Individual differences in reinforcement learning: Behavioral, electrophysiological, and neuroimaging correlates

Diane L. Santesso, Daniel G. Dillon, Jeffrey L. Birk, Avram J. Holmes, Elena Goetz, Ryan Bogdan, and Diego A. Pizzagalli*

Department of Psychology, Harvard University, 1220 William James Hall, 33 Kirkland Street, Cambridge, MA 02138, USA

ARTICLE INFO

ABSTRACT

During reinforcement learning, phasic modulations of activity in midbrain dopamine neurons are conveyed to the dorsal anterior cingulate cortex (dACC) and basal ganglia (BG) and serve to guide adaptive responding. While the animal literature supports a role for the dACC in integrating reward history over time, most human electrophysiological studies of dACC function have focused on responses to single positive and negative outcomes. The present electrophysiological study investigated the role of the dACC in probabilistic reward learning in healthy subjects using a task that required integration of reinforcement history over time. We recorded the feedback-related negativity (FRN) to reward feedback in subjects who developed a response bias toward a more frequently rewarded ("rich") stimulus ("learners") versus subjects who did not ("non-learners"). Compared to non-learners, learners showed more positive (i.e., smaller) FRNs and greater dACC activation upon receiving reward for correct identification of the rich stimulus. In addition, dACC activation and a bias to select the rich stimulus were positively correlated. The same participants also completed a monetary incentive delay (MID) task administered during functional magnetic resonance imaging. Compared to non-learners, learners displayed stronger BG responses to reward in the MID task. These findings raise the possibility that learners in the probabilistic reinforcement task were characterized by stronger dACC and BG responses to rewarding outcomes. Furthermore, these results highlight the importance of the dACC to probabilistic reward learning in humans.

© 2008 Elsevier Inc. All rights reserved.

Introduction

Optimal behavior relies on the ability to internally monitor responses and to evaluate external reinforcements in order to learn about the appropriateness of those responses. Mounting evidence suggests that this reinforcement learning may depend on the basal ganglia and midbrain dopamine system. Accordingly, non-human primate studies have shown that negative reinforcement elicits phasic decreases in neuronal activity of midbrain dopaminergic neurons (i.e., negative prediction error), whereas positive reinforcement elicits increases of dopaminergic activity (i.e., positive prediction error) (Montague et al., 2004; Schultz, 2007). These phasic modulations are thought to act as teaching signals for the anterior cingulate cortex (ACC) and basal ganglia (BG) to implement goal-directed behaviors and update predictions of success or failure (Holroyd and Coles, 2002). This model has received support in the human electrophysiology literature with respect to negative reinforcement (Holroyd and Coles, 2002; Holroyd and Krigolson, 2007; Hajcak et al., 2007), but fewer studies have examined positive reinforcement. In particular, the role of the human dorsal region of the ACC (dACC) in probabilistic reward learning is not well understood.

The dACC appears critical for encoding rewards and using reinforcement histories to guide behavior (Akitsuki et al., 2003; Amiez et al., 2006; Ernst et al., 2004; Rushworth et al., 2007). In non-human primates, ACC lesions impair the ability to integrate reinforcement history over time and choose advantageous responses (Kennerley et al., 2006). In humans, modulation of behavior by reinforcement history can be investigated using two-alternative probabilistic reward tasks in which correct responses to the two stimuli are differentially rewarded; the development of a response bias towards the more frequently rewarded ("rich") stimulus indicates reward sensitivity (Pizzagalli et al., 2005, 2008). Impaired learning on this task has been demonstrated in anhedonic individuals (Pizzagalli et al., 2005), mood disorder patients characterized by dysfunctional reward processing (Pizzagalli et al., in press-a, in press-b), and in healthy participants receiving a pharmacological challenge hypothesized to disrupt phasic DA signaling

(Pizzagalli et al., 2008). This task thus appears suitable for examining reward learning mediated by the midbrain dopamine system. Consistent with this assumption, in a computational model of striatal–cortical function (Frank, 2005), blunted response bias was accounted for by reduced DA bursts to reward (Santesso et al., unpublished), suggesting that this task is sensitive to learning mediated by the midbrain DA system. The primary goal of the present study was to examine reward learning during this task using the feedback-related negativity (FRN) as an electrophysiological index of ACC reward-related activity.

We recorded the feedback-related negativity (FRN) as an index of ACC reward-related activity. The FRN peaks 200–400 ms following feedback and has been localized to various regions of the cingulate cortex, including the dorsal ACC (dACC; Miltner et al., 1997; Gehring and Willoughby, 2002), medial prefrontal cortex (Muller et al., 2005; Nieuwenhuis et al., 2005; Van Veen et al., 2004), and the posterior cingulate cortex (PCC), particularly in response to positive versus negative feedback (Muller et al., 2005; Nieuwenhuis et al., 2005). The FRN is thought to reflect transmission of a DA signal from the BG (Holroyd and Coles, 2002). Although commonly used to study negative reinforcement, the FRN is reliably elicited by positive feedback (Hajcak et al., 2005; Holroyd and Coles, 2008; Muller et al., 2005; Oliveira et al., 2007), and appears as a relatively more positive ERP deflection (compared to that elicited by negative feedback). We predicted that (1) reward feedback delivered after correctly identifying the rich stimulus would elicit more positive FRNs and greater dACC activation in individuals who developed a response bias toward the rich stimulus ("learners") versus those who did not ("non-learners"); and (2) dACC activation would correlate positively with reward learning and the FRN.

A secondary goal of this study was to test whether "learners" and "non-learners" would differ in brain activation in the BG, which includes the globus pallidus and three striatal regions (nucleus accumbens, caudate, and putamen), in response to reward feedback. We were able to address this issue because a sub-set of the ERP participants also participated in an fMRI session that featured a monetary incentive delay (MID) task, which has been used to probe reward-related activity in the BG (Dillon et al., 2008; Knutson et al., 2003). Relevant to the present study, recent neuroimaging findings indicate that optimal performance in probabilistic reward-learning tasks is accompanied by recruitment of striatal regions. Accordingly, in a probabilistic reward-learning task, learners (but not non-learners) showed significant correlations between prediction errors and fMRI signal in dorsal and ventral striatal regions (Schonberg et al., 2007). Along similar lines, participants who learned contingencies between specific cues and the reward probabilities and used them adaptively in a gambling task showed robust striatal responses to reward feedback, particularly at early stages of learning (Delgado et al., 2005). Based on these findings, we predicted that, relative to non-learners, learners in the probabilistic reward task would show larger BG responses to reward feedback during the MID task.

Materials and methods

Participants

Two hundred and thirty-seven adults between 18 and 40 years old (105 men, mean age = 24.5 years) were recruited from Harvard University and the surrounding community for a larger study investigating the neurobiology and molecular genetics of reward processing. Participants meeting the following criteria were excluded: present medical or neurological illness (ADHD, head injury, loss of consciousness, seizures), current alcohol/substance abuse or smoking, claustrophobia, use of psychotropic medications during the last 2 weeks, and pregnancy. All eligible participants were right-handed (Chapman and Chapman, 1987).

The study included three sessions. During the first session, all participants completed the probabilistic reward task at the Affective Neuroscience Laboratory, Harvard University. Sixty-seven subjects were excluded due to failure to meet inclusion criteria ($n=31$), prior task exposure ($n=4$), non-compliance and/or performance below chance level ($n=31$), and outlier status ($n=1$). Of the remaining 170 eligible subjects, 47 were invited to complete an electroencephalogram (EEG) and fMRI session (the order of which was counterbalanced). These 47 subjects were selected to cover a wide range of individual differences in reward learning, which was measured by a response bias difference score (block 3 – block 1; see below). To this end, we first selected participants in the upper and lower 20% of the distribution of reward learning; next, remaining subjects were selected in order to achieve a continuum in reward learning, so that selected participants would be representative of the general population. Of the 47 participants, 41 agreed to perform the probabilistic reward task while EEG was recorded, whereas 38 completed the monetary incentive delay (MID) task during functional scan acquisition at the Martinos Center for Biomedical Imaging. For both the EEG and fMRI datasets, 30 participants had usable data; data from remaining participants were lost due to an insufficient number of artifact-free EEG trials, equipment failure, incomplete data, non-compliance, motion artifacts (fMRI), and technical difficulties. Of the 30 participants with EEG data, 21 had usable data from all three sessions.

Participants received $5 for the first session plus $5.80–$6.20 in earnings in the probabilistic reward task. For the EEG session, participants received $20 plus $24.60 (fixed amount) in task earnings. For the fMRI session, participants received $60 plus $20–$22 in earnings for the MID task. Participants provided written informed consent. All procedures were approved by the Committee on the Use of Human Subjects at Harvard University and the Partners-Massachusetts General Hospital Internal Review Board.

Procedures and tasks

Probabilistic reward task (EEG session)

During the EEG session, participants repeated the reward-learning task used during subject selection, which has been described in detail elsewhere (e.g., Pizzagalli et al., 2005, 2008; see also Tripp and Alsop, 1999). Briefly, the task included 300 trials, divided into 3 blocks of 100 trials. Each trial started with the presentation of a fixation point for 1400 ms. A mouthless cartoon face was then presented for 500 ms followed by the presentation of this face with either a short mouth or a long mouth for 100 ms. Participants were asked to indicate whether a short or long mouth was presented by pressing one of two keys (counterbalanced across subjects). For each block, only 40 correct responses were followed by positive feedback ("Correct!! You won 20 cents"), displayed for 1500 ms in the center of the screen followed by a blank

screen for 250 ms. [Unlike the EEG session, 5-cent rewards were used for the behavioral pre-screening session involving 237 participants.] To induce a response bias, an asymmetrical reinforcer ratio was used: correct responses for the rich stimulus were rewarded three times (30:10) more frequently than correct responses for the other ("lean") stimulus. Participants were informed at the outset that not all correct responses would be rewarded but were not aware that one of the stimuli would be rewarded more frequently. For 16 participants, the same stimulus (e.g., rich mouth) was disproportionally rewarded in both sessions; for the remaining 14 participants, the more frequently rewarded stimulus was switched across the behavioral and EEG session.

After completing the task, participants filled out various questionnaires, including the BDI-II (Beck et al., 1996) and the 62-item version of the Mood and Anxiety Symptom Questionnaire (MASQ; Watson et al., 1995) to assess depressive symptoms, anxiety symptoms, anhedonic depression, and general distress.

Monetary incentive delay task (fMRI session)

The MID task was identical to one recently used by our group in an independent study to dissociate anticipatory versus consummatory phases of incentive processing and reliably elicits activity in brain reward circuitry, including the four components of the BG (nucleus accumbens, caudate, putamen, and globus pallidus) (Dillon et al., 2008; Knutson et al., 2003). Participants completed 5 blocks of 24 trials. Each trial began with the presentation of one of three equally probable cues (duration: 1.5 s) that signaled potential monetary rewards (+\$), no incentive (0\$), or monetary losses (−\$). Following a jittered inter-stimulus interval (ISI: 3–7.5 s), a red square was presented; participants responded to the target with a button press. Following a second jittered ISI (4.4–8.9 s), feedback was presented indicating a gain, no change, or loss: successful reward trials yielded a gain (range: \$1.96 to \$2.34; mean: \$2.15); unsuccessful reward trials yielded no gain; successful punishment trials yielded no loss; and unsuccessful punishment trials yielded a loss (range: −\$1.81 to −\$2.19; mean: −\$2.00). No-incentive trials were always followed by no change feedback. The task design and timing were optimized using a genetic algorithm that maximized the statistical orthogonality of the conditions under investigation (Wager and Nichols, 2003).

Participants were told that their reaction time (RT) to the target affected trial outcomes, such that rapid RTs increased the probability of winning money on reward trials and decreased the probability of losing money on loss trials. To achieve a balanced design, delivery of outcomes was decoupled from RT such that 50% of reward and loss trials resulted in delivery of gains and losses, respectively. However, to maximize task believability, target presentation duration was different for successful and unsuccessful trials. To this end, participants were instructed to perform a practice block of the MID task involving 40 trials while in the scanner; RT were collected and subsequently used to titrate target duration during the experimental blocks. Thus, when a successful or unsuccessful trial was scheduled, the target was presented for a duration corresponding to the 85th or 15th percentiles, respectively, of RTs collected during the practice. This subtle manipulation allowed participants to be generally "successful" on scheduled success trials, and "unsuccessful" on scheduled unsuccessful trials. Finally, to boost task engagement, participants were informed that good performance throughout the task would allow them to qualify for a sixth "bonus" block (not analyzed here) involving larger gains (\$3.63–\$5.18) and few penalties (all participants "qualified" for this bonus block). In two prior samples, we have shown that the combination of instructions and task parameters used in the current version of the MID task leads to sustained motivated behavior (i.e., significantly shorter RT for reward and loss trials compared to no-incentive trials across the five blocks), and robust activation in reward-related brain regions (Dillon et al., 2008).

Data collection and reduction

Behavioral data

For behavioral analyses, the main variables of interest were response bias and reward learning during the probabilistic reward task administered at the EEG session. Response bias (b) assesses the systematic preference for the response paired with the more frequent reward (rich stimulus), and was computed as:

$$\log b = \frac{1}{2}\log\left(\frac{Rich_{correct}*Lean_{incorrect}}{Rich_{incorrect}*Lean_{correct}}\right)$$

Following prior recommendations, 0.5 was added to every cell of the detection matrix to allow calculation of response bias in cases with a zero in one cell of the formula (Hautus, 1995). Reward learning was computed as the response bias score from block 3 minus the response bias score from block 1, as this calculation captures the development of response bias across the task. Negative values represent poor reward learning (i.e., failure to develop a response bias) and have been associated with elevated self-reported anhedonic symptoms (Pizzagalli et al., 2005) and purportedly reduced phasic dopaminergic transmission (Pizzagalli et al., 2008), whereas positive values indicate increased sensitivity to reward feedback. On the basis of this difference score, two groups were formed for the ERP analyses: a non-learners group ($n=14$), comprising individuals who failed to develop a response bias (i.e., a negative score); and a learners group ($n=16$), comprising those individuals displaying successful reward learning from block 1 to block 3.

Scalp ERP data

EEG was recorded continuously using a 128-channel Electrical Geodesics system (EGI Inc., Eugene, OR) at 250 Hz with 0.1–100 Hz analog filtering referenced to the vertex. Impedance of all channels was kept below 50 kΩ. Data were segmented and re-referenced off-line to an average reference. EEG epochs were extracted beginning 200 ms before and ending 600 ms after feedback presentation during each block. Data were processed using Brain Vision Analyzer (Brain Products GmbH, Germany). Each trial was visually inspected for movement artifacts and manually removed followed by automatic artifact removal with a ±75 μV criterion. Eye-movement artifacts were corrected by Independent Component Analysis (e.g., Makeig et al., 1997). A pre-stimulus baseline between −200 and 0 ms was used. The amplitude of the ERP was derived from each individual's average waveform for the midline sites Fz and FCz, where the FRN is typically largest, and filtered at 1–30 Hz. The FRN was defined as the most negative peak 200–400 ms after reward feedback following correct identification of the rich stimulus.

To allow participants to be exposed to the differential reinforcement schedule, primary analyses focused on ERPs computed by averaging artifact-free EEG epochs time-locked to reward feedback for the rich stimulus from blocks 2 and 3 ("blocks 2 and 3"). For analyses evaluating FRN changes over time (see below), secondary analyses also considered ERP peak data from block 1. To ensure that findings were not affected by the relatively low number of trials available for some of the ERP averaging (e.g., 30 rewarded rich trials in block 1), analyses were re-run by considering both rich and lean rewarded trials. Findings were essentially identical to the ones presented in the main text (results available upon request). Given the asymmetric reinforcement ratio used in probabilistic reward task, it was not possible to obtain a sufficient number of trials to analyze reward feedback following *lean* stimuli.

Source localization of ERP data

Low Resolution Electromagnetic Tomography (LORETA; Pascual-Marqui et al., 1999) was used to estimate intracerebral current density underlying the reward-related FRN following previously published procedures (e.g., Pizzagalli et al., 2002; see Pizzagalli, 2007 for a summary of LORETA core assumptions and prior validation findings). Current density was computed within a 200–280 ms post-feedback time window, which captured the mean peak latency of the FRN across frontocentral sites (274 ms). At each voxel ($n=2394$; voxel resolution: 7 mm^3), current density was computed as the linear, weighted sum of the scalp electric potentials (units are scaled to amperes per square meter, A/m^2). For each subject, LORETA values were normalized to a total power of 1 and then log-transformed before statistical analyses.

fMRI data

The imaging protocol has been described in detail in an independent study from our laboratory (Dillon et al., 2008). Briefly, fMRI data were acquired on a 1.5 T Symphony/Sonata scanner (Siemens Medical Systems; Iselin, NJ) using an optimized acquisition protocol (Deichmann et al., 2003). During functional imaging, gradient echo T2*-weighted echoplanar images were acquired using the following parameters: TR/TE: 2500/35 ms; FOV: 200 mm; matrix: 64×64; 35 slices; 222 volumes; voxels: 3.125×3.125×3 mm. A high-resolution T1-weighted MPRAGE structural volume was also collected for anatomical localization and extraction of structural regions-of-interest (ROIs) using standard parameters (TR/TE: 2730/3.31 ms; FOV: 256 mm; matrix: 192×192; 128 slices; voxels: 1.33×1.33×1 mm). Padding was used to minimize head movement.

Analyses were conducted using FS-FAST (http://surfer.nmr.mgh.harvard.edu) and FreeSurfer (Fischl et al., 2002, 2004). Functional pre-processing included motion and slice-time correction, removal of slow linear trends, intensity normalization, and spatial smoothing with a Gaussian filter (6 mm FWHM). A canonical hemodynamic response function (a gamma function) was convolved with stimulus onsets, and the general linear model was used to assess the fit between the model and the data. A temporal whitening filter was used to estimate and correct for autocorrelation in the noise. Participants with incremental (volume-to-volume) or cumulative head movement greater than 3.75 mm or degrees were removed from the analysis ($n=5$); for the remaining participants, motion parameters were included in the model as nuisance regressors. Of the subjects with usable ERP data, functional MRI data for 21 subjects were available and included in the statistical analyses.

Regression coefficients ("beta weights") indicating the fit of the model to the data were extracted from ROIs obtained from FreeSurfer's parcellation. For the purposes of the present study, we focused on data from four BG ROIs (nucleus accumbens, caudate, putamen, and globus pallidus), consistent with prior fMRI studies implicating BG regions in reward processing and reinforcement learning (e.g., Delgado et al., 2005; Dillon et al., 2008; Knutson and Cooper, 2005; Schonberg et al., 2007).

Statistical analyses

Test–retest reliability of behavioral data

The EEG session took place, on average, 39.30 days (S.D.: 23.88) after the initial behavioral pre-screening session. In a prior study using the same probabilistic reward task in an independent sample, we showed that the test–retest reliability for the reward-learning score (i.e., response bias block 3 minus response bias block 1) over a 38-day period was $r=0.57$ ($p<0.004$; Pizzagalli et al., 2005). In our prior study, 20 of the 24 participants were allocated to opposite keys for the rich stimulus. Thus, for participants allocated to a different bias across the two sessions, reward learning was used to estimate test–retest reliability. For participants allocated to the same bias, we did not expect a significant test–retest correlation when considering reward learning. Participants developing a strong response bias toward the long mouth in the first session, for example, were expected to show a robust response bias toward this stimulus already in block 1 of the second session, minimizing the amount of additional learning that could be achieved. Accordingly, for participants allocated to the same bias, the overall response bias (averaged across the 3 blocks) was used for test–retest computations.

In addition, in our prior study (Pizzagalli et al., 2005), we did not account for possible fluctuations in mood/affect between the two sessions. Because reward learning has been found to negatively correlate with anhedonic symptoms (e.g., Pizzagalli et al., 2005; Bogdan and Pizzagalli, 2006), fluctuations in mood across the two sessions might diminish the test–retest estimates. To this end, we also computed residualized reward-learning scores in which variance associated with anhedonic symptoms (MASQ AD subscore) was removed.

ERP data

For the primary analyses, mixed-model ANOVAs were used to analyze the FRN collapsed across blocks 2 and 3 with *Group* as a between-subject factor and *Site* (Fz, FCz) as a within-subject factor. Moreover, to examine the temporal characteristics of reward learning, a secondary mixed-model ANOVA was performed using *Group* as between-subject factor and *Learning Phase* (early: block 1 vs. late: blocks 2 and 3) as the within-subject factor. For the LORETA data, the groups were contrasted on a voxel-wise basis using unpaired t-tests comparing current density in response to rewarded rich trials at the time of the scalp FRN. Statistical maps were thresholded at $p<0.020$ with a minimum cluster size of 5 contiguous voxels (1.715 cm^3), and displayed on a standard MRI template. Pearson correlations were performed among behavioral, scalp ERP, LORETA, and fMRI data.

fMRI data

Although the MID task has been used to dissociate anticipatory versus consummatory phases of incentive processing

(e.g., Dillon et al., 2008), our interest in reward-related reinforcement learning led us to focus exclusively on responses to outcomes (gains versus no gains) on reward trials in the present study. For each participant, mean beta weights were extracted from the four BG ROIs for delivery of monetary gains (successful reward trial, or "win") and omission of potential gains (unsuccessful reward trial, or "no win") and entered into a *Group* (learners, $n=12$; non-learners, $n=9$)×*Region* (caudate, putamen, pallidus, nucleus accumbens)×*Outcome* (win, no win)×*Hemisphere* (left, right) mixed-model ANOVA.

Across the analyses of the behavioral, ERP, and fMRI data, the Greenhouse–Geisser correction was used when applicable. Significant ANOVA effects were follow-up by Newman–Keuls post-hoc tests.

Results

Demographic and behavioral data

Learners ($n=16$) and non-learners ($n=14$) did not differ with respect to age (21.38±2.01 vs. 21.51±4.51 years; $t(28)=$ 0.11, $p>0.90$), education (14.69±1.30 vs. 14.21±1.81 years; $t(28)=-0.83$, $p>0.40$), sex ratio (9 male/7 females vs. 8 males/6 females; $\chi^2(1)=0.002$, $p>0.90$), ethnicity (75% vs. 71.4% Caucasian; $\chi^2(2)=2.06$, $p>0.36$), employment status (87.5% vs. 85.7% undergraduate students; Fisher's exact test $p>0.39$), and length between the behavioral pre-screening and EEG session (36.06±24.06 vs. 43.00±23.45; $t(28)=-0.80$, $p>0.40$). One non-learner had a past history of major depressive episode, whereas 2 learners had a past history of subthreshold major depressive episode; no participants had received psychotropic medication in the past 6 months.

Replicating prior findings from an independent sample (Pizzagalli et al., 2005), non-learners reported higher anhedonic symptoms at the EEG session, as assessed by an anhedonic BDI-II subscore [loss of pleasure (item #4), loss of interest (item #12), loss of energy (item #15), and loss of interest in sex (item #21)] (0.68±0.82 vs. 0.25±0.45; $t(28)=$ 1.74, $p=0.049$, one-tailed) and MASQ anhedonic subscore (51.57±17.02 vs. 43.16±9.98; $t(28)=1.62$, $p=0.06$, one-tailed). Groups did not differ in general distress (MASQ General Distress Anxiety: 15.86±4.52 vs. 15.25±3.47; $t(28)=0.42$, $p=0.68$; General Distress Depression: 20.07±10.54 vs. 15.44± 2.73; $t(28)=1.62$, $p=0.12$) or anxiety symptoms (MASQ anxious arousal: 19.79±5.25 vs. 18.19±1.56; $t(28)=1.16$, $p=0.26$). Per design, learners had significantly higher response bias difference scores (block 3 – block 1) than non-learners (0.15±0.13 vs. −0.17±0.14; $t(28)=4.43$ $p<0.00001$).

For participants allocated to a different bias for the behavioral and EEG session, the test–retest correlation for reward learning over the two sessions was $r=0.50$ ($p=0.068$, $n=14$). When residualized reward-learning values were considered, in which variance associated with anhedonic symptoms (MASQ AD scores) was removed, the test–retest correlation was $r=0.56$ ($p<0.035$). As expected, for participants allocated to the same bias, overall response bias ($r=0.62$, $p<0.12$, $n=16$) but not reward learning ($r=-0.02$, $p>0.55$) was significantly correlated across the two sessions.

Scalp ERP data

The FRN was larger at FCz compared to Fz, as evident from a main effect of *Site*, $F(1,28)=10.56$, $p<0.004$, partial $\eta^2=0.37$. A

main effect of *Group* also emerged: as hypothesized, learners had significantly more positive FRNs to rich reward feedback than non-learners across sites, $F(1,28)=5.23$, $p<0.035$, partial $\eta^2=0.16$. Follow-up post-hoc Newman–Keuls tests confirmed that learners had more positive FRNs compared to non-learners at Fz (1.72±2.89 μV vs. −0.14±2.15 μV; $p<0.010$) and FCz (0.76±3.33 μV vs. −1.59±2.13 μV; $p<0.005$) (Fig. 1A). An ANOVA considering FRN values at Fz as a function of learning phase revealed a significant *Group* by *Learning Phase* interaction, $F(1,28)=4.29$, $p<0.050$, partial $\eta^2=0.13$. As shown in Fig. 1B, the FRN became more negative from early (block 1) to later phases (blocks 2 and 3) of the task for non-learners ($p<0.050$), whereas the FRN did not change for learners ($p>0.39$). Group differences emerged only for the later phase ($p<0.001$). An analogous ANOVA on FRN values at FCz revealed only a main effect of *Group*, $F(1,28)=4.20$, $p<0.05$, partial $\eta^2=0.15$; learners had a significantly more positive FRN than non-learners, particularly at later phases of learning (block 1: $p=0.051$; blocks 2 and 3: $p<0.002$). Finally, Pearson correlations confirmed that the amplitude of the FRN to rich reward
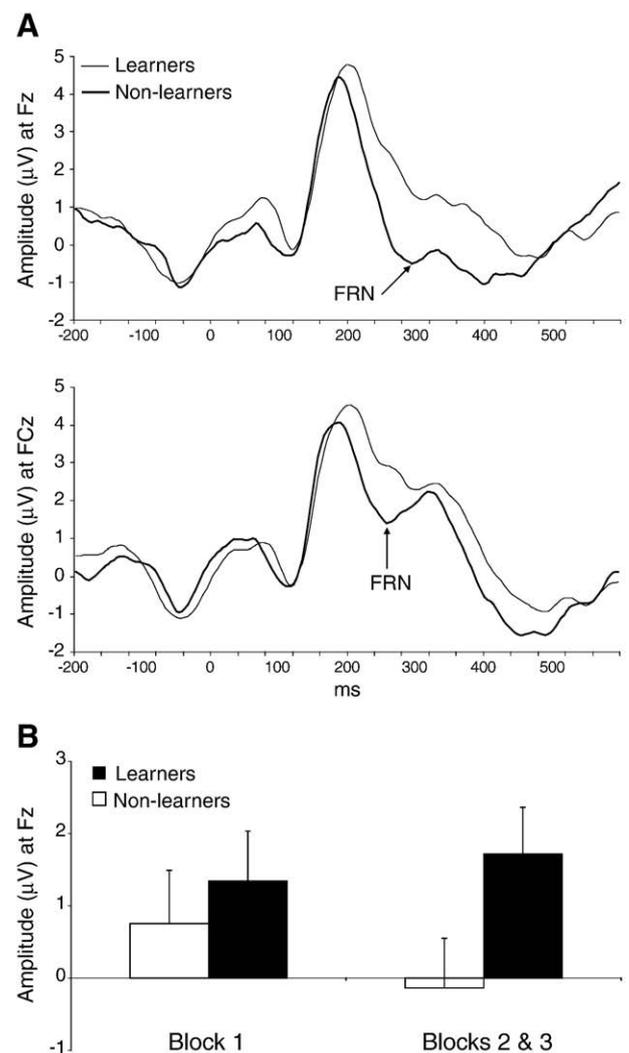
Fig. 1. (A) Averaged ERP waveforms at Fz and FCz from 200 ms before to 600 ms after the presentation of reward feedback for the rich stimulus for learners (light line) and non-learners (heavy line) in the probabilistic reward task; and (B) amplitude of the FRN at Fz during early (block 1) and late phases (blocks 2 and 3) of learning. Error bars refer to standard errors.

Table 1
Summary of significant results emerging from whole-brain LORETA analyses contrasting learners ($n=16$) and non-learners ($n=14$) in their response to reward feedback after correctly identifying the stimulus associated with more frequent reward in the probabilistic reward task
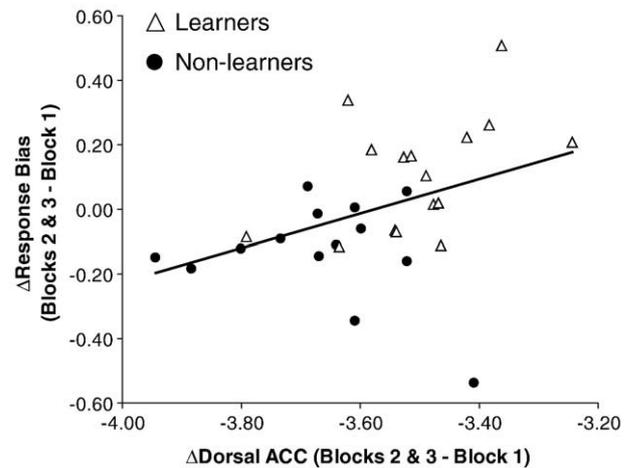
Region	MNI coordinates			Brodmann areas	Voxels	t-value	p-value
	X	Y	Z				
Dorsal anterior cingulate cortex	−3	17	22	24, 32, 33	7	2.769	.009
Posterior cingulate cortex	4	−46	15	29, 30, 31	34	−3.074	.005

The anatomical regions, MNI coordinates, and Brodmann areas of extreme t-values are listed. Positive t-values are indicative of stronger current density for the learners than non-learners, and vice versa for negative t-values. The numbers of voxels exceeding the statistical threshold are also reported ($p<0.02$; minimum cluster size: 5 voxels). Coordinates in mm (MNI space), origin at anterior commissure; (X)=left (−) to right (+); (Y)=posterior (−) to anterior (+); (Z)=inferior (−) to superior (+).

Fig. 3. Scatterplot and Pearson correlation between increases in dACC activation and response bias from early (block 1) to late phases (blocks 2 and 3) of learning. Relatively increased dACC current density in response to reward feedback for the rich stimulus is associated with greater reward learning ($r=0.40$, $p<0.030$). When the subject with the lowest reward learning was omitted, the correlation was $r=0.59$, $p<0.001$.

feedback correlated positively with differences in response bias over time [i.e., response bias (blocks 2 and 3 – block 1)] at Fz ($r=0.46$, $p<0.01$) and FCz ($r=0.35$, $p=0.06$), indicating that the positivity of the FRN is a reliable index of reward learning.

Source localization data

LORETA was used to estimate intracerebral current density underlying the FRN, specifically during blocks 2 and 3 compared to block 1. As hypothesized, learners showed relatively higher activity to rich reward feedback than non-learners in the dACC (Brodmann areas (BAs) 24, 32, 33; $t(28)=2.769$, $p<0.009$)
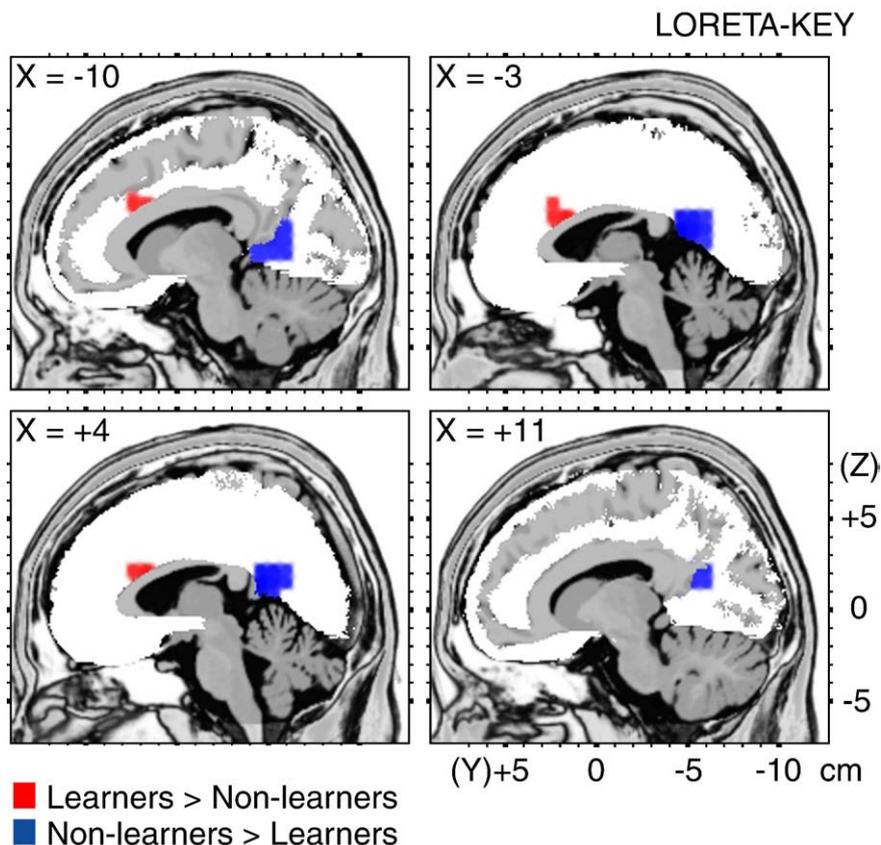
(Table 1, Fig. 2). By contrast, non-learners showed relatively higher activity in the posterior cingulate cortex (PCC; BAs 29, 30, 31; $t(28)=3.074$, $p<0.005$).

Inter-correlations among behavioral and ERP variables

Because the dACC is implicated in representing reinforcement histories to guide behavior (Amiez et al., 2006; Holroyd


Fig. 2. Results of voxel-by-voxel independent t-tests contrasting current density for the learners and non-learners in response to reward feedback for the rich stimulus on the probabilistic reward task. Red: relatively higher activity for learners. Blue: relatively higher activity for non-learners. Statistical map is thresholded at $p<0.020$ (minimum cluster size: 5 voxels) and displayed on the MNI template. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

and Coles, 2008; Kennerley et al., 2006), a positive correlation between dACC activation to reward feedback and the ability to develop a response bias was expected. As shown in Fig. 3, higher current density in the dACC region was indeed associated with greater reward learning [response bias (blocks 2 and 3 – block 1)] ($r=0.40$, $p<0.030$). Also, more positive FRNs were associated with higher current density in the dACC (Fz: $r=0.41$, $p<0.030$; FCz: $r=0.38$, $p<0.040$). In contrast, higher current density in the posterior cingulate was associated with poor reward learning ($r=-0.43$, $p<0.020$). No correlations emerged between PCC current density and FRNs.

fMRI data

No differences emerged between learners and non-learners with respect to the 15th (275.42 ± 35.09 ms vs. 267.89 ± 34.20 ms; $t(19)=0.49$, $p>0.62$) and 85th (382.50 ± 47.47 ms vs. 393.44 ± 69.14 ms; $t(19)=-0.43$, $p>0.65$) percentile RTs, which were used to titrate target duration for "unsuccessful" and "successful" trials, respectively. As in a prior study using this version of the MID task (Dillon et al., 2008), the differences between the short and long duration targets (learners: $\Delta=107.08$ ms; non-learners: $\Delta=125.55$ ms) were different enough to foster task engagement while being similar enough to elicit a comparable BOLD response.

In addition, a *Group* x *Trial Type* (reward, loss) ANOVA performed on the percentage of trials with a mismatch between RT and outcome revealed no significant effects (all $Fs<1.21$, all $ps>0.29$). Thus, no behavioral differences emerged between learners and non-learners during the MID task (% mismatched loss trials: 0.21 ± 0.08 vs. 0.19 ± 0.15; % mismatched reward trials: 0.20 ± 0.07 vs. 0.22 ± 0.13), indicating that fMRI findings were not confounded by group differences in performance during the MID task.
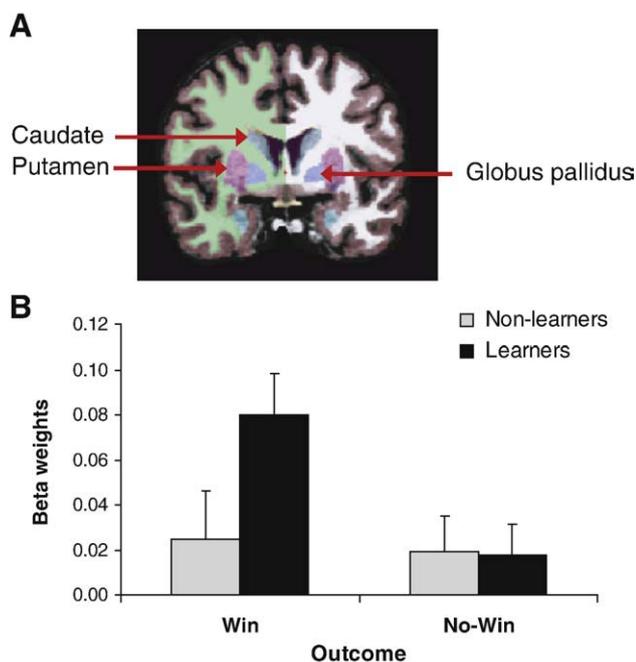
Fig. 4. (A) Parcellation of basal ganglia structures in a representative participant; only the caudate, putamen, and globus pallidus are shown in this coronal slice. (B) Mean beta weights (averaged across regions and hemispheres) in response to win feedback and no-win feedback in learners and non-learners (significant *Group×Outcome* interaction). Error bars refer to standard errors.

For the fMRI data, the *Group×Region×Outcome×Hemisphere* ANOVA revealed a main effect of *Condition*, $F(1,19)=9.36$, $p<0.007$, partial $\eta^2=0.33$, due to significantly higher activation following win than no-win feedback. More importantly, this effect was qualified by a significant *Group×Condition* interaction, $F(1,19)=6.57$, $p<0.02$, partial $\eta^2=0.26$. Neuman–Keuls post-hoc tests indicated that, as hypothesized, learners had significantly higher BG activation than non-learners in response to wins (0.080 ± 0.074 vs. 0.025 ± 0.045; $p<0.002$) but not no-win feedback (0.019 ± 0.039 vs. 0.018 ± 0.053; $p>0.91$). Moreover, learners ($p<0.004$) but not non-learners ($p>0.73$), had significantly higher activation to win compared to no-win feedback (Fig. 4). Although group differences for win feedback were significant for both hemispheres ($ps<0.0003$), the strongest differentiation was seen in the right hemisphere, as evident from a significant *Group×Condition×Hemisphere* interaction, $F(1,19)=4.68$, $p<0.045$, partial $\eta^2=0.20$. The only other effect to emerge was a significant *Region×Condition* interaction, $F(3,54)=10.02$, $p<0.001$, partial $\eta^2=0.35$, which was not explored further because it did not involve *Group*. No significant correlations emerged between (1) BG activation to wins, and (2) behavioral or ERP variables.

Discussion

This study investigated the contribution of the dACC to probabilistic reward learning in humans. As predicted, relative to non-learners, learners generated more positive FRNs and greater dACC activity in response to reward feedback following correct identification of the more frequently rewarded stimulus. Consistent with prior studies underscoring the sensitivity of FRN amplitude to learning (e.g., Muller et al., 2005), group differences were largest in later phases of the probabilistic reward task, by which time learners had established a robust response bias. Furthermore, FRN amplitude was positively correlated with current density in the dACC, and both FRN amplitude and dACC activation were positively correlated with reward learning. These correlations support the conclusion that dACC responses to reward feedback are a useful marker of reinforcement learning. Reward-related modulation of activity in the dACC is hypothesized to reflect a DA signal conveyed by the BG (Holroyd and Coles, 2002). Although the limitations of the electrophysiological technique precluded measuring BG activity during the probabilistic reward task, we found that relative to non-learners, learners showed a stronger BG response to rewarding outcomes in the MID task. Potentiated recruitment of BG regions in subjects developing a response bias toward the rich stimulus is consistent with the hypothesis that BG regions are critically implicated in feedback-based learning (Delgado, 2007; O'Doherty et al., 2004; Seymour et al., 2007). Collectively, the present findings extend a well-established model of human learning (Holroyd and Coles, 2002) into the domain of positive reinforcement, and highlight the importance of the human dACC in probabilistic reward learning.

The observation of relatively greater dACC activation in learners, as well as the relationship between dACC activation and reward learning, is consistent with emerging animal and neuroimaging evidence implicating the dACC in encoding reward probability and mediating the link between reinforcement history and upcoming behavior (Akitsuki et al., 2003; Amiez et al., 2006; Ernst et al., 2004; Ito et al., 2003; Nishijo et al., 1997; Shima and Tanji, 1998; Rushworth et al., 2007).

First, Ito et al. (2003) described dACC neurons that were particularly responsive to unexpected reinforcement; learners in the current study may have recruited this population of neurons, as only 40% of trials in the probabilistic reward task were rewarded. Second, Shima and Tanji (1998) identified a region of the rostral cingulate motor area (rCMA) that fired when monkeys voluntarily switched from one response to another in order to obtain greater reward, and this finding has been replicated in humans (Bush et al., 2002). The human homologue for rCMA is the anterior motor cingulate cortex (BA 24b; Vogt, 2005), thus these findings suggest that the human dACC – specifically, BA 24 – might play an important role in updating response selection based on reward feedback. Indeed BA 24 is the region identified by LORETA as more strongly activated by rewards in learners versus non-learners (Fig. 2). Third, Nishijo et al. (1997) identified dACC neurons that not only responded to rewarding objects but whose magnitude of response correlated with the monkey's object preferences. This result mirrors the present demonstration of a positive correlation between dACC activation and response bias.

Although the current results are consistent with findings highlighting the role of the dACC in using reward information to optimize behavior, we note that the positive relationship between FRN amplitude and dACC activation observed here appears inconsistent with an influential model of the FRN (Holroyd and Coles, 2002). The model proposes that the dACC is tonically inhibited by dopaminergic BG signals, such that when an event is worse than expected (negative prediction error), the resultant DA dip disinhibits the dACC and a relatively negative FRN is generated. The same model predicts that when events are better than expected (positive prediction error), the resultant DA burst will yield a more positive FRN (Holroyd and Coles, 2008). Although a relationship between the dACC and BG data presented here must be considered speculative given important differences between the probabilistic reward learning and MID tasks, this is essentially what was observed in the current study: learners, who showed a more vigorous BG response to unpredictable rewarding outcomes than non-learners (Fig. 4), also showed more positive FRNs (Fig. 1). However, along with more positive FRNs, learners also showed relatively greater dACC activation (Fig. 2). This seems to contradict the model (Holroyd and Coles, 2002), because although it is not explicitly stated that the relationship between DA bursts and more positive FRNs must be mediated by inhibition of the dACC, this seems logically implied by the fact that excitation of the dACC yields a more negative FRN.

We are not currently able to resolve this discrepancy, but it should be noted that we have observed this pattern of results previously. Using the same paradigm, we found that administration of a DA agonist (hypothesized to activate DA autoreceptors and thus decrease reward-related DA bursting) impaired reward learning and led to a more negative FRN along with *decreased* dACC activity (Santesso et al., unpublished). By contrast, participants who received a placebo demonstrated better reward learning, a more positive FRN, and greater dACC activity. Thus, in two studies examining probabilistic reward learning, we have observed positive correlations between dACC activity and FRN amplitude, rather than the negative correlation that has been described in situations when performance and/or outcomes are worse than expected (Holroyd and Coles, 2002). Future research will be needed to specify how the relationship

between DA signals, dACC activation, and scalp FRN differs for unpredicted negative vs. positive outcomes. Positive and negative prediction errors appear to be partially segregated to different regions of the striatum, with ventral anterior regions relatively more implicated in positive prediction errors and dorsal posterior striatal regions relatively more involved in negative prediction errors (Seymour et al., 2007); this raises the possibility that different sub-regions within the dACC may mediate FRNs to unpredicted positive vs. negative outcomes.

Relative to learners, non-learners unexpectedly showed relatively greater activation in the posterior cingulate cortex (PCC) in response to rewards. The PCC is connected with reward-related areas of the brain such as the ACC, medial PFC, and caudate nucleus (Vogt et al., 1992). Furthermore, PCC activity has been noted during the expectation and delivery of reward in monkeys (McCoy et al., 2003) and in response to positive compared with negative feedback in humans (Marco-Pallares et al., 2007; Nieuwenhuis et al., 2005), and implicated in the acquisition of response-outcome associations in rodents (Tabuchi et al., 2005). However, the reason for a stronger PCC response to rewards in non-learners versus learners is currently unclear and will require additional research.

The present study has five main limitations. First, negative feedback was not included in the probabilistic reward task. The FRN deflection is notably larger following negative versus positive feedback, and FRNs elicited by positive and negative feedback may be generated by distinct areas in the medial PFC/ACC (Nieuwenhuis et al., 2005); because our task involved only positive feedback, we could not test this hypothesis. Second, although we were able to investigate the spatio-temporal dynamics of brain mechanisms underling reinforcement learning with millisecond time resolution, we could not examine activity in subcortical regions (e.g., BG), or interactions between BG and cingulate regions, during the probabilistic reward task. Thus, while we show that relative to non-learners, learners demonstrated increased dACC and BG activation to reward feedback, it is important to emphasize that these data came from different tasks, only one of which (the probabilistic reward task) has a learning component. Although the disparate nature of the tasks might explain the lack of correlations between the EEG and fMRI data, we note that one of the strengths of the present study was our ability to show that non-learners were characterized by reduced activation in brain regions implicated in reinforcement learning (BG and ACC) in two rather distinct tasks, highlighting convergence and promising generalizability across the findings. Nevertheless, the implied relationship between the ERP and fMRI data is tentative and must be interpreted with caution. Third, while the LORETA algorithm has received important cross-modal validation (Pizzagalli, 2007), the spatial resolution of this source localization technique (1–2 cm) remains relatively coarse. Fourth, recent studies focusing on individual differences in reinforcement learning have provided compelling evidence that genetic variations affecting dopaminergic function can have profound influences on behavior (Frank et al., 2007) and brain activation (Klein et al., 2007), critically extending theoretical models of reinforcement learning. Unfortunately, for the present analyses, genetic information was not available. Finally, no data were collected about socioeconomical status — a variable that has been found to modulate monetary reward prediction error responses in a recent fMRI study (Tobler et al., 2007). However, among these 30 participants, 26 were Harvard undergraduate students (12

non-learners and 14 learners), 3 were graduate students, and one had graduated from college and was employed. Despite these similarities, results should be replicated with samples directly evaluated with respect to economic status.

Nonetheless, the present study provides important electrophysiological evidence of the critical role of the dACC in positive reinforcement learning in humans, and suggests that the differences in dACC activity in learners versus non-learners may be related to differences in the vigor of BG responses to rewards. The positive relationship between FRN amplitude and dACC activation is at odds with a prominent model of human reinforcement learning (Holroyd and Coles, 2002). Overall, however, the findings are consistent with two of the model's main hypotheses: (1) that phasic DA bursts act as signals that reinforce rewarding behaviors (Bayer and Glimcher, 2005; Garris et al., 1999), and (2) that these signals "teach" the dACC to select among various response options (Holroyd and Coles, 2002). Moreover, these results add to emerging evidence indicating that the dACC plays an important role in integrating reinforcement history over time to guide adaptive behavior (Amiez et al., 2006; Kennerley et al., 2006; Rushworth et al., 2007).

References

Akitsuki, Y., Sugiura, M., Watanabe, J., Yamashita, K., Sassa, Y., Awata, S., Matsuoka, H., Maeda, Y., Matsue, Y., Fukuda, H., Kawashima, R., 2003. Context-dependent cortical activation in response to financial reward and penalty: an event-related fMRI study. NeuroImage 19, 1674–1685.

Amiez, C., Joseph, J.P., Procyk, E., 2006. Reward encoding in the monkey anterior cingulate cortex. Cereb. Cortex 16, 1040–1055.

Bayer, H.M., Glimcher, P.W., 2005. Midbrain dopamine neurons encode a quantitative reward prediction error signal. Neuron 47, 129–141.

Beck, A.T., Steer, R.A., Brown, G.K., 1996. Beck Depression Inventory Manual, 2nd ed. The Psychological Corporation, San Antonio.

Bogdan, R., Pizzagalli, D.A., 2006. Acute stress reduces reward responsiveness: implications for depression. Biol. Psychiatry 60, 1147–1154.

Bush, G., Vogt, B.A., Holmes, J., Dale, A.M., Greve, D., Jenike, M.A., Rosen, B.R., 2002. Dorsal anterior cingulate cortex: a role in reward-based decision making. Proc. Natl. Acad. Sci. U S A 99, 523–528.

Chapman, L.J., Chapman, J.P., 1987. The measurement of handedness. Brain Cogn. 6, 175–183.

Delgado, M.R., 2007. Reward-related responses in the human striatum. Ann. N. Y. Acad. Sci. 1104, 70–88.

Delgado, M.R., Miller, M.M., Inati, S., Phelps, E.A., 2005. An fMRI study of reward-related probability learning. NeuroImage 24, 862–873.

Deichmann, R., Gottfried, J.A., Hutton, C., Turner, R., 2003. Optimized EPI for fMRI studies of the orbitofrontal cortex. NeuroImage 19, 430–441.

Dillon, D.G., Holmes, A.J., Jahn, A.L., Bogdan, R., Wald, L.L., Pizzagalli, D.A., 2008. Dissociation of neural regions associated with anticipatory versus consummatory phases of incentive processing. Psychophysiology 45, 46–59.

Ernst, M., Nelson, E.E., McClure, E.B., Monk, C.S., Munson, S., Eshel, N., Zarahn, E., Leibenluft, E., Zametkin, A., Towbin, K., Blair, J., Charney, D., Pine, D.S., 2004. Choice selection and reward anticipation: an fMRI study. Neuropsychologia 42, 1585–1597.

Fischl, B., Salat, D., Busa, E., Albert, M., Dieterich, M., Haselgrove, C., van der Kouwe, A., Killiany, R., Kennedy, D., Klaveness, S., Montillo, A., Makris, N., Rosen, B., Dale, A.M., 2002. Whole brain segmentation: Automated labeling of neuroanatomical structures in the human brain. Neuron 33, 341–355.

Fischl, B., van der Kouwe, A., Destrieux, C., Halgren, E., Segonne, F., Salat, D.H., Busa, E., Seidman, L.J., Goldstein, J., Kennedy, D., Caviness, V., Makris, N., Rosen, B., Dale, A.M., 2004. Automatically parcellating the human cerebral cortex. Cereb. Cortex 14, 11–22.

Frank, M.J., 2005. Dynamic dopamine modulation in the basal ganglia: a neurocomputational account of cognitive deficits in medicated and non-medicated Parkinsonism. J. Cogn. Neurosci. 17, 51–72.

Frank, M.J., Moustafa, A.A., Haughey, H.M., Curran, T., Hutchison, K.E., 2007. Genetic triple dissociation reveals multiple roles for dopamine in reinforcement learning. Proc. Nat. Acad. Sci. U.S.A. 104, 16311–16316.

Garris, P.A., Kilpatrick, M., Bunin, M.A., Michael, D., Walker, Q.D., Wightman, R.M., 1999. Dissociation of dopamine release in the nucleus accumbens from intracranial self-stimulation. Nature 398, 67–69.

Gehring, W.J., Willoughby, A.R., 2002. The medial frontal cortex and the rapid processing of monetary gains and losses. Science 295, 2279–2282.

Hajcak, G., Holroyd, C.B., Moser, J.S., Simons, R.F., 2005. Brain potentials associated with expected and unexpected good and bad outcomes. Psychophysiology 42, 161–170.

Hajcak, G., Moser, J.S., Holroyd, C.B., Simons, R.F., 2007. It's worse than you thought: The feedback negativity and violations of reward prediction in gambling tasks. Psychophysiology 44, 905–912.

Hautus, M.J., 1995. Corrections for extreme proportions and their biasing effects on estimated values of d'. Behav. Res. Methods, Instrum. Comput. 27, 46–51.

Holroyd, C.B., Coles, M.G.H., 2002. The neural basis of human error processing: Reinforcement learning, dopamine, and the error-related negativity. Psychol. Rev. 109, 679–709.

Holroyd, C.B., Coles, M.G.H., 2008. Dorsal anterior cingulate cortex integrates reinforcement history to guide voluntary behavior. Cortex 44, 548–559.

Holroyd, C.B., Krigolson, O.E., 2007. Reward prediction error signals associated with a modified time estimation task. Psychophysiology 44, 913–917.

Ito, S., Stuphorn, V., Brown, J.W., Schall, J.D., 2003. Performance monitoring by the anterior cingulate cortex during saccade countermanding. Science 302, 120–122.

Kennerley, S.W., Walton, M.E., Behrens, T.E.J., Buckley, M.J., Rushworth, M.F.S., 2006. Optimal decision making and the anterior cingulate cortex. Nat. Neurosci. 9, 940–947.

Klein, T.A., Neumann, J., Reuter, M., Hennig, J., von Cramon, D.Y., Ullsperger, M., 2007. Genetically determined differences in learning from errors. Science 318, 1642–1645.

Knutson, B., Cooper, J.C., 2005. Functional magnetic resonance imaging of reward prediction. Curr. Opin. Neurol. 18, 411–417.

Knutson, B., Fong, G.W., Bennett, S.M., Adams, C.M., Hommer, D., 2003. A region of mesial prefrontal cortex tracks monetarily rewarding outcomes: characterization with rapid event-related fMRI. NeuroImage 18, 263–272.

Makeig, S., Jun, T.P., Bell, A.J., Ghahremani, D., Sejnowski, T.J., 1997. Blind separation of auditory event-related brain responses into independent components. Proc. Natl. Acad. Sci. U. S. A. 94, 10979–10984.

Marco-Pallares, J., Muller, S.V., Munte, T.F., 2007. Learning by doing: An fMRI study of feedback-related brain activations. NeuroReport 18, 1423–1426.

McCoy, A.N., Crowley, J.C., Haghighian, G., Dean, H.L., Platt, M.L., 2003. Saccade reward signals in posterior cingulate cortex. Neuron 40, 1031–1040.

Miltner, W.H.R., Braun, C.H., Coles, M.G.H., 1997. Event-related brain potentials following incorrect feedback in a time-estimation task: evidence for a "generic" neural system for error detection. J. Cogn. Neurosci. 9, 788–798.

Montague, P.R., Hyman, S.E., Cohen, J.D., 2004. Computational roles for dopamine in behavioural control. Nature 431, 760–767.

Muller, S.V., Rodriguez-Fornells, A., Munte, T.F., 2005. Brain potentials related to self-generated information used for performance monitoring. Clin. Neurophysiol. 116, 63–74.

Nieuwenhuis, S., Slagter, H.A., von Geusau, A., Heslenfeld, D.J., Holroyd, C.B., 2005. Knowing good from bad: differential activation of human cortical areas by positive and negative outcomes. Eur. J. Neurosci. 21, 3161–3168.

Nishijo, H., Yamamoto, Y., Ono, T., Uwano, T., Yamashita, J., Yamashima, T., 1997. Single neuron responses in the monkey anterior cingulate cortex during visual discrimination. Neurosci. Lett. 227, 79–82.

O'Doherty, J., Dayan, P., Schultz, J., Deichmann, R., Friston, K., Dolan, R.J., 2004. Dissociable roles of ventral and dorsal striatum in instrumental conditioning. Science 304, 452–454.

Oliveira, F.T.P., McDonald, J.J., Goodman, D., 2007. Performance monitoring in the anterior cingulate is not all error related: expectancy deviation and the representation of action-outcome associations. J. Cogn. Neurosci. 19, 1–11.

Pascual-Marqui, R.D., Lehmann, D., Koenig, T., Kochi, K., Merlo, M.C., Hell, D., Koukkou, M., 1999. Low resolution brain electromagnetic tomography LORETA functional imaging in acute, neuroleptic-naive, first-episode, productive schizophrenia. Psychiatry Res.: Neuroimaging 90, 169–179.

Pizzagalli, D.A., 2007. Electroencephalography and high-density electrophysiological source localization, In: Cacioppo, J.T., Tassinary, L.G., Berntson, G.G. (Eds.), Handbook of Psychophysiology, 3rd Edition. Cambridge University Press, Cambridge, U.K., pp. 56–84.

Pizzagalli, D.A., Lehmann, D., Hendrick, A.M., Regard, M., Pascual-Marqui, R.D., Davidson, R.J., 2002. Affective judgments of faces modulate early activity (~160 ms) within the fusiform gyri. NeuroImage 16, 663–677.

Pizzagalli, D.A., Jahn, A.L., O'Shea, J.P., 2005. Toward an objective characterization of an anhedonic phenotype: a signal-detection approach. Biol. Psychiatry 57, 319–327.

Pizzagalli, D.A., Goetz, E., Ostacher, M., Iosifescu, D., Perlis, R.H., in press-a. Euthymic patients with Bipolar Disorder show decreased reward learning in a probabilistic reward task. Biol. Psychiatry. doi:10.1016/j.biopsych.2007.12.001.

Pizzagalli, D.A., Iosifescu, D., Hallett, L.A., Ratner, K.G., Fava, M., in press-b. Reduced hedonic capacity in Major Depressive Disorder: Evidence from a probabilistic reward task. J. Psychiatr. Res. doi:10.1016/j.jpsychires.2008.03.001.

Pizzagalli, D.A., Evins, A.E., Schetter, E., Frank, M.J., Pajtas, P.E., Santesso, D.L., Culhane, M., 2008. Single dose of a dopamine agonist impairs reinforcement learning in humans: behavioral evidence from a laboratory-based measure of reward responsiveness. Psychopharmacology 196, 221–232.

Rushworth, M.F., Buckley, M.J., Behrens, T.E., Walton, M.E., Bannerman, D.M., 2007. Functional organization of the medial frontal cortex. Curr. Opin. Neurobiol. 17, 220–227.

Schonberg, T., Daw, N.D., Joel, D., O'Doherty, J.P., 2007. Reinforcement learning signals in the human striatum distinguish learners from nonlearners during reward-based decision making. J. Neurosci. 27, 12860–12867.

Schultz, W., 2007. Behavioral dopamine signals. Trends Neurosci. 30, 203–210.

Seymour, B., Daw, N., Dayan, P., Singe`r, T., Dolan, R., 2007. Differential encoding of losses and gains in the human striatum. J. Neurosci. 27, 4826–4831.

Shima, K., Tanji, J., 1998. Role for cingulate motor area cells in voluntary movement selection based on reward. Science 282, 1335–1338.

Tabuchi, E., Furusawa, A.A., Hori, E., Umeno, K., Ono, T., Nishijo, H., 2005. Neural correlates to action and rewards in the rat posterior cingulate cortex. NeuroReport 16, 949–953.

Tobler, P.N., Fletcher, P.C., Bullmore, E.T., Schultz, W., 2007. Learning-related human brain activations reflecting individual finances. Neuron 54, 167–175.

Tripp, G., Alsop, B., 1999. Sensitivity to reward frequency in boys with attention deficit hyperactivity disorder. J. Clin. Child Psychol. 28, 366–375.

Van Veen, V., Holroyd, C.B., Cohen, J.D., Stenger, V.A., Carter, C.S., 2004. Errors without conflict: Implications for performance monitoring theories of anterior cingulate cortex. Brain Cogn. 56, 267–276.

Vogt, B.A., 2005. Pain and emotion interactions in subregions of the cingulate gyrus. Nat. Rev. Neurosci. 6, 533–544.

Vogt, B.A., Finch, D.M., Olsen, C.R., 1992. Functional heterogeneity in cingulate cortex: The anterior executive and posterior evaluative regions. Cereb. Cortex 2, 435–443.

Wager, T.D., Nichols, T.E., 2003. Optimization of experimental design in fMRI: a general framework using a genetic algorithm. Neuroimage 18, 293–309.

Watson, D., Weber, K., Assenheimer, J.S., Clark, L.A., Strauss, M.E., McCormick, R.A., 1995. Testing a tripartite model: I. Evaluating the convergent and discriminant validity of anxiety and depression symptom scales. J. Abnorm. Psychol. 104, 3–14.