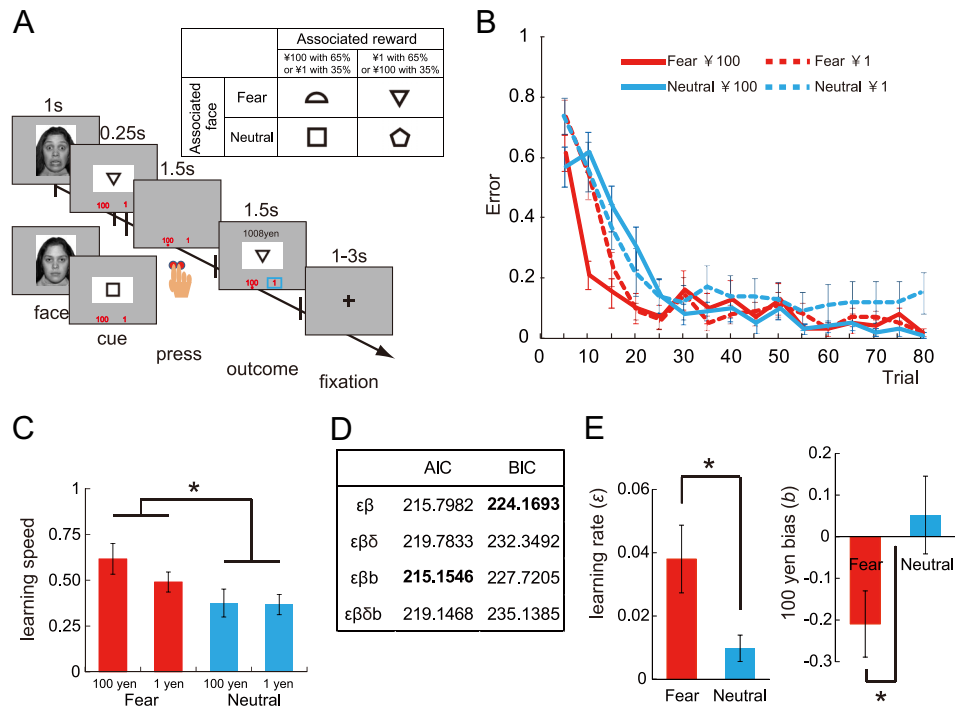


Figure 1. Experimental design and behavioral results. **A**, Inset, An example combination of the facial expression, cue, and reward. Each of the four cues was associated probabilistically (65%) with one of the two different reward amounts, and also with one of the two facial expressions. **A**, The participants were required to press a button, depending on their expectation regarding which reward would be associated with the presented cue. **B**, Learning curves. Each data point represents the average of five trials, and the error bar indicates the SEM. **C**, Exponential fit of learning speed (mean \pm SEM). **D**, The results of the parameter estimation by AIC and BIC. ϵ , β , δ , and b represent the learning rate, exploration, reward sensitivity, and ¥100 bias, respectively. **E**, Model-based analysis of the learning rate and the ¥100 bias (mean \pm SEM by $\epsilon\beta b$). * $p < 0.05$, ** $p < 0.01$, and *** $p < 0.005$ throughout the figures.

to predict the magnitude of the reward, the actual reward was unaffected by their choice, which is a form of classical conditioning. Before the experiments, we confirmed that the participants fully understood the task characteristics. They were instructed that the face presentation would signal the immediate appearance of a cue, and that there was no fixed relationship between the face and the reward. No participants reported noticing any associations between particular facial expressions and the cue. The associations between the four visual cues we used and the facial expressions and rewards were counterbalanced across participants (Fig. 1A, inset). The total number of trials was 320 (80×4 conditions).

fMRI task. Before the fMRI experiment, 58 experimentally naive participants completed the learning task outside of the scanner. This phase includes presentation of a mask image instead of the facial expressions. This was to avoid habituation in the subsequent scanning session from repeated exposure to the faces. Learning was regarded as successful if the participant chose the correct option in 5 consecutive trials for all cues, and the participants who did not reach this criterion within 200 trials were not invited to the scanning session. Following a 5 min break, 34 participants (17 male, 17 female; mean age 19.8 ± 1.2 years; all right handed) underwent scanning with the presentation of faces. Thus, in the fMRI experiments, we requested that the participants confirm their already learned cue–reward associations. There were two main reasons for this. First, cue–reward learning elicits learning-related attention, which gradually decreases during learning (Maunsell, 2004). Therefore, to dissociate the RPE from the decrease in learning-related attention, we scanned them after they had completed the learning, when we could assume a constant attention level during the outcome. Second, it would take longer than 70 min for some participants to complete the learning process in a task design that would allow us to discriminate brain activity during reward feedback and cue presentation. Our probabilistic task design allows us to detect the RPE signal in the brain even after the completion of learning (Fiorillo et al., 2003; McClure et al., 2003; Haruno and Kawato, 2006). The time intervals between presentation of the visual cue and the button press, the button press and presentation of the out-

come, and the intertrial interval were all randomized to be between 4.5 and 8.5 s. The total number of trials was 80 (20×4 conditions).

Postexperiment ratings of facial expressions. After completing the experiment, all participants rated their subjective emotional responses to each facial expression based on its valence and arousal level (Bradley and Lang, 1994) with a 5-point scale. More negative valences and higher excitement were reported for fearful faces than for neutral faces (valence: fear = -1.19 , neutral = -0.10 , paired $t_{(19)} = -8.79$, $p < 0.001$; arousal: fear = 0.96 , neutral = -0.36 , paired $t_{(19)} = 8.71$, $p < 0.001$). This tendency was also seen in the fMRI experiments (valence: fear = -1.18 , neutral = -0.12 , paired $t_{(33)} = -11.08$, $p < 0.001$; arousal: fear = 1.16 , neutral = -0.54 , paired $t_{(33)} = 15.70$, $p < 0.001$). These data demonstrate that the emotional effects of the fearful face were equivalent in the behavioral and fMRI experiments.

Short-term memory task. One block of the task consisted of two phases. In the encoding phase, participants were required to memorize six visual stimuli, each presented for 250 ms, with 3 following fearful faces presented for 1 s, and 3 following neutral faces. In the retrieval phase, after 35–38 s the subjects reported which of the 12 presented stimuli appeared in the encoding phase, in which only 6 appeared. The entire task comprised 24 blocks.

Exponential curve fitting

We fitted each participant's learning data in the four different conditions (F100, F1, N100, and N1) separately by an exponential function ($c + la \times \exp(-ls \times \text{trial})$), where ls represents learning speed ($0-\infty$), la represents learning amplitude ($0-1$), and c means the convergent value ($0-1$). No goodness of fit was significantly different from the others ($p > 0.5$).

Reinforcement learning model-based analysis

To investigate the learning process in each participant, we adopted a reinforcement learning model (Sutton and Bart, 1998; Daw, 2011). This model assumes that each participant assigns the value function $V_t(S_t)$ to the cue S_t at time t . Learning proceeds by increasing the accuracy of the value representation by updating the value in proportion to the RPE $R_t -$

$V_t(S_t)$, in other words the difference between the expected and actual reward at time t (Eq. 1):

$$V_{t+1}(S_t) = V_t(S_t) + \varepsilon_f(R_t - V_t(S_t)). \quad (1)$$

Our learning model contains the free parameters of the learning rate (ε_f), reward sensitivity (δ_f), value-independent bias for the choice of ¥100 (a_t), ($b_f(a_t)$), and exploration parameter (β_f). The learning rate controls for the effects of the RPE, and reward sensitivity transforms the actual reward (r_t) in yen into a subjective reward (R_t) for each participant (Eq. 2):

$$R_t = \delta_f r_t. \quad (2)$$

In relation to behavioral choice (Eq. 3), the bias term represents a value-independent bias or inclination toward the choice of ¥100, and the exploration parameter controls how deterministically a value function leads to an advantageous behavior:

$$P(a_t | S_t) = \frac{\exp(\beta_f V_t(S_t) + b_f(a_t))}{\sum_{a'} \exp(\beta_f V_t(S_t) + b_f(a'))}. \quad (3)$$

We estimated these four free parameters (denoted as the vector θ) for each participant from their trial-by-trial learning using the maximum likelihood estimation method, which minimizes the negative log-likelihood of the participant's behavior (D), as shown in Equations 4 and 5. This nonlinear minimization of Equation 4 was conducted using the MATLAB function "fmincon":

$$\min - \log P(D | \theta) \quad (4)$$

$$P(D | \theta) = \prod_t P(a_t | S_t). \quad (5)$$

The probability of choosing an action, a_t (¥100 or ¥1), given a visual cue, s_t , was computed based on Equation 3.

We evaluated the significance of each parameter using the Akaike information criteria (AIC) and Bayesian information criteria (BIC) by comparing four models using the learning rate and exploration parameter ($\varepsilon\beta$), $\varepsilon\beta$ with reward sensitivity ($\varepsilon\beta\delta$), $\varepsilon\beta$ with ¥100 bias ($\varepsilon\beta b$), and $\varepsilon\beta$ with both ($\varepsilon\beta\delta b$). For statistical tests of ε and β , we used the nonparametric Wilcoxon signed-rank test because the Kolmogorov–Smirnov test rejected normality.

fMRI data acquisition and analysis

The fMRI scanning was conducted at Tamagawa University using a Siemens Trio TIM 3T scanner with a 12-channel head coil. T2*-weighted echoplanar images (EPIs) were acquired with blood oxygen level-dependent (BOLD) contrast (42 slices; TR 2500 ms; TE 25 ms; flip angle 90°; FOV 192 mm; and resolution $3.0 \times 3.0 \times 3.0$ mm). For each participant, functional data were acquired in one scanning session that consisted of 730 volumes. Two initial volumes in each session were discarded to provide steady-state magnetization. Anatomical images were acquired using a T2-weighted protocol (42 slices; TR 4200 ms; TE 84 ms; flip angle 145°; FOV 199.4 mm; and resolution $0.6 \times 0.4 \times 3.0$ mm horizontal slices) after the EPI session.

Preprocessing and data analysis were performed using SPM8 (Friston et al., 1995). The images were corrected using slice timing correction and realignment, and the volumes were normalized to the MNI space and then resliced into $2 \times 2 \times 2$ mm voxels and spatially smoothed with an 8 mm Gaussian kernel. The data for each participant were analyzed with a standard event-related design.

Together with six-dimensional head movement regressors, we used six constant and parametric regressors for the analysis, including constants for cue, button pressing, and outcome timing, and parametric modulators of value, RPE, and reward size. When we analyzed face timing, we used a face-timing constant instead of a cue-timing constant. In addition, in our split analysis, we separated the RPE at the outcome timing into two components: "reward size" and "minus value." This analysis was done to determine whether a signal correlated with RPE was attributable to either reward size or negative expected value or only to their conjunction (Behrens et al., 2008; Li and Daw 2011). As we assumed that BOLD signal

changes reflected energy consumption, the absolute value of the reward prediction error was used for RPE (Haruno and Kawato, 2006). These regressors were convolved with canonical, time, and dispersion derivative hemodynamic response functions, and the second-level group analysis was conducted using a full-factorial model. The subtraction contrast between the fearful versus neutral conditions was masked with the voxels that were correlated positively with both contrasts (fearful and neutral; $p < 0.005$ uncorrected with an extended threshold of 15 voxels). The anatomical masks and regions of interest (ROIs) were constructed from IBASPM (<http://www.thomaskoenig.ch/Lester/ibaspm.htm>) and Automated Anatomical Labeling map templates, and the statistics within the ROIs were obtained using MarsBaR (<http://marsbar.sourceforge.net/>).

Psychophysiological interaction analysis

We conducted psychophysiological interaction (PPI) analysis to test whether activity in the amygdala during face presentation modulates RPE-correlated activity in the ventral striatum (vStr). The psychological factor was defined as the product of the emotion contrast (fearful versus neutral) and the RPE size. To link striatum activity for outcome timing with amygdala activity for face timing, BOLD signals in the dorsal part of the amygdala (4 mm sphere around the peak at coordinates $x = -26, y = -2, z = -12$, which is the peak of the face-timing activity) during every face presentation with a 5 s delay, to account for the peak BOLD response, were extracted and convolved with the hemodynamic response function to yield a parametric modulator for outcome timing. Finally, we examined whether the interaction between the psychological factor and the parametric modulator explained activity in the vStr.

Results

Behavioral experiment

Figure 1B shows the learning curve for each cue averaged separately over all participants. Learning was faster for the cues associated with fearful faces than those associated with neutral faces. A within-participant three-way ANOVA (trial, reward size, and emotion) revealed significant main effects of trial ($F_{(75,1425)} = 52.937, p < 0.001$) and emotion ($F_{(1,19)} = 4.618, p = 0.045$). The interactions between trial \times emotion \times reward ($F_{(75,1425)} = 2.331, p < 0.001$), trial \times emotion ($F_{(75,1425)} = 3.674, p < 0.001$), and reward \times trial ($F_{(75,1425)} = 1.430, p = 0.011$) were significant. We also fitted each participant's learning with an exponential function. A within-participant ANOVA revealed a statistically significant effect of facial expression ($F_{(1,19)} = 6.597, p = 0.019$) for learning speed. There were no other significant effects (Fig. 1C).

This behavioral effect of fearful faces suggests that emotion acts on learning. However, an alternative explanation is the enhancement of short-term memory by emotion, unrelated to learning (Phelps et al., 2006; Schaefer et al., 2006). To address this possibility, we found no significant differences in the error rate between fearful (9.6%) and neutral (10.3%) conditions ($t_{(19)} = -1.415, p = 0.173$) in a short-term memory task (see Materials and Methods).

To examine the computational process behind the acceleration of learning by emotion, we analyzed learning using a reinforcement learning model. Using model selection, the $\varepsilon\beta b$ model was selected by AIC, and the $\varepsilon\beta$ model by BIC (Fig. 1D). Analysis with the $\varepsilon\beta b$ model showed that the learning rate was higher in the fearful than in the neutral condition [$\varepsilon_F = 0.038 \pm 0.011$ (mean \pm SEM), $\varepsilon_N = 0.010 \pm 0.004$, Wilcoxon signed-rank test: $Z = -2.053, p = 0.040$] (Fig. 1E, left). In addition, the bias for ¥100 was less than zero in the fearful condition ($b_F = -0.211 \pm 0.080, t_{(19)} = -2.628, p = 0.017$) but not different from zero in the neutral condition (Fig. 1E, right). The $\varepsilon\beta$ model yielded a highly consistent statistical difference for ε and β ($\varepsilon_F = 0.034 \pm 0.010, \varepsilon_N = 0.006 \pm 0.003, Z = -2.062, p = 0.039$; and

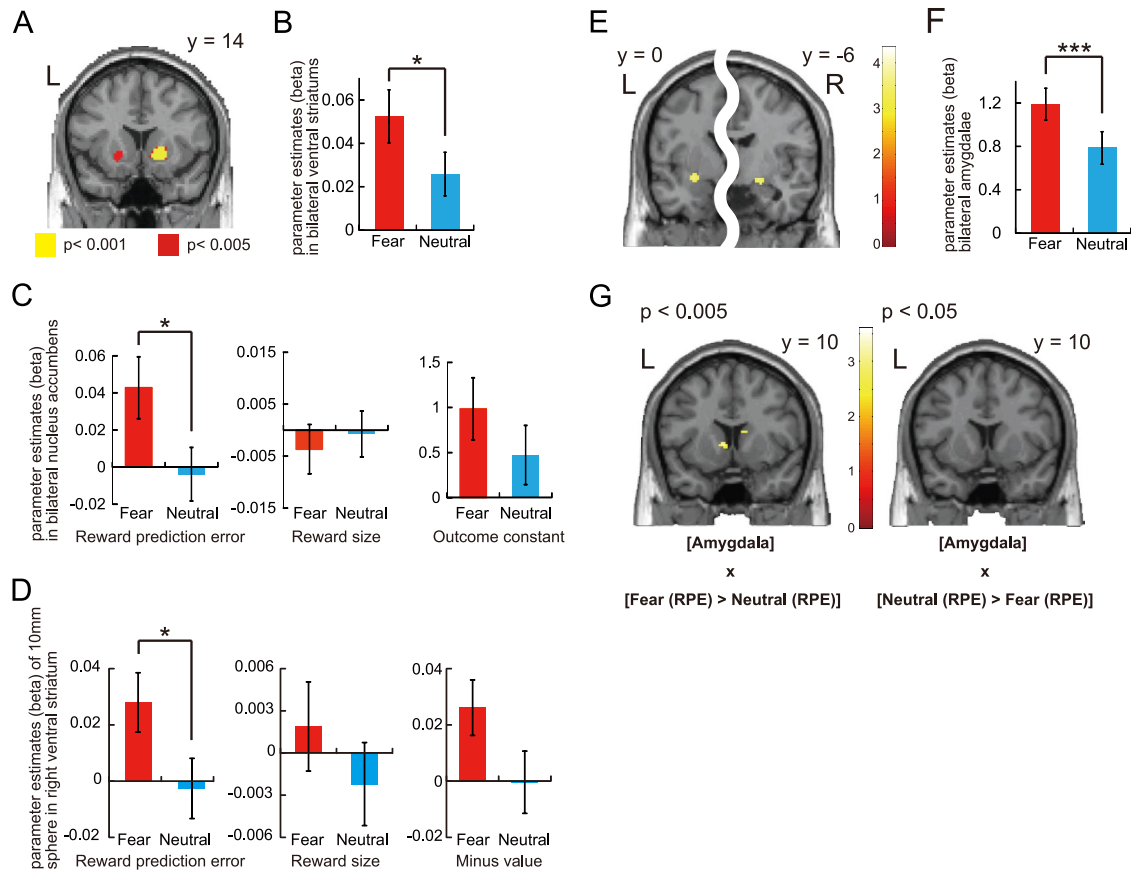


Figure 2. fMRI results. **A**, Activity in the right vStr (18, 16, -4) and activity in the left vStr (-20, 14, -8) that were correlated with the RPE (Fearful + Neutral). **B**, Activity in the vStr 10 mm radius ROI. **C**, Comparison of the correlations in the anatomically defined bilateral NAcc for the RPE, outcome constant, and reward size. **D**, Comparison of the correlations in 10 mm sphere ROI for the RPE (= reward size - value), reward size, and minus value revealed by the split analysis. **E**, Activity in the left (-28, 0, -10) and right (20, -6, -10) amygdala at the time of face presentation (Fearful - Neutral, $p < 0.001$, uncorrected). **F**, Activity in the anatomically defined bilateral amygdala ROI. **G**, Activity in the striatum as determined by PPI analysis ($p < 0.005$, uncorrected).

$\beta_F = 9.917 \pm 2.246$, $\beta_N = 14.556 \pm 1.920$, $Z = -1.643$, $p = 0.100$, respectively).

These results indicate that the presentation of a fearful face enhances the effects of the RPE by increasing the learning rate. To test this computational hypothesis, we used fMRI to examine whether the emotional stimulus changed the size of the RPE signal in the brain.

fMRI experiment

The estimated learning speeds in the pre-scan learning session were not different between the pre-fearful cues (which would later be presented with the fearful face), and the pre-neutral cues (which would later be presented with the neural face). These results validate that there were no differences among the four visual cues before the scanning session. Additionally, we did not find any significant differences in the error rate or rate of not responding among the four conditions in the scanning session.

Activity correlated with the RPE

We first searched for the brain regions in which activity was correlated with RPEs at the time of the reward outcome (Fig. 1A). Significant activity was found in the right vStr ($x = -20$, $y = 14$, $z = -8$, $p < 0.001$; uncorrected for multiple comparisons; Fig. 2A), and in the left vStr with a slightly lower statistical threshold ($x = 18$, $y = 16$, $z = -4$, $p < 0.005$). Importantly, activity averaged within a 10 mm sphere around the bilateral peak voxels

revealed a higher correlation (β value) with the RPE in the fearful than in the neutral condition (paired $t_{(33)} = 2.132$, $p = 0.041$; Fig. 2B). This was also true for an anatomically defined ROI in the bilateral nucleus accumbens (NAcc) (paired $t_{(33)} = 2.409$, $p = 0.022$; Fig. 2C, left).

However, a pertinent question arises about whether the dissociable correlation between the fearful and neutral conditions in the NAcc is specific to the RPE, or more generally, is seen at the time of the reward outcome. To address this, we compared the fearful and neutral conditions for NAcc activity correlated with the constant and reward size regressors at the time of the reward outcome and found no difference (paired $t_{(33)} = 1.197$, $p = 0.24$; and paired $t_{(33)} = -0.573$, $p = 0.57$, respectively; Figure 2C, middle and right). Although we found brain activity correlated with RPEs in several areas (Table 1, A), none of these areas, except for the vStr, showed differences between the fearful and neutral conditions (threshold: $p < 0.05$). The results for the constant and reward size at the time of the outcome are shown in Table 1, B and C.

We also conducted the split analysis of the RPE signal, which examined whether either of “reward size (R) or “minus value (-V)” was modulated by the emotion or this modulation was significant only for total RPE (Behrens et al., 2008; Li and Daw 2011) (see also Materials and Methods). The contrast of Fear(R - V) > Neutral(R - V) exhibited a significant difference in activity in right vStr, which overlapped with Figure 2A (peak voxel $t_{(33)} = 3.27$, $p = 0.001$, $x = 16$, $y = 18$, $z = -4$), and the Beta

Table 1. Summary of the fMRI peak voxels

Area	L/R	<i>x</i>	<i>y</i>	<i>z</i>	<i>k</i>	<i>t</i> value
A—Prediction error activation at Outcome timing						
Common regions of fearful and neutral conditions						
Ventral striatum	R	18	16	−4	155	4.957
Primary motor area	L	−42	−10	54	103	4.042
Inferior frontal operculum	L	−44	6	30	33	3.734
Primary motor area	R	46	−8	56	32	3.675
Cuneus	L	−8	−88	40	28	3.436
Fear-neutral masked by $p < 0.005$ common regions ^a						
Neutral-fear masked by $p < 0.005$ common regions ^a						
B—Reward size activation at Outcome timing						
Common regions of fearful and neutral conditions						
Calcarine sulcus	L	−14	−90	−6	483	6.615
Lingual gyrus	R	18	−88	−4	78	4.456
Fear-neutral masked by $p < 0.005$ common regions ^a						
Neutral-fear masked by $p < 0.005$ common regions ^a						
C—Constant activation at outcome timing						
Common regions of fearful and neutral conditions						
Middle occipital lobe	L	−20	−92	4	7502	16.573
Inferior occipital lobe	R	32	−84	−6	8017	15.598
Primary motor area	R	46	8	32	3953	13.329
Insula cortex	L	−28	24	0	962	12.263
Primary motor area	L	−44	4	30	1867	10.775
Supplementary motor area	R	8	20	46	2312	8.652
Pallidum	L	−10	6	−4	126	5.377
Pallidum	R	10	6	−6	42	4.094
Middle frontal lobe	L	−26	54	14	23	3.571
Fear-neutral masked by $p < 0.005$ common regions						
Insula cortex	R	30	18	−16	22	4.089
Neutral-fear masked by $p < 0.005$ common regions ^a						
D—Expected value						
Common regions of fearful and neutral conditions						
Putamen	L	−30	12	−4	59	3.905
Anterior cingulate	R	14	48	26	22	3.631
Fear-neutral ^a						
Neutral-fear ^a						
E—Face timing activation						
Fear-neutral masked by $p < 0.005$ common regions						
Inferior frontal triangularis	R	56	34	8	85	4.325
Amygdala	L	−28	0	−10	47	3.555
Amygdala	R	20	−6	−10	16	3.457
Neutral-fear masked by $p < 0.005$ common regions ^a						
F—PPI analysis						
Fear-neutral masked by striatum-anatomical ROI						
Dorsal striatum	R	10	16	12	58	3.590
Ventral striatum	L	−6	10	2	24	3.097
Neutral-fear masked by striatum-anatomical ROI ^a						

The uncorrected threshold was $p < 0.001$ (A–E) or $p < 0.005$ (F), and the cluster size threshold was 15 voxels. *k* represents the cluster size.

^aNo suprathreshold clusters.

values for the conjunction of R and $-V$ averaged in a 10 mm sphere around the peak voxel showed a significant difference between fearful and neutral conditions for RPE ($t_{(33)} = 2.097$, $p = 0.044$; Fig. 2D, left). However, we could not detect the significant differences for R ($t_{(33)} = 1.116$, $p = 0.272$; Fig. 2D, middle) or $-V$ ($t_{(33)} = 1.861$, $p = 0.072$; Fig. 2D, right). These results from the split analysis are consistent with those from the original RPE-based analysis (Fig. 2C) that the emotion affected RPE rather than reward size or value. All of these were consistent with the model selection results in the computational model-based analysis, where reward sensitivity did not help explain the behavioral data.

Activity correlated with the expected value

Consistent with previous studies, activity in the left putamen was correlated with the expected value ($p < 0.001$ uncorrected; Table

1, D). Importantly, this correlation was not different between the fearful and neutral conditions.

Activity during face presentation

We analyzed brain activity with respect to the time of face presentation. Consistent with previous human imaging studies on the processing of facial expressions (Whalen et al., 2009), the bilateral amygdala showed higher activity in the fearful than the neutral condition ($p < 0.001$ uncorrected; Fig. 2E; Table 1, E). Furthermore, this characteristic activity of the amygdala was also evident when we masked and averaged bilateral amygdala activity using an anatomical ROI (paired $t_{(33)} = 3.64$, $p < 0.001$; Fig. 2F). The same subtraction between the fearful and neutral conditions revealed activity in the inferior frontal triangularis (Table 1, E), but no difference was observed using the opposite subtraction (i.e., neutral minus fear), even with a weak threshold ($p < 0.01$).

Psychophysiological interactions between amygdala and striatum activity

We have seen that the RPE signal in the vStr was enhanced in the fearful condition, and activity in the amygdala was greater in the fearful condition. Accordingly, because we hypothesized that amygdala activation has a functional link with RPE signals in the striatum, we conducted a PPI analysis. Figure 2G, left, displays the activity in the left vStr ($t = 3.02$, $p = 0.001$, $x = -6$, $y = 10$, $z = 2$, cluster size = 24) that is correlated with the interaction between the amygdala activity and psychological factor ([Fear (RPE) > Neutral (RPE)]) (compare with Materials and Methods). This analysis also identified the right dorsal striatum ($t = 3.59$, $p < 0.001$, $x = 10$, $y = 16$, $z = 12$, cluster size = 58). Importantly, the inverse contrast ([Amygdala] \times [Neutral (RPE) > Fear (RPE)]) did not detect difference in striatum activity even with a moderate threshold ($p < 0.05$) (Fig. 2G, right; Table 1, F).

Similarly, we searched for the functional link with the right inferior frontal cortex triangularis, but no significant activity was found, even with a weak threshold ($p < 0.01$). These results suggest an important contribution of amygdala activity to the enhancement of RPE signals in the vStr by emotion.

Discussion

Herein, we report that the presentation of a task-independent fearful face increased learning speed during probabilistic reward learning in comparison with the presentation of a neutral face. In parallel, we demonstrated an increased RPE signals in the ventral striatum (mainly the NAcc). We also found that amygdala activity during face presentation modulates RPE-correlated activity in the vStr in a trial-by-trial analysis.

Previous studies on decision making showed that emotional factors affect decision making and human brain activity, especially in the vStr. For example, people tend to be more sensitive to losses than gains that are of an equivalent magnitude (Kahneman and Tversky, 1979). An fMRI study reported that activity in the vStr reflected individual differences in expected loss aversion (Tom et al., 2007). It has also been reported that the expectation of the reward value in the vStr is modulated by the conscious control of a participant's own emotional reaction to a reward (Delgado et al., 2008). Related to this, Knutson et al. (2008) demonstrated that the task-independent pleasant picture presented before cues enhanced participant's risky choices and value signals in the vStr. In contrast, this study examined the effects of a task-unrelated emotional face on learning and revealed modulation of RPE signals in the vStr by emotion.

Related to this, one may ask whether enhanced RPE brain activity can be interpreted in the context of pavlovian instrumental transfer (PIT) (Bray et al., 2008; Talmi et al., 2008), which is a change in a cue value when it follows a positive or negative cue. Talmi et al. (2008) reported that a pavlovian-conditioned cue changed the value signal in the NAcc, and in the amygdala, during subsequent instrumental conditioning. These two phenomena are comparable in the sense that a task-independent cue changes the reward representations of subsequent cues and brain activity. However, fearful faces are likely to be perceived as negative stimuli and should slow cue–reward associations, which contradicts the results of this study. In addition, we found an increase in the RPE signal rather than an increase in the value signal (reward prediction). Therefore, we do not think that the observed accelerated learning can be linked to PIT. Interestingly, our model-based analysis revealed that the participants had a negative bias for ¥100 following the presentation of a fearful face, and this

tendency could be related to PIT. Similarly, a recent fMRI study reported a facial expression-dependent bias using positive and negative facial expressions as reward-predicting cues (Evans et al., 2011). It is also known that novel objects can increase value representations and reward responses in the striatum (Guitart-Masip et al., 2010). However, we matched the novelty of the faces used in the current study.

Subjects simply confirmed their already learned association in the fMRI experiments, and there is a possibility that the behavioral and fMRI results do not have one-to-one interpretation. However, previous electrophysiological studies showed (Fiorillo et al., 2003) that dopamine neuron firing in a probabilistic task represents RPE even after the learning, and similar RPE activity was reported in the striatum in a fMRI study (McClure et al., 2003). These results indicate that activity in the striatum in the current study successfully detected RPE signal. It is also indicated from the task design that learning-related attention does not have a major role in the emotional modulation of the RPE signal (Maunsell, 2004). However, the dynamic property of the modulation during learning remains an open question for future study.

We showed that activity in vStr is correlated with both RPE and amygdala activity during face presentation (Fear > Neutral). This suggests that the amygdala plays a role in modulating RPE signals, although it is difficult to draw strong conclusions regarding causality with fMRI. Consistent with this view, a recent optogenetic study has revealed that excitatory glutamatergic transmissions from the basolateral amygdala to the NAcc are indispensable for cue–reward association learning (Stuber et al., 2011). Similarly, some recent fMRI studies have suggested the possibility that emotional stimuli enhance amygdala–striatum interactions and learning (Delgado et al., 2009; Schlund and Cataldo, 2010; Li et al., 2011) in the context of aversive conditioning.

Considering the temporally separated timing of amygdala and RPE-correlated striatum activity, they could be related through indirect pathways between the vStr and the amygdala, via the midbrain (Ahn and Phillips, 2002) or orbitofrontal cortex (Schoenbaum et al., 2003). However, we think that direct interactions are more likely, as neither the midbrain nor the orbitofrontal cortex showed significant activity differences between the fearful and neutral conditions.

Finally, a natural question may arise as to why task-independent emotion enhances RPEs and makes learning faster. One possible explanation may be that this mechanism contributes to the individual obtaining more rewards in complex environments by using the facial expressions of others and the surrounding atmosphere. Such an advantage may have allowed the development of the computational and neural processes found in this study.

References

- Ahn S, Phillips AG (2002) Modulation by central and basolateral amygdala nuclei of dopaminergic correlates of feeding to satiety in the rat nucleus accumbens and medial prefrontal cortex. *J Neurosci* 22:10958–10965. [Medline](#)
- Behrens TE, Hunt LT, Woolrich MW, Rushworth MF (2008) Associative learning of social value. *Nature* 456:245–249. [CrossRef Medline](#)
- Bradley MM, Lang PJ (1994) Measuring emotion: the self-assessment manikin and the semantic differential. *J Behav Ther Exp Psychiatry* 25:49–59. [CrossRef Medline](#)
- Bray S, Rangel A, Shimojo S, Balleine B, O'Doherty JP (2008) The neural mechanisms underlying the influence of pavlovian cues on human decision making. *J Neurosci* 28:5861–5866. [CrossRef Medline](#)
- Daw DN (2011) Trial-by-trial data analysis using computational models. In: *Decision making, affect, and learning: attention and performance*

- XXIII (Delgado MR, Phelps EA, Robbins TW, eds), pp 3–38. Oxford, UK: Oxford UP.
- Dayan P (2008) The role of value systems in decision making. In: *Better than conscious? Decision making, the human mind, and implications for institutions* (Engel C, Singer W, eds), pp 51–70. Cambridge, MA: MIT Press.
- Delgado MR, Gillis MM, Phelps EA (2008) Regulating the expectation of reward via cognitive strategies. *Nat Neurosci* 11:880–881. [CrossRef Medline](#)
- Delgado MR, Jou RL, Ledoux JE, Phelps EA (2009) Avoiding negative outcomes: tracking the mechanisms of avoidance learning in humans during fear conditioning. *Front Behav Neurosci* 3:33. [CrossRef Medline](#)
- Evans S, Fleming SM, Dolan RJ, Averbeck BB (2011) Effects of emotional preferences on value-based decision-making are mediated by mentalizing and not reward networks. *J Cogn Neurosci* 23:2197–2210. [CrossRef Medline](#)
- Fiorillo CD, Tobler PN, Schultz W (2003) Discrete coding of reward probability and uncertainty by dopamine neurons. *Science* 299:1898–1902. [CrossRef Medline](#)
- Friston KJ, Holmes AP, Worsley KJ, Poline JB, Frith CD, Frackowiak RSJ (1995) Statistical parametric maps in functional brain imaging: a general linear approach. *Hum Brain Mapp* 2:189–210. [CrossRef](#)
- Guitart-Masip M, Bunzeck N, Stephan KE, Dolan RJ, Düzel E (2010) Contextual novelty changes reward representations in the striatum. *J Neurosci* 30:1721–1726. [CrossRef Medline](#)
- Haruno M, Kawato M (2006) Different neural correlates of reward expectation and reward expectation error in the putamen and caudate nucleus during stimulus-action-reward association learning. *J Neurophysiol* 95:948–959. [CrossRef Medline](#)
- Kahneman D, Tversky A (1979) Prospect theory: an analysis of decision under risk. *Econometrica* 47:263–291. [CrossRef](#)
- Knutson B, Wimmer GE, Kuhnen CM, Winkielman P (2008) Nucleus accumbens activation mediates the influence of reward cues on financial risk taking. *Neuroreport* 19:509–513. [CrossRef Medline](#)
- Li J, Daw ND (2011) Signals in human striatum are appropriate for policy update rather than value prediction. *J Neurosci* 31:5504–5511. [CrossRef Medline](#)
- Li J, Schiller D, Schoenbaum G, Phelps EA, Daw ND (2011) Differential roles of human striatum and amygdala in associative learning. *Nat Neurosci* 14:1250–1252. [CrossRef Medline](#)
- Maunsell JH (2004) Neuronal representations of cognitive state: reward or attention? *Trends Cogn Sci* 8:261–265. [CrossRef Medline](#)
- McClure SM, Berns GS, Montague PR (2003) Temporal prediction errors in a passive learning task activate human striatum. *Neuron* 38:339–346. [CrossRef Medline](#)
- O'Doherty JP, Dayan P, Friston K, Critchley H, Dolan RJ (2003) Temporal difference models and reward-related learning in the human brain. *Neuron* 38:329–337. [CrossRef Medline](#)
- Phelps EA, Ling S, Carrasco M (2006) Emotion facilitates perception and potentiates the perceptual benefits of attention. *Psychol Sci* 17:292–299. [CrossRef Medline](#)
- Schaefer A, Braver TS, Reynolds JR, Burgess GC, Yarkoni T, Gray JR (2006) Individual differences in amygdala activity predict response speed during working memory. *J Neurosci* 26:10120–10128. [CrossRef Medline](#)
- Schlund MW, Cataldo MF (2010) Amygdala involvement in human avoidance, escape and approach behavior. *Neuroimage* 53:769–776. [CrossRef Medline](#)
- Schoenbaum G, Setlow B, Saddoris MP, Gallagher M (2003) Encoded predicted outcome and acquired value in orbitofrontal cortex during cue sampling depends upon input from basolateral amygdala. *Neuron* 39:855–867. [CrossRef Medline](#)
- Schultz W, Dayan P, Montague PR (1997) A neural substrate of prediction and reward. *Science* 275:1593–1599. [CrossRef Medline](#)
- Stuber GD, Sparta DR, Stamatakis AM, van Leeuwen WA, Hardjoprajitno JE, Cho S, Tye KM, Kempadoo KA, Zhang F, Deisseroth K, Bonci A (2011) Excitatory transmission from the amygdala to nucleus accumbens facilitates reward seeking. *Nature* 475:377–380. [CrossRef Medline](#)
- Sutton RS, Barto AG (1998) *Reinforcement learning*. Cambridge, MA: MIT Press.
- Talmi D, Seymour B, Dayan P, Dolan RJ (2008) Human pavlovian-instrumental transfer. *J Neurosci* 28:360–368. [CrossRef Medline](#)
- Tom SM, Fox CR, Trepel C, Poldrack RA (2007) The neural basis of loss aversion in decision-making under risk. *Science* 315:515–518. [CrossRef Medline](#)
- Whalen PJ, Davis FC, Oler JA, Kim H, Kim MJ, Neta M (2009) Human amygdala responses to facial expressions of emotion. In: *The human amygdala* (Whalen PJ, Phelps EA, eds), pp 265–288. New York: Guilford Press.