Human Midbrain Sensitivity to Cognitive Feedback and Uncertainty During Classification Learning

A. R. Aron,[1] D. Shohamy,[2] J. Clark,[3] C. Myers,[4] M. A. Gluck,[2] and R. A. Poldrack[1,3]

[1]*Department of Psychology and Brain Research Institute, University of California, Los Angeles, California 90065;* [2]*Center for Molecular and Behavioral Neuroscience and* [4]*Department of Psychology, Rutgers University, New Brunswick, New Jersey 08903; and* [3]*Martinos Center for Biomedical Imaging, Massachusetts General Hospital, Massachusetts 02129*

Aron, A. R., D. Shohamy, J. Clark, C. Myers, M. A. Gluck, and R. A. Poldrack. Human midbrain sensitivity to cognitive feedback and uncertainty during classification learning. *J Neurophysiol* 92: 1144–1152, 2004. First published March 10, 2004; 10.1152/jn.01209. 2003. Mesencephalic dopaminergic system (MDS) neurons may participate in learning by providing a prediction error signal to their targets, which include ventral striatal, orbital, and medial frontal regions, as well as by showing sensitivity to the degree of uncertainty associated with individual stimuli. We investigated the mechanisms of probabilistic classification learning in humans using functional magnetic resonance imaging to examine the effects of feedback and uncertainty. The design was optimized for separating neural responses to stimulus, delay, and negative and positive feedback components. Compared with fixation, stimulus and feedback activated brain regions consistent with the MDS, whereas the delay period did not. Midbrain activity was significantly different for negative versus positive feedback (consistent with coding of the "prediction error") and was reliably correlated with the degree of uncertainty as well as with activity in MDS target regions. Purely cognitive feedback apparently engages the same regions as rewarding stimuli, consistent with a broader characterization of this network.

INTRODUCTION

The means by which the brain responds to rewarding events has been a subject of intense investigation, focused particularly on the mesencephalic dopaminergic system (MDS) in the midbrain. Neurophysiological investigations have shown that midbrain dopaminergic neurons of the substantia nigra pars compacta (SN) and ventral tegmental area (VTA) fire in response to salient or rewarding events. Recent work has provided compelling evidence for the hypothesis that the phasic firing of these neurons codes for errors in reward prediction, which indexes the degree to which behavior should be altered (Waelti et al. 2001; and see for review Schultz and Dickinson 2000). However, other work has suggested that dopamine neurons may fire more generally in response to behaviorally salient events, even in the absence of reward (Horvitz 2000; Redgrave et al. 1999). In humans, the primary target regions of these dopaminergic systems, particularly the ventral striatum, orbital, and medial frontal cortex, have also been identified as playing a role in the processing of diverse kinds of rewarding stimuli, including cocaine (Breiter et al. 1997), money (Delgado et al. 2000; Elliott et al. 2003; Gehring and Willoughby 2002; Knutson et al. 2001a), taste rewards (Berns et al. 2001; McClure et al. 2003; O'Doherty et al. 2003; Pagnoni et al. 2002), and beauty (Aharon et al. 2001).

One important question is whether the MDS and its targets are also activated by tasks providing purely cognitive feedback without primary or secondary rewards, such as feedback-driven classification learning. In such settings, subjects learn to sort stimuli into two or more categories by observing outcomes over many trials. Typically the outcome is probabilistic to prevent subjects from adopting declarative strategies, instead forcing them to rely on gradually acquired stimulus-outcome associations. Demonstrating that learning on the basis of purely cognitive feedback *does* engage the MDS and its targets would have implications for understanding the neural basis of fundamental aspects of human cognition related to classification such as categorization and concept formation (Estes 1994). Neuroimaging of probabilistic classification learning suggests the striatum and midbrain are significantly activated compared with baseline and significantly more so when subjects learn based on response-contingent feedback versus learning the same materials without feedback (Poldrack et al. 2001). Moreover patients with Parkinson's disease—a basal ganglia disorder—are impaired relative to controls on a feedback-based classification task but normal on a nonfeedback version (Shohamy et al. 2004). One characterization of these findings—which we test here—may be that the MDS and its targets participate in a wider network by representing expectancies and adjusting them on the basis of feedback, i.e., coding the prediction error that drives learning (Schultz 2002; Schultz and Dickinson 2000).

The MDS may also be modulated by uncertainty regarding stimulus-outcome associations. A neurophysiological study found that sustained activity of dopamine neurons during the delay period in an associative-learning task was modulated by the uncertainty of the stimulus-reward relation with greatest activity under greatest uncertainty (Fiorillo et al. 2003). Uncertainty is of particular interest because it is a central concept in associative-learning theories (Dayan et al. 2000; Pearce and Hall 1980; Schultz and Dickinson 2000). In humans, temporally unpredicted delivery of reward has been shown (relative to predictable reward) to activate ventral striatum as well as orbitofrontal cortex in some cases (Berns et al. 2001; McClure et al. 2003; O'Doherty et al. 2003; Pagnoni et al. 2002). A recent functional MRI (fMRI) study established widespread activation of the midbrain and MDS targets when subjects had

0022-3077/04 $5.00 Copyright © 2004 The American Physiological Society

to make an uncertain compared with a certain prediction (Volz et al. 2003). However, in that study, it was not possible to separate stimulus from feedback-related activation nor to parametrically examine the role of uncertainty.

We studied how the brain's reward systems respond to expectancy, feedback, and uncertainty by examining neural activity using fMRI while subjects learned a probabilistic classification task using trial-by-trial feedback. Unlike previous research treating entire classification-learning trials as single events, the current study used rapid-presentation event-related fMRI to separate the activity associated with stimulus, delay, and feedback components of each trial (Fig. 1). When non-human primates learn to correctly perform stimulus-reward tasks, the phasic dopaminergic response shifts backward in time from the occurrence of the reward to the stimulus that predicts reward (Mirenowicz and Schultz 1994). If a similar mechanism applies in human classification learning, there should be significantly different MDS activation between stimulus and positive feedback events. Moreover, as subjects usually reach high classification accuracy early in the course of such experiments, the MDS response to negative feedback (large error signal; mismatch between prediction and outcome) should be significantly different compared with positive feedback (small error signal; match between prediction and outcome). The probabilistic design also allowed us to investigate a role for the midbrain in processing uncertainty because each of the 14 stimulus combinations in the experiment was differentially associated with the potential outcomes and therefore contained a different average level of information. If it is the case that the midbrain codes uncertainty (Fiorillo et al. 2003), then there should be significant correlations between activity in that region and increasing uncertainty. Finally, we investigated functional connectivity between the midbrain and other brain regions to characterize the wider network associated with midbrain function.
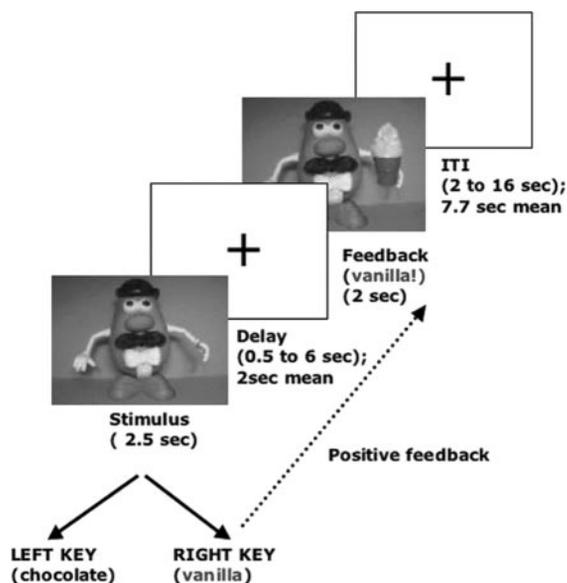

FIG. 1. A single positive-feedback trial. The Mr. Potato Head stimulus is presented with hat, moustache, and bowtie features for 2.5 s, and the subject responds within that time with a left or right button press (indicating, respectively, predictions for chocolate or vanilla outcomes). There is a delay period (fixation cross). Feedback follows; Mr. Potato Head is shown with vanilla ice cream. Inter-trial interval (ITI).

METHODS

Subjects

Twenty right-handed healthy English-speaking subjects participated (7 males; age range: 20–33). All subjects gave informed consent according to a Massachussetts General Hospital Human Subjects Committee protocol. Four subjects were excluded from fMRI data analysis because they did not successfully learn to perform the task (mean accuracy: <60% over the last 100 trials). One additional subject was excluded due to extensive signal dropout, leaving 15 subjects in the analyzed dataset.

Imaging procedures

Imaging was performed using a 3.0T Siemens Allegra MR scanner. Blood-oxygen-level-dependent (BOLD)-sensitive functional images were collected using a gradient-echo echo-planar pulse sequence (TR = 2,000 ms, TE = 30 ms, 64 × 64 matrix, 200-mm field of view, 21 slices, 5-mm-thick, 1-mm interslice gap). Six runs of fMRI scanning were performed for each subject, lasting 384 s each, with an additional four images at the beginning of the run discarded to allow T1 equilibration.

Task and stimuli

Subjects participated in a feedback-driven classification-learning task (Knowlton et al. 1996; Poldrack et al. 2001) consisting of six blocks of 25 trials each. Each subject was instructed to pretend that they were working in an ice-cream store and that they were to learn to predict which individual figures preferred vanilla or chocolate ice cream. On each trial, the subject was presented with a toy figure (Mr. Potato Head, Playschool/Hasbro) with a subset of four features (hat, eyeglasses, moustache, and bow tie). Stimulus presentation lasted 2.5 s, within which time the subject responded with a left button press for chocolate or a right button press for vanilla. There followed a variable interval of visual fixation (0.5–6 s, mean = 2 s; sampled randomly from an exponential distribution), after which feedback was presented (2 s) by showing the stimulus figure holding either a vanilla or chocolate ice cream cone (Fig. 1). The interval between the feedback offset and onset of the next stimulus varied between 2 and 16 s (mean = 7.7 s; also randomly sampled from an exponential distribution). The interstimulus-interval length was determined by jointly minimizing the correlation between stimulus- and feedback-evoked responses and optimizing the efficiency of the design for differentiation between these two classes of responses (Dale 1999). Category labels (chocolate/vanilla) were probabilistically associated with feature combinations (Table 1). As there were four features, and these could occur in any combination with at least one feature being present on every Mr. Potato Head (but never all 4), there were 14 combinations of features (stimuli). The cue strengths between individual features and the chocolate outcome were chosen such that the probability of being associated with chocolate was 0.85, 0.66, 0.44, and 0.24 for the different features. A combination of features constituted a particular "stimulus," and the probability that a stimulus was associated with the chocolate outcome, and the frequency with which this occurred (of 150 trials total), are shown in Table 1. Entropy values are also shown, computed using the formula in the following text (see Entropy analysis).

Data analysis

Behavioral data were analyzed for accuracy and reaction time (RT). A response was judged correct if the probability of its matching the feedback (over all 150 trials) was >0.5. Stimuli that were equally associated with both outcomes [i.e., P(Chocolate) = 0.5; see Table 1], were excluded from the accuracy analysis. No subject failed to respond on more than one trial, and around half the subjects missed no trials at all. Given the small number of misses, these trials were not

TABLE 1. *Relation among stimuli, features, and outcomes*

Stimulus	Moustache	Hat	Glasses	Bow Tie	P(Choc)	Frequency	Entropy
1	0	0	0	1	0.94	16	0.33
2	0	1	0	0	0.75	8	0.81
3	0	1	0	1	0.91	22	0.44
4	0	0	1	0	0.12	8	0.53
5	0	0	1	1	0.88	8	0.53
6	0	1	1	0	0.5	4	1.00
7	0	1	1	1	0.94	16	0.33
8	1	0	0	0	0.07	15	0.37
9	1	0	0	1	0.5	6	1.00
10	1	1	0	0	0.17	6	0.66
11	1	1	0	1	0.71	7	0.87
12	1	0	1	0	0.07	15	0.37
13	1	0	1	1	0.6	5	0.97
14	1	1	1	0	0.14	14	0.58

There were 14 stimuli, consisting of one to three of the following features: moustache, hat, glasses, and bow tie. During the course of the experiment (150 trials), each stimulus had a particular frequency, was associated with chocolate outcome with a particular probability P(Choc), and had an entropy value between 0 and 1 (0 indicates 0 uncertainty, 1 indicates maximum uncertainty).

separately coded in the imaging analysis. Preprocessing and statistical analysis of the fMRI data were performed using SPM99 software (Wellcome Dept. of Cognitive Neurology, London, UK) and included slice timing correction, motion correction, spatial normalization to the Montreal Neurological Insititue 305 (MNI305) stereotactic space (using linear affine registration followed by nonlinear registration, resampling to 3-mm cubic voxels), and spatial smoothing with an 8 mm Gaussian kernel. Stimulus, positive feedback, negative feedback, and delay (a box-car starting at stimulus-offset and lasting the delay duration) were modeled using the canonical hemodynamic response function and its temporal derivative. Low-frequency signal components (66-s cutoff) were treated as confounding covariates. The model-fit was performed individually for each subject. Contrast images were generated for each of the four event types (against the explicitly unmodelled baseline/fixation) as well as for contrasts between event types. The contrast images were then used in a second-level analysis treating subject as a random effect.

Validation

Because of the correlation between stimulus and feedback events, a validation was performed to determine whether the estimate of feedback activation would be contaminated by preceding stimulus-evoked activation. A synthetic dataset was created using the actual event timings from a single subject in the study, injecting signal for each stimulus and feedback event (according to a canonical hemodynamic response) into a large region of interest (along with Gaussian noise). The signal injected for each stimulus event was 1/2 the size of the signal injected for each feedback event. The data were analyzed using the methods outlined in the preceding text, using a set of finite impulse response basis functions (instead of the canonical hemodynamic response function used for the statistical analyses). The analysis was validated by comparing the mean signal estimates (beta weights) at the maximum of the hemodynamic response for each event type. The beta weights were accurately estimated for each event type (1.9921 estimated ratio of mean feedback to stimulus activity vs. 2.0 actual ratio). In addition, the estimates for negative versus positive feedback differed by only 0.0099%, suggesting that estimated differences between positive and negative feedback events were not artifacts of the design or analysis.

Entropy analysis

The average information contained by a particular stimulus over the course of the experiment was represented by an entropy value on each trial according to the formula

$$E = -x * \log_2(x) - [(1 - x) * \log_2(1 - x)]$$

where x was the probability of the chocolate outcome given that particular stimulus over the course of the experiment (Table 1). A new model was fit for each subject including entropy as a parametric regressor for each trial component and using the canonical hemodynamic response function and its temporal derivative. Contrast images were generated for each event type, reflecting the relationship between BOLD activity and entropy at each voxel, and these were used in a second-level analysis treating subject as a random effect. As the entropy analysis was motivated by the finding that uncertainty related to firing of dopaminergic neurons in the monkey midbrain (Fiorillo et al. 2003), voxel responses were analyzed within a spherical midbrain region of interest (ROI), radius 15 mm, centered at MNI coordinates 0, −15, −9 [x, y, z], according to an anatomical atlas (Lucerna et al. 2002). An ROI of this size encompasses the entire midbrain, including SN, VTA and other structures. Although fMRI (using an 8-mm filter) cannot distinguish these structures, we assumed, given our a priori hypothesis concerning uncertainty, that any BOLD response within this region was likely to relate to SN or VTA activity.

Functional connectivity analysis

For each subject, for the comparison of negative feedback with fixation, average activity was computed from the same midbrain ROI as in the preceding text (radius: 15 mm, centered at MNI coordinates 0, −15, −9 [x, y, z]). A random-effects correlational analysis was performed using average midbrain activity as a regressor; thus assessing which brain regions correlated reliably with activity for this ROI.

Learning analysis

For each of the reliable foci reported for the comparisons of stimulus and positive feedback and negative feedback and positive feedback, we used the SPM toolbox (http://spm-toolbox.sourceforge.net/) to investigate how average activation changed over scanning session.

RESULTS

Behavioral results

A large increase in classification accuracy occurred between blocks 1 and 2 with a slight decrease in block 3 and relatively stable performance in the remaining blocks (Fig. 2). Accuracy was analyzed using ANOVA with block as a repeated measure.
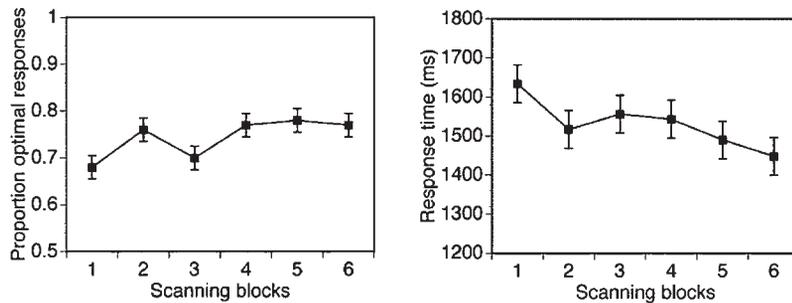
FIG. 2. Behavioral data. Averaged over subjects, a large increase in classification accuracy occurred between blocks 1 and 2 with a decrease for block 3 and stable performance for the remainder. Response times decreased across the scanning blocks.

There was a significant effect of block, $F(5,75) = 2.714$, $P = 0.032$ (a test for linear trend was marginally significant, $P = 0.079$). RT decreased with learning on the task, with ANOVA showing a marginally significant effect of block, $F(5,75) = 2.087$, $P = 0.091$ (a test for linear trend was also marginally significant, $P = 0.072$).

Functional imaging results

Except where indicated, the statistical significance of reported fMRI activations survived correction for multiple comparisons at the wholebrain level, using the false discovery rate (FDR) correction (Genovese et al. 2002). The FDR procedure ensures that on average no more than 5% of activated voxels for each contrast are expected to be false positives.

STIMULUS VERSUS FIXATION. Significant activation was found in an extensive set of bilateral cortical regions, including prefrontal (premotor, inferior and middle frontal gyri, lateral orbital, and anterior cingulate), anterior insula, intraparietal, and occipital cortices (Fig. 3A). In addition, activity was observed bilaterally in a number of subcortical regions, includ-

ing thalamus, caudate/putamen, ventral striatum/nucleus accumbens, globus pallidus, midbrain (including SN/VTA), and cerebellum (superior and inferior cerebellar cortex and vermis). Significant negative activation was observed in the medial prefrontal cortex, precuneus, medial temporal lobe, anterior hippocampus, and inferior parietal cortex.

DELAY VERSUS FIXATION. There was significant activation in the right inferior frontal cortex, caudate nucleus, parietal cortex, and cerebellum (Fig. 3B). Significant de-activation was observed in the medial prefrontal, medial temporal, and parietal cortex.

POSITIVE-FEEDBACK VERSUS FIXATION AND NEGATIVE-FEEDBACK VERSUS FIXATION. Significant feedback activation [positive feedback vs. fixation (Fig. 3C) and negative feedback vs. fixation (Fig. 3D)] was observed in widespread areas similar to those activated by the stimulus, including bilateral orbitofrontal, right inferior and middle frontal, occipital, and parietal cortical regions and thalamus, caudate/putamen, ventral striatum/nucleus accumbens, globus pallidus, midbrain (including SN/VTA), and cerebellum. Significant de-activation was ob-
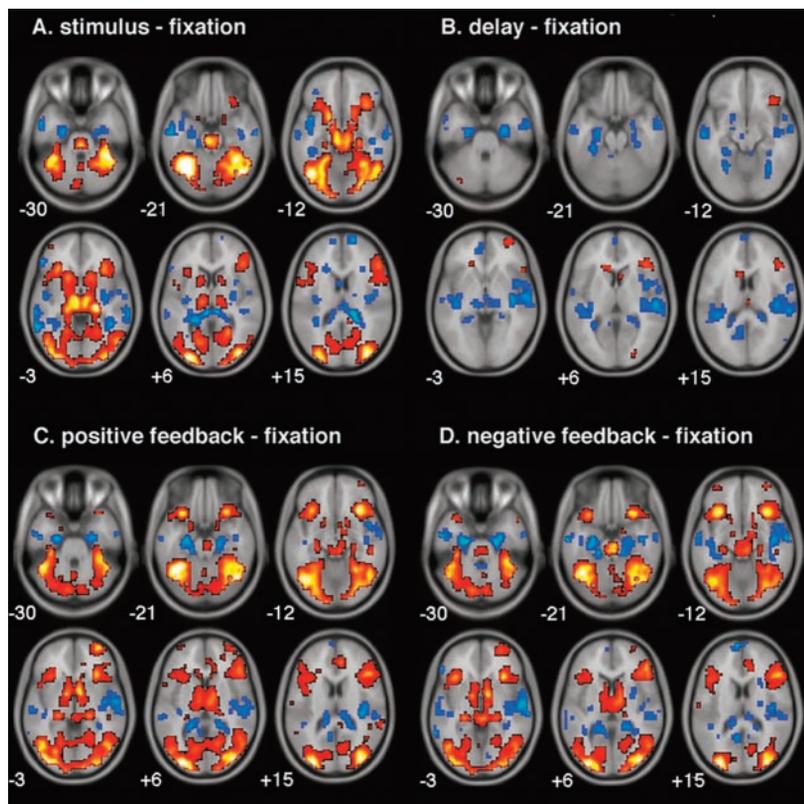

FIG. 3. Subcortical and cortical activations (red-scale) and deactivations (blue-scale) shown on axial slices for the four main contrasts. A: stimulus minus fixation produces occipital, cerebellar, midbrain, basal ganglia, and orbitofrontal activation, whereas deactivations occur for the medial temporal lobe (MTL). B: delay minus fixation produces right frontal, cerebellar, and caudate activation and deactivations in MTL. The positive (C) and negative (D) feedback events produce occipital, cerebellar, midbrain, basal-ganglia, orbital, and inferior frontal activations and deactivations of the MTL. All maps are corrected for multiple comparisons [false dectection rate (FDR) correction, $P < 0.05$].
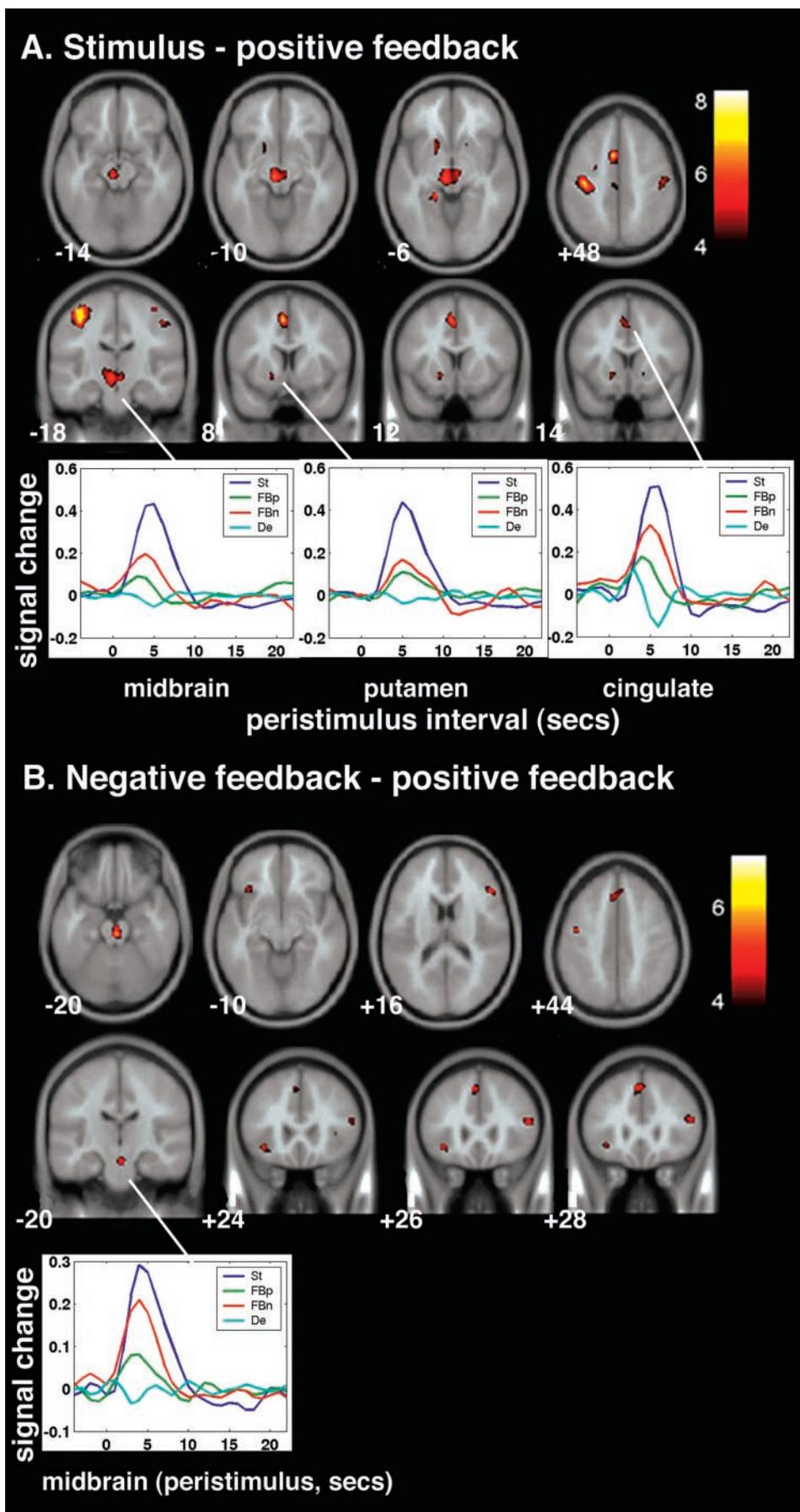
FIG. 4. Midbrain activity codes prediction error. *A*: the stimulus event leads to significantly more activation than positive feedback for the midbrain region, left motor/premotor cortex, posterior anterior cingulate, and ventral striatum (FDR corrected for whole brain, $P < 0.05$). Peristimulus plots show activity change at midbrain and MDS target foci for the 4 event types. *B*: negative feedback significantly activates the midbrain region more than positive feedback (whole-brain cluster correction according to Gaussian random field theory, $P < 0.05$). Medial and orbital frontal foci are also evident at a lower, nonsignificant, threshold ($P < 0.001$). A peristimulus plot showing activity change for the different event types is shown for the midbrain focus alone. St, stimulus; FBp, positive feedback; FBn, negative feedback; De, delay period.

TABLE 2. *Activations for the contrast (stimulus − positive feedback)*

Description	Cluster Size	t Value	x	y	z
Cingulate gyrus	100	8.29	−6	9	51
Calcerine sulcus	18	8.21	−18	−39	−3
Pre-motor/motor cortex	272	7.53	−36	−18	54
Cerebellum	59	7.04	18	−57	−27
Midbrain (SN/VTA)	163	6.60	−6	−21	−9
Precentral gyrus	21	5.44	39	−15	63
Parieto-temporal fissure	23	5.39	−15	−63	21
Precentral gyrus	26	5.11	51	−15	48
Putamen	24	4.88	−15	12	−6

All foci are false discovery rate (FDR) corrected for wholebrain multiple comparisons ($P < 0.05$). Only foci with cluster size >10 are reported, along with MNI coordinates. For the x dimension, negative denotes left hemisphere. SN, substantia nigra pass compacta; VTA, ventral tegmental area.

served in superior frontal, medial temporal, inferior parietal, and occipital regions.

STIMULUS VERSUS NEGATIVE-FEEDBACK AND STIMULUS VERSUS POSITIVE-FEEDBACK. At no focus was there significantly more activity for stimulus than negative feedback, but there was significantly more activity for stimulus than positive feedback in the midbrain, ventral striatum, left motor/premotor cortex, and cingulate gyrus, among other regions (Fig. 4*A*, Table 2). There were no foci at which activity was reliably greater for positive feedback than stimulus. Although the finding of reliable midbrain and striatal activation change for stimulus compared with positive feedback supported the idea that the MDS and its targets may code predictions (Mirenowicz and Schultz 1994; O'Doherty et al. 2003), this particular contrast was beset by two potential confounds. First, the stimulus event included both stimulus processing and the motor response, whereas positive feedback only included stimulus processing. The fact that left motor/premotor cortex was activated for this contrast, when subjects made right-handed responses, suggests that activation of this focus (at least) may have been related to motor execution. Second, the stimulus event represented the first viewing of a stimulus, whereas the positive feedback event represented the second viewing, resulting in a novelty/saliency confound. Such considerations motivated the following, unconfounded, contrast.

NEGATIVE-FEEDBACK VERSUS POSITIVE-FEEDBACK. A critical contrast for identifying the neural correlates of the prediction error is that between negative and positive feedback. For this contrast, no foci survived a whole-brain FDR correction. However, using a weaker threshold ($P < 0.001$), midbrain, bilateral orbital frontal, and medial frontal foci were more activated for negative than positive feedback. Only the midbrain focus (peak MNI coordinates: 3, −27, −21 [x, y, z], and ranging from z = −21 to z = −14) survived an alternative correction for multiple comparisons (a cluster-level correction based on Gaussian random field theory), which is sensitive to different aspects of the fMRI signal than FDR (e.g., broader clusters with lower peak activations; Fig. 4*B*, Table 3). There were no foci at which activation was significantly greater for positive feedback than negative feedback even at an uncorrected threshold of $P < 0.001$. We further assessed midbrain activation with respect to the midbrain ROI used for the entropy and functional

TABLE 3. *Activations for the contrast (negative feedback − positive feedback)*

Description	Cluster Size	t Value	x	y	z
Midbrain	49	7.11**	3	−27	−21
Inferior frontal juncture	22	5.37	−45	−6	45
Medial frontal cortex	39	5.90	0	30	48
Inferior frontal gyrus	23	5.47	51	27	15
Orbital frontal gyrus	20	5.05	−36	24	−9

Only the midbrain focus (**) was reliable ($P < 0.05$: SPM99 implementation of Gaussian random field theory whole-brain correction).

connectivity analyses (see following text), by performing a confirmatory negative versus positive feedback analysis within the predefined midbrain ROI (radius: 15 mm, centered at MNI coordinates 0, −15, −9 [x, y, z]). Two clusters (centered at: 0, −24, −21; t = 5.46 and 3, −15, −15; t = 5.37) showed significantly greater activation for negative than positive feedback according to multiple comparison corrections within this ROI using both Gaussian random field theory family-wise correction ($P < 0.05$) and the FDR method ($P < 0.0001$).

UNCERTAINTY (CORRELATING INCREASING BOLD WITH INCREASING ENTROPY). Uncertainty was defined in terms of entropy for each stimulus pattern across all trials (see METHODS). Because of the previous results of Fiorillo et al. (2003), we focused on an ROI in the midbrain (radius: 15 mm, centered MNI coordinates 0, −15, −9 [x, y, z]). There was a significant correlation between increasing midbrain activity (at focus 12, –15, –12) and increasing entropy during the delay period [$t(1,14) = 3.5$, $P < 0.05$, FDR corrected for midbrain ROI; Fig. 5]. No activation foci survived a whole-brain correction using FDR. There were no reliable inverse correlations between entropy and midbrain activity. However, as entropy was confounded with *novelty* of visual stimulus—i.e., those stimuli that occurred infrequently had higher entropy (see Table 1)—we also examined the midbrain ROI for voxels where activity for the stimulus event increased with entropy. The rationale for this analysis was that if it were novelty rather than entropy that drove the increasing activity, then this should also (and particularly) be the case for the stimulus event. There were, however, no significant activations for this analysis (at $P < 0.05$, FDR corrected). Moreover, a paired *t*-test showed that activity for
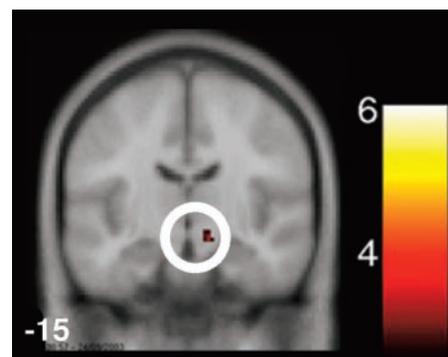

FIG. 5. Midbrain activity correlates with uncertainty. A midbrain region of interest (ROI) analysis, centered on the substantia nigra (sphere of 15-mm radius), motivated a priori by the findings of Fiorillo et al. (2003), reveals a significant relationship between increasing BOLD activation and increasing entropy for the delay period (FDR corrected within ROI, $P < 0.05$).

the entropy modulation of the delay event [cluster 12, −15, −12] was greater than for the entropy modulation of the stimulus event ($t = 2.7$; although this interaction was no longer significant when correcting for multiple comparisons). We therefore argue that activity in this region reflects the coding of uncertainty rather than stimulus novelty, in accordance with the findings of Fiorillo et al. (2003).

FUNCTIONAL CONNECTIVITY (CORRELATING MIDBRAIN ACTIVITY WITH THE REST OF THE BRAIN). We entered average midbrain activity (for each subject) as a regressor for the contrast of negative feedback versus fixation. There were statistically reliable (whole-brain FDR corrected) correlations between midbrain and MDS target areas such as ventral striatum, orbital, and dorsomedial frontal foci as well as other brain regions (Fig. 6). The specificity of the revealed network (i.e., the activation of MDS regions but not many other regions activated by the contrast of negative feedback vs. fixation; such as widespread visual cortex, see Fig. 3) makes it unlikely that this result reflects global correlations in the amount of activation across subjects. Instead, it appears likely that co-activation of MDS regions during the probabilistic-learning task occurred because those regions are functionally connected.

LEARNING ANALYSIS. There were no significant linear trends across scanning sessions, probably reflecting the fact that learning occurred rapidly, with subjects achieving ~70% accuracy within the first session (Fig. 2).

DISCUSSION

We scanned subjects while they learned which combination of features on a Mr. Potato Head figure predicted a chocolate or vanilla outcome. Feedback consisted in showing the Mr. Potato Head figure with chocolate or vanilla ice cream where the outcome was probabilistically associated with the feature combinations. When the subject's prediction (indicated by a button press) matched the outcome, the feedback was coded as positive; but when there was no match, the feedback was coded as negative. We separately modeled the BOLD response to

FIG. 6. Midbrain activity is correlated with activity in other—encephalic dopaminergic system (target areas. A functional connectivity analysis, across 15 subjects, for the comparison of negative feedback with fixation, reveals that ventral striatal (vSTR), lateral orbital (lOFC), and dorsomedial frontal (dmPFC) activity correlates reliably with midbrain (MB) activity (FDR corrected for whole-brain, $P < 0.05$).

stimulus, delay, and positive and negative feedback events. We performed contrasts between these event types and show assesses how brain activity changed with the level of informational uncertainty in the stimulus-outcome relation.

Compared with fixation, stimulus and positive and negative feedback events showed a similar pattern of activity (consisting of widespread areas of cortex and subcortical regions, including the MDS and its targets), but activation for the delay event was different. The delay event instead activated a fronto/striatal/parietal system consistent with a working memory role, as would be expected by the requirement to hold "on-line" the stimulus-related prediction. This difference between delay and other events, combined with a validation analysis, confirmed that the event-related design could efficiently separate BOLD responses to different trial components. Despite similar overall patterns of activity for stimulus and feedback events, direct comparisons between these events revealed important activation differences at midbrain, ventral striatal, and medial frontal foci consistent with the MDS and its targets. In particular, midbrain activity was significantly greater for negative than positive feedback, was significantly greater the greater the uncertainty of stimulus-outcome associations, and was reliably correlated with activity in other regions such as ventral striatum and medial frontal cortex.

The results provide novel human evidence for the role of midbrain neurons in coding uncertainty and implicate the midbrain within a wider system underlying feedback learning. The results suggest that activation of these brain regions—typically associated with reward processing (Berns et al. 2001; Delgado et al. 2000; Hollerman et al. 2000; Knutson et al. 2001a,b; McClure et al. 2003; O'Doherty et al. 2003; Pagnoni et al. 2002)—should be conceptualized more generally in terms of informationally salient events rather than specifically in terms of "reward." It should be noted that although the feedback consisted in a picture of the Mr. Potato Head figure holding ice cream (perhaps a secondary reinforcer), this was the case on both positive and negative feedback trials, and so this in itself could not account for the significant differences in brain activation evinced by the contrast of these events. Before we discuss the findings in detail, we note the caveat that fMRI cannot uniquely identify dopaminergic neurons as the source of the activation in our study. It is known, for example, that a substantial proportion of SN/VTA neurons that project to the striatum and PFC are GABAergic (Carr and Sesack 2000). Nevertheless, there are strong parallels between our imaging results and those from neurophysiological studies of SN/VTA dopamine neurons as well as substantial overlap between the regions shown here and those identified as dopaminergic targets using positron emission tomography (Martinez et al. 2003).

Midbrain activity codes the prediction error and uncertainty

Consistent with neurophysiological recordings suggesting a role for the midbrain in mediating learning (Hollerman and Schultz 1998), we found significant activation change in the midbrain for negative compared with positive feedback. Positive feedback should evoke relatively little learning-related activity because it is of little informational value (especially after block 2 when performance was near maximal, Fig. 2). Negative feedback, by contrast, is particularly salient because
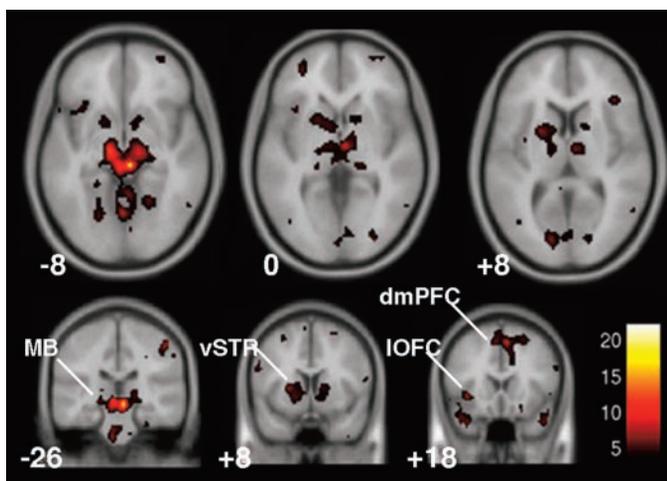
it may generate surprise, and it suggests the subject should change expectancy (Rescorla and Wagner 1972). However, while results from primate neurophysiology demonstrate *decreased* firing of dopaminergic neurons in the absence of reward (Hollerman and Schultz 1998) (the assumed corollary of negative feedback), we found *increased* fMRI signal for negative feedback compared with positive feedback. This discrepancy may be explained by critical differences between the nature of the fMRI and neurophysiological signals (Logothetis 2003). As opposed to unit recordings of single neurons, fMRI indirectly measures integrated synaptic activity over large ($>10^6$) pools of neurons, leaving many features of the underlying activity (e.g., which transmitter is involved, projection neurons versus interneurons) unknown. The interpretation of fMRI results must also take into account the fact that fMRI signals are thought to arise primarily from postsynaptic processes because they are better correlated with local field potentials (reflecting synaptic input and local interneuron processing) than spiking activity (Logothetis 2003; but see Smith et al. 2002). Thus fMRI activation in the midbrain may reflect active projections into the region, or active interneurons, more directly than the firing level of dopamine projection neurons. In particular, increased signal during negative feedback may reflect the activity of GABAergic signals arising from the ventral striatum (Bolam and Smith 1990), which could result in decreased firing of MDS neurons when an expected reward does not occur. Research has shown that inhibitory signals can result in decreased neural firing but increased fMRI signal (e.g., Lauritzen 2001). Notably, several fMRI studies of reward-related activity have found signal change in MDS regions but in inconsistent directions (Delgado et al. 2000; McClure et al. 2003; O'Doherty et al. 2003; Pagnoni et al. 2002), perhaps reflecting differential effects for primary versus secondary reinforcers or for different tasks.

Further support for the hypothesis of MDS involvement in feedback learning was the finding that a midbrain ROI, centered on the substantia nigra, evidenced a significant relationship between increasing BOLD response and increasing uncertainty for the delay period. This provides human data in support of recent research into dopaminergic neurons in the monkey midbrain (Fiorillo et al. 2003) and is compatible with associative-learning theories relating attention (and learning) to uncertainty about reinforcers (Dayan et al. 2000; Pearce and Hall 1980).

Midbrain activity is correlated with ventral striatal and medial frontal foci

Much animal research suggests the phasic dopaminergic input to the striatum, orbital, and medial frontal cortex relates to neural coding of prediction errors (Schultz and Dickinson 2000). fMRI studies have reported significant activation change of the ventral striatum, in particular the left ventral putamen and bilateral nucleus accumbens, related to mismatches in predicted and actual outcomes of reward (Berns et al. 2001; McClure et al. 2003; O'Doherty et al. 2003; Pagnoni et al. 2002) and to differences between reward and punishment (or nonreward) (Delgado et al. 2000; Knutson et al. 2001b).

Although we did not find reliable activation change in the ventral striatum for the comparison of negative and positive feedback, there was reliable activation change for negative feedback compared with fixation, and across subjects this was reliably correlated with midbrain activity. A medial frontal region was also more active for negative feedback than positive feedback, although it was only marginally reliable following a wholebrain correction for multiple comparisons. However, like the ventral striatum, there was reliable activation change in this region for negative feedback compared with fixation, and across subjects this region was reliably correlated with midbrain activity. The coordinates of this focus (0, 30, 48; for the comparison of negative and positive feedback) were very close to those identified in a study of uncertain versus certain predictions (Volz et al. 2003) (coordinates: 4, 30, 46), as well as a further fMRI study of hypothesis testing (Elliott and Dolan 1998). Several event-related brain potential studies have identified a medial frontal event-related potential component, now known as the error-related negativity, which appears to follow closely (<300 ms) after the subject makes an error. The error-related negativity has been proposed to result from reinforcement-learning signals carried by the MDS on anterior cingulate cortex (Holroyd and Coles 2002). Overall, the functional connectivity analysis supported the hypothesis that phasic dopaminergic input from the midbrain to the ventral striatum and medial frontal cortex may relate to neural coding of prediction errors (Schultz and Dickinson 2000).

Conclusions

This study provides a link between the midbrain/ventral striatal/orbito-frontal/medial frontal system (i.e., the MDS and its targets) and decision-making under uncertainty. It suggests that activity in this network may be related not just to processing of direct rewards but more generally to any form of decision making involving environmental feedback. More specifically, we have shown that the midbrain participates in a network involved in learning associations between stimuli and outcomes when this learning involves predicting an outcome in response to stimuli, making a related response, and getting feedback for that response. Furthermore, the midbrain area was most active under conditions of maximum uncertainty. This is compatible with a role for the SN/VTA in mediating learning by the dopamine system (possibly by impacting on attention or arousal), as well as potentially explaining the putatively reinforcing property of uncertainty itself during gambling and other behavior (Fiorillo et al. 2003). As MRI is capable of detecting hemodynamic changes induced by stimulation with dopaminergic compounds such as amphetamine and cocaine (e.g., Chen et al. 1997), it would be interesting to investigate midbrain and MDS target responses in future fMRI studies of feedback learning employing pharmacological challenge.

REFERENCES

Aharon I, Etcoff N, Ariely D, Chabris CF, O'Connor E, and Breiter HC. Beautiful faces have variable reward value: fMRI and behavioral evidence. *Neuron* 32: 537–551, 2001.

Berns GS, McClure SM, Pagnoni G, and Montague PR. Predictability modulates human brain response to reward. *J Neurosci* 21: 2793–2798, 2001.

Bolam JP and Smith Y. The GABA and substance P input to dopaminergic neurones in the substantia nigra of the rat. *Brain Res* 529: 57–78, 1990.

Breiter HC, Gollub RL, Weisskoff RM, Kennedy DN, Makris N, Berke JD, Goodman JM, Kantor HL, Gastfriend DR, Riorden JP, Mathew RT, Rosen BR, and Hyman SE. Acute effects of cocaine on human brain activity and emotion. *Neuron* 19: 591–611, 1997.

Carr DB and Sesack SR. GABA-containing neurons in the rat ventral tegmental area project to the prefrontal cortex. *Synapse* 38: 114–123, 2000.

Chen YC, Galpern WR, Brownell AL, Matthews RT, Bogdanov M, Isacson O, Keltner JR, Beal MF, Rosen BR, and Jenkins BG. Detection of dopaminergic neurotransmitter activity using pharmacologic MRI: correlation with PET, microdialysis, and behavioral data. *Magn Reson Med* 38: 389–398, 1997.

Dale AM. Optimal experimental design for event-related fMRI. *Hum Brain Mapp* 8: 109–114, 1999.

Dayan P, Kakade S, and Montague PR. Learning and selective attention. *Nat Neurosci* 3: 1218–1223, 2000.

Delgado MR, Nystrom LE, Fissell C, Noll DC, and Fiez JA. Tracking the hemodynamic responses to reward and punishment in the striatum. *J Neurophysiol* 84: 3072–3077, 2000.

Elliott R and Dolan RJ. Activation of different anterior cingulate foci in association with hypothesis testing and response selection. *Neuroimage* 8: 17–29, 1998.

Elliott R, Newman JL, Longe OA, and Deakin JF. Differential response patterns in the striatum and orbitofrontal cortex to financial reward in humans: a parametric functional magnetic resonance imaging study. *J Neurosci* 23: 303–307, 2003.

Estes WK. *Classification and Cognition.* New York: Oxford, 1994.

Fiorillo CD, Tobler PN, and Schultz W. Discrete coding of reward probability and uncertainty by dopamine neurons. *Science* 299: 1898–1902, 2003.

Gehring WJ and Willoughby AR. The medial frontal cortex and the rapid processing of monetary gains and losses. *Science* 295: 2279–2282, 2002.

Genovese CR, Lazar NA, and Nichols T. Thresholding of statistical maps in functional neuroimaging using the false discovery rate. *Neuroimage* 15: 870–878, 2002.

Hollerman JR and Schultz W. Dopamine neurons report an error in the temporal prediction of reward during learning. *Nat Neurosci* 1: 304–309, 1998.

Hollerman JR, Tremblay L, and Schultz W. Involvement of basal ganglia and orbitofrontal cortex in goal-directed behavior. In: *Cognition, Emotion and Autonomic Responses: The Integrative Role of the Prefrontal Cortex and Limbic Structures*, Amsterdam: Elsevier, 2000, p. 193–215.

Holroyd CB and Coles MG. The neural basis of human error processing: reinforcement learning, dopamine, and the error-related negativity. *Psychol Rev* 109: 679–709, 2002.

Horvitz JC. Mesolimbocortical and nigrostriatal dopamine responses to salient non-reward events. *Neuroscience* 96: 651–656, 2000.

Knowlton BJ, Mangels JA, and Squire LR. A neostriatal habit learning system in humans. *Science* 273: 1399–1402, 1996.

Knutson B, Adams CM, Fong GW, and Hommer D. Anticipation of increasing monetary reward selectively recruits nucleus accumbens. *J Neurosci* 21(17): 3683–3687, 2001a.

Knutson B, Fong GW, Adams CM, Varner JL, and Hommer D. Dissociation of reward anticipation and outcome with event-related fMRI. *Neuroreport* 12: 3683–3687, 2001b.

Lauritzen M. Relationship of spikes, synaptic activity, and local changes of cerebral blood flow. *J Cereb Blood Flow Metab* 21: 1367–1383, 2001.

Logothetis N. The underpinnings of the BOLD functional magnetic resonance imaging signal. *J Neurosci* 23: 3963–3971, 2003.

Lucerna S, Salpietro FM, Alafaci C, and Tomasello F. *In Vivo Atlas of Deep Brain Structures.* Berlin: Springer-Verlag, 2002.

Martinez D, Slifstein M, Broft A, Mawlawi O, Hwang DR, Huang Y, Cooper T, Kegeles L, Zarahn E, Abi-Dargham A, Haber SN, and Laruelle M. Imaging human mesolimbic dopamine transmission with positron emission tomography. II. Aamphetamine-induced dopamine release in the functional subdivisions of the striatum. *J Cereb Blood Flow Metab* 23: 285–300, 2003.

McClure SM, Berns GS, and Montague PR. Temporal prediction errors in a passive learning task activate human striatum. *Neuron* 38: 339–346, 2003.

Mirenowicz J and Schultz W. Importance of unpredictability for reward responses in primate dopamine neurons. *J Neurophysiol* 72: 1024–1027, 1994.

O'Doherty JP, Dayan P, Friston K, Critchley H, and Dolan RJ. Temporal difference models and reward-related learning in the human brain. *Neuron* 38: 329–337, 2003.

Pagnoni G, Zink CF, Montague PR, and Berns GS. Activity in human ventral striatum locked to errors of reward prediction. *Nat Neurosci* 5: 97–98, 2002.

Pearce JM and Hall G. A model for Pavlovian learning: variations in the effectiveness of conditioned but not of unconditioned stimuli. *Psychol Rev* 87: 532–552, 1980.

Poldrack RA, Clark J, Pare-Blagoev EJ, Shohamy D, Moyano JC, Myers C, and Gluck MA. Interactive memory systems in the human brain. *Nature* 414: 546–550, 2001.

Redgrave P, Prescott TJ, and Gurney K. Is the short-latency dopamine response too short to signal reward error? *Trends Neurosci* 22: 146–151, 1999.

Rescorla R and Wagner A. A theory of Pavlovian conditioning: variations in the effectiveness of reinforcement and nonreinforcement. In: *Classical Conditioning. II. Current Research and Theory*, edited by Black A and Prokasy W. New York: Appleton Century Crofts, 1972, p. 64–99.

Schultz W. Getting formal with dopamine and reward. *Neuron* 36: 241–263, 2002.

Schultz W and Dickinson A. Neuronal coding of prediction errors. *Annu Rev Neurosci* 23: 473–500, 2000.

Shohamy D, Myers CE, Grossman S, Sage J, Gluck MA, and Poldrack RA. Cortico-striatal contributions to feedback-based learning: Converging data from neuroimaging and neuropsychology. *Brain* 127(Pt 4): 851–859, 2004.

Smith AJ, Blumenfeld H, Behar KL, Rothman DL, Shulman RG, and Hyder F. Cerebral energetics and spiking frequency: the neurophysiological basis of fMRI. *Proc Natl Acad Sci USA* 99: 10765–10770, 2002.

Volz KG, Schubotz RI, and Von Cramon DY. Predicting events of varying probability: uncertainty investigated by fMRI. *Neuroimage* 19: 271–280, 2003.

Waelti P, Dickinson A, and Schultz W. Dopamine responses comply with basic assumptions of formal learning theory. *Nature* 412: 43–48, 2001.