# Parallel contributions of distinct human memory systems during probabilistic learning

Kathryn C. Dickerson [a], Jian Li [b], Mauricio R. Delgado [a,*]

[a] Center for Molecular and Behavioral Neuroscience & Department of Psychology, Rutgers University, 101 Warren Street, Newark, NJ 07102, USA
[b] Department of Psychology, New York University, 6 Washington Place, New York, NY 10003, USA

## ABSTRACT

Regions within the medial temporal lobe and basal ganglia are thought to subserve distinct memory systems underlying declarative and nondeclarative processes, respectively. One question of interest is how these multiple memory systems interact during learning to contribute to goal directed behavior. While some hypotheses suggest that regions such as the striatum and the hippocampus interact in a competitive manner, alternative views posit that these structures may operate in a parallel manner to facilitate learning. In the current experiment, we probed the functional connectivity between regions in the striatum and hippocampus in the human brain during an event related probabilistic learning task that varied with respect to type of difficulty (easy or hard cues) and type of learning (via feedback or observation). We hypothesized that the hippocampus and striatum would interact in a parallel manner during learning. We identified regions of interest (ROI) in the striatum and hippocampus that showed an effect of cue difficulty during learning and found that such ROIs displayed a similar pattern of blood oxygen level dependent (BOLD) responses, irrespective of learning type, and were functionally correlated as assessed by a Granger causality analysis. Given the connectivity of both structures with dopaminergic midbrain centers, we further applied a reinforcement learning algorithm often used to highlight the role of dopamine in human reward related learning paradigms. Activity in both the striatum and hippocampus positively correlated with a prediction error signal during feedback learning. These results suggest that distinct human memory systems operate in parallel during probabilistic learning, and may act synergistically particularly when a violation of expectation occurs, to jointly contribute to learning and decision making.

© 2010 Elsevier Inc. All rights reserved.

## Introduction

The theory of multiple memory systems has been prominently displayed across species suggesting that the medial temporal lobes (MTL), in particular the hippocampus, are involved in declarative learning while the basal ganglia (BG) system, primarily the striatum, supports nondeclarative learning (Sherry and Schacter, 1987; Squire, 1992; Squire and Zola, 1996). A question of debate, however, is how these distinct structures interact to contribute to learning and goal directed behavior. Some research suggests that this interaction may be competitive, such that when one system is engaged, it may inhibit or decrease activation of the other system (e.g., Poldrack and Packard, 2003). Evidence for this type of interaction has been observed in animal research (Lee et al., 2008; Packard and Knowlton, 2002; Packard and McGaugh, 1996) and supported by several human functional magnetic resonance imaging (fMRI) studies of probabilistic

learning (Foerde et al., 2006; Poldrack et al., 2001; Seger and Cincotta, 2006).

Accumulating data, however, suggests that the interactions between these two systems may not be purely competitive. Rather, an alternative hypothesis posits that the BG and MTL may make parallel contributions to learning (e.g., Atallah et al., 2008; Cincotta and Seger, 2007; Tricomi and Fiez, 2008; Voermans et al., 2004; White and McDonald, 2002). We adopt White and McDonald's definition of parallel learning systems, in which the authors state that information passes independently through each system. Each system receives the same information but specializes in representing different aspects of the information and may either simultaneously influence behavior in parallel or may interact directly in either a cooperative or a competitive manner (2002). Part of this argument stems from the anatomical and functional neuroconnectivity between the hippocampus, striatum, and dopaminergic cells in the midbrain (Lisman and Grace, 2005). Lisman and Grace (2005) propose a dynamic model by which the hippocampus detects the entrance of novel information and sends this novelty signal to the ventral tegmental area (VTA) in the midbrain via projections through the subiculum, nucleus accumbens, and ventral pallidum. The dopaminergic neurons located

in the VTA fire in response to this novel information, releasing dopamine into the hippocampus where it enhances long term potentiation; thus forming a functional loop between the hippocampus, nucleus accumbens, and midbrain which specializes in novelty detection. However, this is not the sole anatomical connection between the striatum and the hippocampus. There is also a direct projection from the hippocampus to the ventral medial caudate nucleus in the rodent (Jung et al., 2003). Combining this information with the knowledge of the existence of multiple spiral loops between the striatum and the midbrain dopaminergic centers (Haber, 2003), it is plausible that there may be interactions between the hippocampus, midbrain DA areas, and more dorsal regions of the striatum, via a ventromedial to dorsolateral movement of information through these spiral loops.

To expand upon these interactions, it has been proposed that dopamine's role in reward related learning, which has been typically associated with striatal function (for review see Schultz, 2002), also facilitates long term memory formation in humans, which is more traditionally associated with hippocampal function (Adcock et al., 2006; Shohamy and Adcock, 2010; Shohamy and Wagner, 2008; Wittmann et al., 2005). While a role for dopamine facilitating hippocampal memory formation is well characterized in studies involving non human animals (Packard and White, 1991), only recently has this evidence been extended into human neuroimaging experiments (Adcock et al., 2006; Shohamy and Wagner, 2008; Wittmann et al., 2005). Considering the conflicting evidence regarding the nature of the interactions between these memory systems in humans, it is important to revisit the idea of competitive systems while incorporating reinforcement learning concepts that account for potential dopaminergic influences. Thus, the goal of this paper was to probe the interactions between the BG and MTL during probabilistic learning in humans using fMRI and a reinforcement learning model.

Probabilistic category learning has been commonly used in the literature to examine both nondeclarative and declarative learning (e.g., Knowlton et al., 1994; Poldrack et al., 2001; Shohamy et al., 2004). Here, we used a variant of a probabilistic learning task (Delgado et al., 2005), where participants learned the value of easy and hard cues either through trial and error (feedback learning) or via paired association (observation learning), while tracking BOLD responses in both the BG and MTL during an initial learning phase and in a subsequent test phase. Additionally, given the hypothesized role of dopaminergic midbrain structures in reward processing and long term memory formation, we employed a reinforcement learning algorithm to assess the involvement of the BG and MTL during the generation of prediction error signals during probabilistic learning. We hypothesized that irrespective of the type of probabilistic learning, both structures would show a pattern of activity more characteristic of parallel processing rather than competition. Specifically we hypothesized that both the striatum and the hippocampus would be engaged during learning, but rather than showing signs of competition (negative correlations) we predicted observing signs of parallel activity (voxels within both ROIs involved in probability learning irrespective of learning type).

## Materials and methods

### Participants

Seventeen right handed adults participated in this study (nine females). All participants were screened for a history of neurologic and psychiatric illness as well as head injury (mean age 24 years, SD 4.1). One participant was excluded due to equipment failure in the middle of the session (scanner malfunction). Final analysis was therefore conducted on 16 participants. This study was approved by the Institutional Review Boards of Rutgers University and the University of Medicine and Dentistry of New Jersey. All participants gave informed consent prior to participating in the experiment.

### Experimental paradigm

The paradigm was a variant of a previously used reward learning task (Delgado et al., 2005) adapted to incorporate features of other probabilistic learning paradigms (e.g., Poldrack et al., 2001; Shohamy et al., 2004). In this "card game," participants were expected to learn the correct response associated with several cues for indirect monetary rewards. There were two distinct parts of the game: a learning phase and a test phase (Fig. 1). During the initial learning phase, participants were instructed to learn the value of several visual cues (e.g. a circle). The participants were informed that the numerical value of each cue was either higher (6–9) or lower (1–4) than the number 5. Participants were not required to learn the exact value of each cue (e.g. 3), but simply indicate if each cue was higher or lower than 5. They were tested for accuracy in a subsequent test phase.

We manipulated two independent variables within each experimental phase: the type of learning (feedback or observation) and the predictive outcome of the visual cues, or cue difficulty (easy or hard; see Fig. 1). One type of learning was feedback-based and designed to be more nondeclarative like (Fig. 1A). During feedback learning blocks, participants saw a visual cue (e.g. a circle) and made a button press to guess the numerical value of the cue (higher or lower than 5). Learning occurred via feedback contingent on the participant's response which indicated a correct (check mark), incorrect (X symbol), or missed (a pound sign) trial.

A second type of learning was observation and was designed to be more declarative like (Fig. 1B). During observation learning blocks, participants saw each visual cue paired with an arrow, which provided information about the probabilistic value of the visual cue. An upward facing arrow indicated a numerical value higher than 5 while a downward facing arrow indicated a numerical value lower than 5. Learning occurred via observation of the association between the cue and the probabilistic information provided by the arrow. Participants were requested to make a button press indicating their prediction regarding the value of the cue (higher or lower than 5), followed by a non-informative message indicating if their answer was or was not recorded (missed trial). We requested that participants make a button press to indicate the value of the cues in the observation blocks to try to equate the difficulty and motor requirements across learning types, as well as provide a comparable measure of learning (i.e., prediction of cue value). Specifically, participants were instructed to optimize their responses. They were told that as the value of the cues was probabilistic, the direction of the arrow would change. For example, if the square is 85% lower than five it is also 15% higher than five. Participants were instructed to pay careful attention to the arrows and to try to determine what the value of each cue was most of the time. On those few inconsistent trials (square is higher than five) they were instructed to push the button indicating "lower," if that is what they believed the value of the cue to be most of the time. Therefore, participants were actively making a decision regarding each cue's value and were not passively pushing the button that always matched the direction of the arrow. Behavioral accuracy was scored according to participants' actual performance. That is, since we instructed participants to optimize their responding, if they always pushed the button indicating a lower than five response, for the square for example, they would receive a score of 100% correct; if they followed the arrows exactly they would probability match and be 85% correct.

The two versions of our task therefore have a few key differences in both the cue presentation and the feedback phases of the trials. In the feedback version, the cue presentation phase consists only of the cue on the screen. In the observation version of the task, the cue presentation phase consists of two stimuli on the screen: the cue as
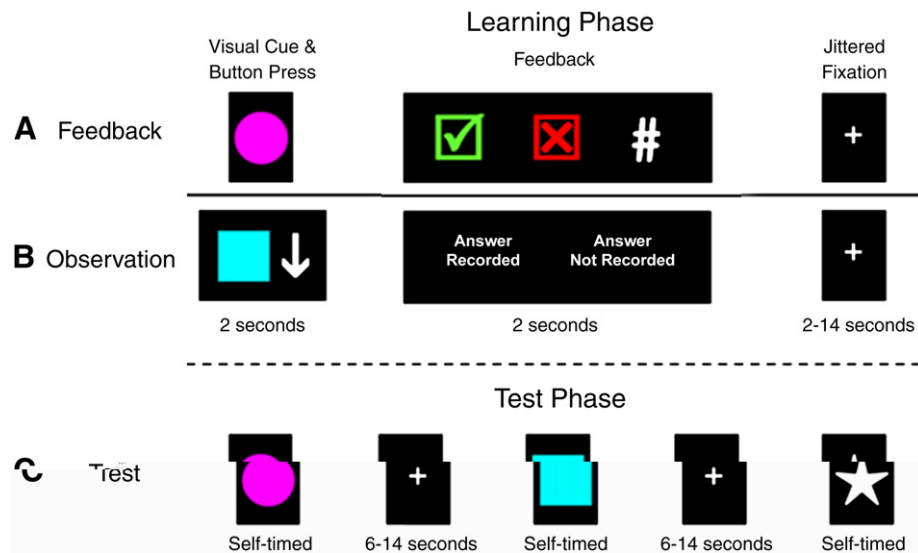
## Learning Phase



**Fig. 1.** Experimental design of probabilistic learning task. (A) Example of a feedback learning trial (FB), where participants are presented with a cue and asked to make a response if the value of that cue was higher or lower than "5" (2 s), followed by a feedback presentation depicting either a correct, incorrect, or missed trial (2 s), and a jittered inter-trial-interval (ITI; 2–14 s). (B) Example of an observation learning trial (OB), where participants are presented with a cue and an arrow that provides information about the value of that cue, while being asked to make a response indicating the probabilistic cue value (2 s). This is followed by a non informative feedback response indicating if a participant's answer was recorded or not (2 s) and a jittered ITI (2–14 s). (C) Example of a test phase trial where participants are presented with a cue and asked to make a response about the value of that cue (self timed), followed by a jittered ITI (6–14 s). Feedback is not provided in the test phase.

well as an arrow which provides information regarding the cue's value. During the feedback phase, feedback is given to the participant following his or her response for the feedback version. In the observation version, participants receive a message indicating whether or not their response has been recorded. It is hypothesized that the feedback version is akin to trial and error learning, considered to be more similar to nondeclarative types of learning, as the participant must learn which cue is associated with what value via guessing (initially) and subsequently receiving feedback. The non-declarative component of this version lies in the trial and error nature of the feedback trial structure, which has been known to recruit regions of the BG (Poldrack et al., 2001; Shohamy et al., 2008). Furthermore, it is thought that the observation version of the task is akin to paired-associate learning, considered to be more similar to declarative types of learning, as the participant views both the cue and its value simultaneously and may therefore overtly memorize the association. The declarative component of this version is the cue presentation phase, where participants may engage in memorization strategies, and has been linked with MTL function, particularly the hippocampus (Poldrack et al., 2001). However, these tasks may not be exclusively solved via one learning mechanism or another. This issue will be addressed further in Discussion.

The other independent variable manipulated was cue difficulty (the probability of the visual cues). The value of each cue was probabilistic and the exact values were unknown to the participants. The cue values were either 85% predictive of the outcome (higher or lower than 5), hereafter referred to as "easy" cues, or 65% predictive of the outcome (higher or lower than 5), hereafter referred to as "hard" cues.

The learning phase consisted of 160 total trials, divided into 4 blocks of 40 trials each. The task was pseudo blocked by learning type such that each participant completed 2 alternating blocks of both feedback and observation learning. Participants were presented with eight different visual cues: four in the observation learning block and four different cues in the feedback learning block (two easy and two hard cues per block). Each trial began with a 2-s presentation of a visual cue which prompted a behavioral response. This was followed by a 2 second feedback time period and a jittered inter trial interval (ITI, 2 to14 seconds) before onset of the next stimulus (Fig. 1). Participants were aware of the type of learning block by the color coded visual cues (e.g. feedback cues were pink and observation cues were blue). Both cue

color and block order were counterbalanced across participants, while cue presentation order was randomized within each block.

The test phase was presented immediately following the learning phase and required participants to make an accuracy judgment (i.e., higher or lower than 5) to the presented visual cues (Fig. 1C). All 8 cues were presented in the test phase; although neither feedback nor observation information (e.g., arrows) was provided. Additionally, two novel cues were presented in the test phase to provide a non studied control condition. Each trial in the test phase consisted of visual cue presentation which prompted a behavioral response (self timed) followed by a jittered ITI (6 to14 s) before onset of the next trial. The test phase consisted of 30 trials, including 3 presentations of each of the 10 visual cues. Participants' accuracy was also recorded during a follow up behavioral only test session administered between 1 and 7 days following initial task participation in the laboratory (M = 2.63 days, SD = 1.63).

The task was programmed with E PRIME V.2.0 (PST, Pittsburgh) and presented to participants in the fMRI scanner via a back projection system. Participants made behavioral responses using a MRI compatible button box. Prior to beginning the experiment, participants were instructed on how to play the game and played a short version of the game with different stimuli for practice purposes. Participants' compensation consisted of a minimum of $50 for the 2-day experimental sessions, but also included additional incentives based on their performance. Specifically, participants were instructed they could earn bonus money based on their test phase accuracy bringing their total compensation to an amount ranging from $50 to $65.

### fMRI acquisition and analysis

A 3 Tesla Siemens Allegra scanner was used to collect the structural (T1 weighted MPRAGE: 256 × 256 matrix; FOV = 256 mm; 176 1 mm sagittal slices) and functional images (single shot echo EPI sequence; TR = 2000 ms, TE = 25 ms; FOV = 192 cm; flip angle = 80°; matrix = 64 × 64; slice thickness = 3 mm). We obtained forty contiguous oblique axial slices (3 × 3 × 3 mm voxels) parallel to the anterior commissure, posterior commissure line. BrainVoyager QX software (Version 1.10; Brain Innovation, Maastricht, The Netherlands) was used to preprocess and analyze the functional data. Preprocessing consisted of motion correction (six parameter, three dimensional

motion correction), spatial smoothing (4 mm, FWHM), voxel wise linear detrending, high pass filtering of frequencies (3 cycles per time course), and normalizing the data to Talairach sterotaxic space (Talairach and Tournoux, 1988). A canonical two gamma hemodynamic response function was used to convolve the events of interest.

During the learning phase, we examined neural activity across the 4 second period where learning occurred (2-s stimulus presentation/participant response + 2-s feedback presentation). We chose to conduct the analysis in this manner due to the different nature of the feedback and observation trials. Relevant information in the feedback trials occurs during the feedback presentation (last 2 s of the trial) whereas relevant information in the observation trials occurs in the cue presentation phase (first 2 s of the trial). Thus, we examined the BOLD activity during the initial 4 second window of a trial to most accurately compare feedback and observation learning. We conducted a random effects general liner model (GLM) analysis using learning type (observation and feedback) and cue difficulty (easy and hard) as predictors. Additionally, 6 motion parameters were included as regressors of no interest to further control for motion related issues. From this GLM, we generated statistical parametric maps (SPM) thresholded at $p < 0.005$, using a cluster threshold with an extent of 5 contiguous voxels (equivalent to 135 mm$^3$ of tissue in $1 \times 1 \times 1$ mm increments). Unless otherwise noted, for all SPMs generated during learning phase analyses, a correction for multiple comparisons was employed via the cluster level statistical threshold estimator plugin in the BrainVoyager analysis package (Forman et al., 1995; Goebel et al., 2006). This tool uses Monte Carlo simulations to determine the likelihood of observing clusters of various sizes. First, the map is thresholded at the desired level (e.g., $p < 0.005$) and then a whole brain correction is performed based on an estimate of the spatial smoothness of the map selected and Monte Carlo simulations which estimate the rate of cluster level false positives. Following 1000 iterations (the recommended number), the minimum cluster size threshold which produces the desired cluster level false positive alpha rate (5% was chosen) is automatically applied to the selected map. For each analysis, all active clusters in the resulting map are used to make a table which summarizes the number of clusters above the desired threshold for each size. When the analysis is complete, each cluster size is assigned an alpha value determined by the frequency of its occurrence in the SPM. Therefore, this method corrects for multiple cluster tests in the map and produces a cluster level false positive rate of 5%.

In order to investigate activity in the brain during learning, an Analysis of Variance (ANOVA) was conducted using learning type (feedback and observation) and cue difficulty (easy and hard) as within subjects factors. The first SPM of interest investigated a main effect of cue difficulty. This primary analysis allowed for a non biased examination of the relative engagement and pattern of activity in the BG, particularly the striatum, and the MTL, with focus on the hippocampus, during the learning phase of the task. Regions of interest (ROIs) were defined based on the results from the main effect of cue difficulty analysis and the BOLD signal (characterized by mean parameter estimates or beta weights) of these a posteriori ROIs (e.g., BG) was examined in order to explore differences between feedback and observation learning in a post hoc analysis. Additionally, a potential effect of time was examined during the learning phase. In order to model time, the learning phase was separated according to block/run. The BOLD signal was extracted from the a posterior regions of interest (e.g., hippocampus and caudate nucleus) for the first learning block/run of each learning type (e.g., observation trials 1 to 40; feedback trials 1 to 40) and again for the second learning block/run of each learning type (e.g., observation trials 41 to 80; feedback trials 41 to 80). The mean BOLD signal from the "early" learning run was then compared to the mean BOLD signal from the "late" learning run in both ROIs for both learning types. The second SPM of interest investigated a main effect of learning type. The third SPM of interest examined potential interactions of learning type and cue difficulty.

During the test phase, we examined neural activity across the entire trial (stimulus onset and participant response). A random effects GLM was used with observation (easy and hard), feedback (easy and hard), and novel cues as predictors, along with 6 motion parameters as regressors of no interest. From this GLM, we generated SPMs thresholded at $p < 0.005$, that contrasted studied (observation and feedback cues collapsed across difficulty) vs. non studied (novel) cues. The cluster level statistical threshold estimator plugin was also used on the SPM generated for the test phase producing a cluster level false positive rate of 5%. Any ROIs which did not withstand correction are clearly labeled and should be interpreted with caution (Poldrack et al., 2008). Parameter estimates were extracted from the resulting ROI and used for further post hoc analyses. Post hoc analyses conducted on the learning phase behavior and neuroimaging data as well as the test phase behavior and neuroimaging data which consisted of more than two t tests within a family of comparisons were corrected for multiple comparisons with the sequential Bonferroni correction (Holm, 1979; Rice, 1989).

Granger causality analysis

A Granger causality analysis was performed on the learning phase data to examine both functional and effective connectivity in the brain. The purpose of this analysis was to determine the relationship between a seed region (x) and activity within the rest of the brain. Geweke (1982) proposed a measure of linear dependence, $F_{x,y}$ between two hypothetical times series of data, x[n] and y[n], using vector autoregressive models. $F_{x,y}$ consists of the sum of three distinct components:

$$F_{x,y} = F_{x \to y} + F_{y \to x} + F_{x \cdot y}$$

$F_{x \to y}$    is a measure of the directed influence from x to y; examining if past values of x improve the current predicted value of y.

$F_{y \to x}$    is a measure of the directed influence from y to x; examining if past values of y improve the current predicted value of x.

$F_{x \cdot y}$    is a measure of the undirected instantaneous influence of x and y. This measure incorporates the current value of x or y into the model which already contains the past values of x and y.

Therefore, the Granger causality analysis measures both directed ($F_{x \to y}$ and/or $F_{y \to x}$) as well as undirected instantaneous influence ($F_{x \cdot y}$) between a specified seed region of the brain (x) and the rest of the brain (see Goebel et al., 2003; Roebroeck et al., 2005 for more details).

A Granger causality analysis was conducted to specifically probe interactionslpast eac.4(s)-5577.9(ip-647.31(caudateTc(1ast)-363(eacs(41)t-2:

extent of 7 voxels for the observation maps and 8 voxels for the feedback maps (correcting to a cluster level false positive rate of 5%). Last, standard second level statistics were performed on the group maps; specifically we ran a t test on the group data. Reported results represent functional connectivity data (instantaneous influence without directionality information; $F_{x\cdot y}$) as effective connectivity ($F_{x\rightarrow y}$ and/or $F_{y\rightarrow x}$) results were not observed between the a posteriori regions of interest. Finally, as a control analysis, a second Granger causality analysis was performed in an identical manner using the caudate nucleus ROI obtained from the main effect of cue difficulty analysis from the learning phase ANOVA. Maps were thresholded at $p<0.0005$ and a cluster threshold of 12 (correcting to a cluster level false positive rate of 5%) unless otherwise stated.

Prediction error analysis

We applied a reinforcement learning model to the behavioral accuracy data for the feedback trials only. The prediction error (PE) regressor was calculated based on a Q learning model (Watkins, 1989). In this model, the expected values for actions (indicating whether the value of the cue was higher or lower than 5) were updated using the Bellman equation and subjects' actual choices were determined by a softmax function of action values. Three parameters were used in the model: (1) learning rate, $\lambda$ (2) action values, $w_\uparrow$ and $w_\downarrow$ (3) softmax function temperature, m. Optimal values for these parameters were estimated using a maximum likelihood estimation algorithm (MLE). A single set of free parameters was used for all participants when modeling the PE regressor. The prediction error is represented by $\delta$; r is equal to the amount of reward, which in the feedback trials was either 0 for incorrect trials or 1 for correct trials; i = trial number and, j represents action type: guessing a high or low value.

$$w_{j,i+1}\leftarrow w_{j,i} + \lambda \tag{1}$$

$$= r - w_{j,i} \tag{2}$$

The prediction errors generated from the above equations using the optimal parameters were used as a regressor in the GLM neuroimaging analysis. We also included two additional predictors, trial event and missed responses, as well as the six motion parameters as regressors of no interest in the analysis. The PE was coded during the 2-s feedback presentation phase for the feedback learning trials only. The resulting SPM was generated at a threshold of $p<0.005$ with a voxel contiguity threshold of 5 continuous voxels (correcting to a cluster level false positive rate of 5%) and probed regions of the brain that correlated with PE during probabilistic learning in the feedback session.

**Re**

Behavioral results

Learning phase: accuracy
In order to examine accuracy differences between the learning types and levels of cue dif

Bonferroni correction), but no differences were observed for observation performance ($t_{(15)} = 1.04$; $p > 0.05$). As expected, participants' performance in the novel condition did not differ over time ($t_{(15)} = 1.01$; $p > 0.05$).

During the individual test phases, no differences between feedback and observation trials or between easy and hard cues were observed. In the immediate test phase, for example, a 2 (learning type: observation vs. feedback) × 2 (cue difficulty: easy vs. hard) repeated measures ANOVA was performed to examine differences in accuracy between learning types and the level of cue difficulty (excluding the novel information and examining only the previously studied material during the immediate test session). This analysis revealed no significant main effect of learning type ($F_{(1,15)} = 0.06$; $p > 0.05$), no significant main effect of cue difficulty ($F_{(1,15)} = 0.23$; $p > 0.05$), nor a significant interaction ($F_{(1,15)} = 1.92$; $p > 0.05$), suggesting that participants successfully learned contingencies independent of learning type and cue difficulty. A similar 2 × 2 repeated measures ANOVA was performed for the follow up test phase examining potential differences between learning type and level of cue difficulty. No significant main effect of learning type ($F_{(1,15)} = 0.51$; $p > 0.05$), no significant main effect of cue difficulty ($F_{(1,15)} = 0.37$; $p > 0.05$), nor a significant interaction ($F_{(1,15)} = 0.00$, $p > 0.05$) were observed.

### Neuroimaging results

#### Learning phase: main effect of cue difficulty

From the learning type × cue difficulty ANOVA, a main effect of cue difficulty was examined (Supplementary Table 1). A region of the left caudate nucleus was involved in processing a main effect of cue difficulty (Fig. 3), along with a cluster in the left hippocampus which was uncorrected. Mean parameter estimates from these two a posteriori ROIs were then extracted for further analyses. In the left caudate nucleus (x, y, z = −15, 20, 7; Fig. 3A and B), the pattern of BOLD responses was similar for both feedback and observation learning sessions, with no differences between learning type when collapsed across cue difficulty ($t_{(15)} = 1.00$; $p > 0.05$). Post hoc t tests reveal a greater BOLD response for easy than hard cues in both learning types [feedback: ($t_{(15)} = 5.02$; $p < 0.025$); observation: ($t_{(15)} = 3.99$; $p < 0.05$]. In the left hippocampus ROI (x, y, z = −36, −28, −8; Fig. 3C and D) a marginally significant effect was observed when comparing mean parameter estimates from the feedback and observation sessions ($t_{(15)} = 2.01$; $p = 0.06$), with a trend towards greater BOLD responses during the observation session. This difference was primarily driven by performance during the hard trials ($t_{(15)} = 2.40$; $p = 0.03$; trend after sequential Bonferroni correction). Post hoc t tests indicate a greater BOLD response for easy than hard cues in the feedback trials ($t_{(15)} = 3.73$; $p < 0.025$) and a trend towards a greater response for easy compared to hard cues in the observation trials: ($t_{(15)} = 1.79$; $p = 0.09$).

Finally, we examined changes in the mean BOLD signal for feedback and observation trials over time by comparing the BOLD signal early (first block) and late (second block) during the learning phase. A 2 (learning type: feedback vs. observation) × 2 (cue difficulty: easy vs. hard) × 2 (time: early vs. late) repeated measures ANOVA conducted on the mean BOLD signal from the caudate nucleus revealed no main effect of learning type ($F_{(1,15)} = 0.82$; ; $p > 0.05$), a main effect of cue difficulty ($F_{(1,15)} = 95.47$; $p < 0.01$), no main effect of time ($F_{(1,15)} = 1.97$; ; $p > 0.05$), and no significant interactions. Mean parameter estimates approached a significant increase as learning progressed for the observation hard cues in the caudate nucleus ($t_{(15)} = 2.25$; $p = 0.04$; trend after sequential Bonferroni correction; Supplementary Fig. 1A).
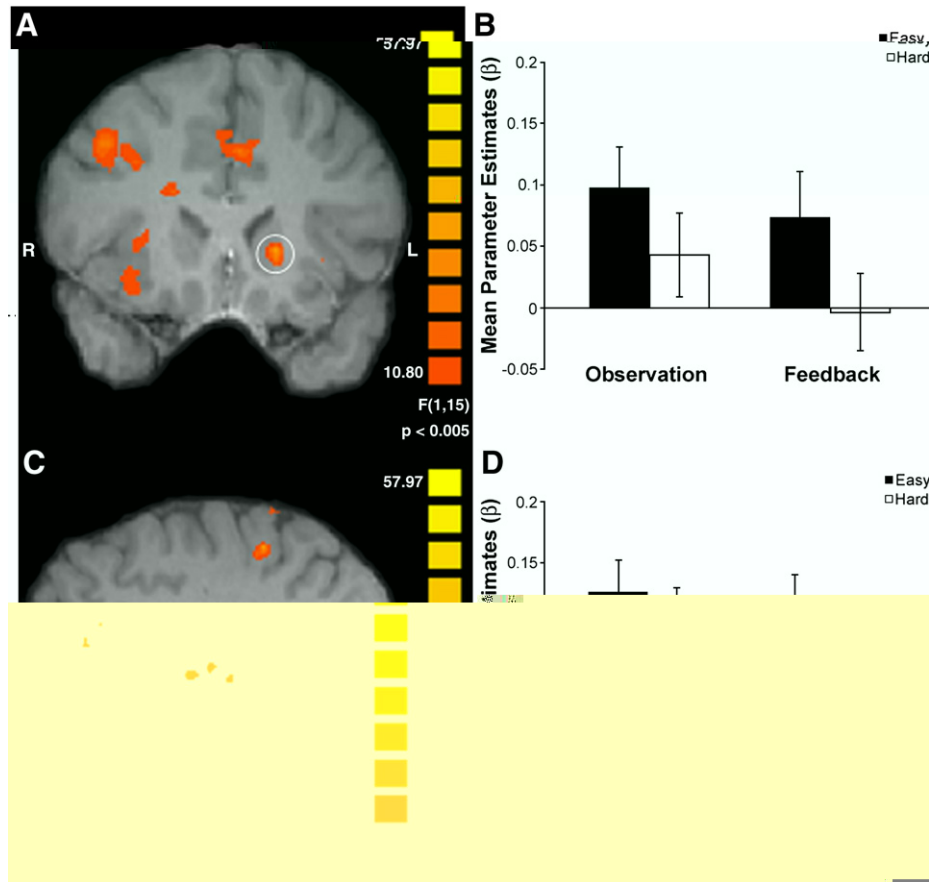


**Fig. 3.** Main effect of cue difficulty analysis during the learning phase identified regions of interest in the (A) left caudate nucleus (x, y, z = −15, 20, 7) and (C) left hippocampus (x, y, z = −36, −28, −8). Graphs depict mean parameter estimates for both the (B) left caudate nucleus and (D) left hippocampus across learning type and cue difficulty. Error bars represent s.e.m.

The same $2 \times 2 \times 2$ repeated measures ANOVA was conducted on the mean BOLD signal from the hippocampus ROI and revealed a trend towards a main effect of learning type ($F_{(1,15)} = 3.36$; ; p = 0.09), a main effect of cue difficulty ($F_{(1,15)} = 20.69$; p < 0.01), a main effect of time ($F_{(1,15)} = 5.35$; ; p < 0.05), and a trend towards a significant interaction between learning type and time ($F_{(1,15)} = 3.76$; ; p < 0.07). Post hoc comparisons revealed a nearly significant increase for the feedback hard cues in the hippocampus as learning progressed ($t_{(15)} = 2.38$; p = 0.03; trend after sequential Bonferroni correction; Supplementary Fig. 1B).

Learning phase: Main effect of learning type

From the same learning type × cue difficulty ANOVA, a main effect of learning type was examined (Supplementary Table 2). This analysis revealed activation in different regions of the basal ganglia, specifically the right ventral portion of the head of the caudate nucleus (x, y, z = 6, 3, 4) and the left ventral caudate nucleus extending into the globus pallidus (x, y, z = −12, 2, 4; Supplementary Fig. 2). No differences with respect to cue difficulty were observed in these ROIs. At the threshold of p < 0.005, no voxels within the hippocampus were observed.

Learning phase: interaction of cue difficulty and learning type

An investigation of the interaction between learning type and cue difficulty from the ANOVA revealed activation in an area of the left medial prefrontal cortex. Post hoc t tests conducted on the mean parameter estimates extracted from this region indicated greater activity for observation easy compared to hard trials ($t_{(15)} = 4.63$; p < 0.025); but no differences in difficulty for the feedback trials ($t_{(15)} = 1.26$; p > 0.05). The interaction is driven by a greater BOLD response to observation compared to feedback easy cues ($t_{(15)} = 2.72$; p < 0.025) and a trend towards a greater response to feedback than observation hard cues ($t_{(15)} = 1.94$; p = 0.07).

Test phase

In the test phase, contrasting studied (observation and feedback cues) versus non-studied (novel cues) information was used to identify ROIs involved in memory processes triggered by cue presentation (Supplementary Table 3). This contrast produced one active region within the medial temporal lobe—the parahippocampal gyrus (x, y, z = −21, −40, −8; Supplementary Fig. 3A and B). Post hoc t tests on the mean parameter estimates extracted from the parahippocampal ROI revealed no significant difference between observation and feedback cues ($t_{(15)} = 0.37$; p > 0.05). Regions of interest just outside the hippocampus (x, y, z = −24, −22, −5), and the right caudate nucleus (x, y, z = 6, 5, 16) were also observed in this contrast, but did not survive correction at the cluster level.

Correlations within the neuroimaging data

We performed a series of Pearson's correlations to explore the relationship between the caudate nucleus and hippocampus during the learning phase. We observed a significant positive correlation between mean parameter estimates from the caudate nucleus and hippocampus ROIs extracted from the main effect of cue difficulty analysis during the observation learning session (r = 0.498, p = 0.05) (Supplementary Fig. 4A). While a similar result was not apparent during the feedback session, a trend towards a positive correlation between the caudate nucleus and hippocampus was observed during later stages of feedback learning, specifically during easy cue trials when participants' expectations were violated by the delivery of incorrect feedback (r = 0.530, p = 0.08) (Supplementary Fig. 4B).

Granger causality analysis

To more effectively assess the level of connectivity between the hippocampus and striatum during probabilistic learning, we conducted a Granger causality analysis. This analysis examined functional connectivity in the brain using the left hippocampus from the main effect of cue difficulty analysis as the principle seed region. The resulting Granger causality maps highlight correlations between the hippocampus (seed region) and regions of the striatum during both feedback and observation probabilistic learning sessions. Specifically, this analysis yielded instantaneous influence between the hippocampus and two regions of the right caudate nucleus during feedback learning (x, y, z = 14, 18, 13) and (x, y, z = 14, 12, 19; not shown), as well as nearly the identical ROIs in the caudate nucleus (x, y, z = 14, 18, 13) and (x, y, z = 14, 10, 18; not shown) and one region of the right ventral putamen during observation learning (x, y, z = 22, 3, −4; not shown) (Fig. 4). The second Granger causality analysis performed as a control using the caudate nucleus as the seed region revealed instantaneous influence between the caudate nucleus and bilateral hippocampal regions during observation learning (x, y, z = −31, −38, −3 and x, y, z = 32, −29, −12; not shown; corrected to a cluster level false positive rate of 5%; Supplementary Fig. 5A) as well as a loci near the right hippocampus during feedback learning (x, y, z = 29, −8, −15; uncorrected for multiple comparisons; Supplementary Fig. 5B).

Prediction error analysis

Midbrain dopaminergic neurons are believed to project to both the striatum and the hippocampus (Lynd-Balta and Haber, 1994; Scatton et al., 1980) and contribute to learning and memory processes (Lisman and Grace, 2005). One correlate of dopaminergic physiological firing during reward related learning is the prediction error (PE) learning signal (Schultz et al., 1997). Based on our observation of mutual activation of the hippocampus and striatum during the learning phase, we applied a reinforcement learning algorithm to our behavioral data to explore purported dopaminergic influences in these regions during probabilistic learning. The PE estimates generated from the reinforcement learning model applied to the feedback session were used as a PE regressor in a GLM. The resulting SPM revealed the left putamen (x, y, z = −30, 2, 4) and the right hippocampus (x, y, z = 27, −28, −14) as regions whose activation positively correlated with the PE signal (Fig. 5). Thus, the use of a reinforcement learning model during feedback learning engaged the typically reported prediction error signals in the striatum (Abler et al., 2006; McClure et al., 2003; Pagnoni et al., 2002), while also engaging the hippocampus, which has been more often associated with mismatch signals (Ploghaus et al., 2000).

## Discussion

In this experiment, we investigated the interaction of multiple memory systems during probabilistic learning while incorporating a reinforcement learning algorithm that is posited to reflect dopaminergic modulation of the basal ganglia. Several studies have suggested that memory systems dependent on the BG and MTL compete with each other during learning (Foerde et al., 2006; Lee et al., 2008; Poldrack et al., 2001). However, our data are inconsistent with the idea of competitive interactions, and instead support the alternative hypothesis that such memory systems work in parallel during probabilistic learning (Atallah et al., 2008; Cincotta and Seger, 2007; Voermans et al., 2004). Using a variant of a probabilistic task that allows for within subjects measures of learning in an event related design, we observed that as learning progressed, accuracy was modulated by difficulty (easy and hard) across type of learning (via feedback or observation). Tracking these behavioral results, ROIs in the BG and MTL, specifically the caudate nucleus and hippocampus, were found to be modulated by cue difficulty, displaying similar patterns of activity to both feedback and observation information during learning. Additionally, BOLD signals in the BG and MTL regions were found to be positively correlated during learning, as assessed by simple comparisons between parameter estimates and a more sophisticated functional connectivity analysis using both the hippocampus and the caudate nucleus as seed regions. This functional
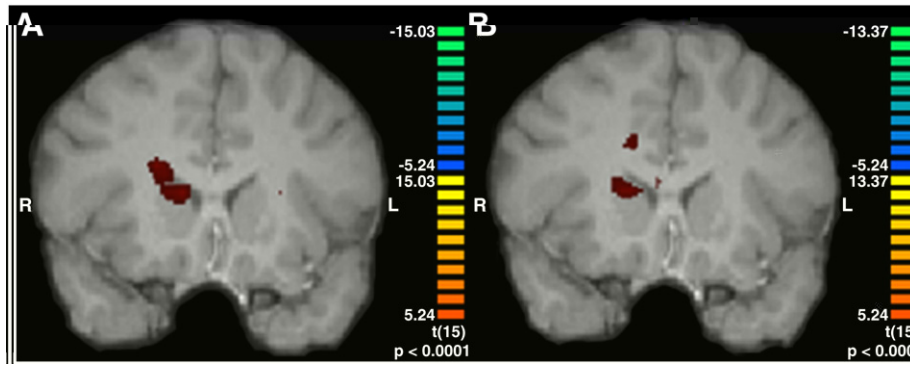
**Fig. 4.** Granger causality analysis using the hippocampus ROI obtained from the main effect of cue difficulty analysis as a reference region. (A) Two regions of the right caudate nucleus (x, y, z = 14, 18, 13) and (x, y, z = 14, 12, 19; not shown) correlated with the hippocampus during the feedback trials, while (B) nearly the same ROIs (x, y, z = 14, 18, 13) and (x, y, z = 14, 10, 18; not shown) and a region of the right ventral putamen (x, y, z = 22, 3, −4; not shown) correlated with the hippocampus during the observation trials.

interaction may potentially be explained by dopaminergic modulation during reward related learning (Lisman and Grace, 2005; Shohamy et al., 2008) as both BG and MTL ROIs were found to correlate with a prediction error signal, further corroborating the hypothesis that these distinct memory systems interact in a parallel manner while processing probabilistic information to facilitate goal directed behavior.

These results complement the visuomotor learning and simple association learning literature, which suggests that the medial temporal lobe and basal ganglia may be online simultaneously and are involved in learning arbitrary visuomotor associations (Amso et al., 2005; Haruno and Kawato, 2006; Law et al., 2005; Toni et al., 2001). In particular, our results are consistent with a recent study by Mattfeld and Stark (2010) which investigated the interaction of the MTL and the BG during an arbitrary visuomotor association task. The authors reported that several regions of the MTL and BG demonstrated an increase in BOLD signal as the strength of memory increased during a trial and error task, suggesting that regions of the MTL and BG are involved in learning arbitrary associations. Our observation of the hippocampus and caudate nucleus exhibiting larger BOLD responses to easy compared to hard cues complements Mattfeld and Stark's result. The authors also employed a functional connectivity analysis which revealed connectivity between both the ventral (nucleus accumbens) and dorsal (caudate nucleus) striatum with the hippocampus during learning, further supporting the interactive nature of these systems during learning and memory processes. One important distinction between our current task design and more traditional visuomotor tasks is the inclusion and investigation of different types of learning (observation and feedback).

Our results enhance recent neuroimaging findings demonstrating noncompetitive interactions between the BG and MTL during category learning (Cincotta and Seger, 2007; Voermans et al., 2004). For instance, a recent blocked design fMRI study which used an

information integration category learning task observed similar patterns of activation to feedback and observation information in the hippocampus and regions of the striatum (Cincotta and Seger, 2007). The focus of this paper was on categorizing the BOLD response of different sub regions of the striatum to feedback and observation information. Interestingly, bilateral hippocampal activity to both types of information was also reported, leading to the interpretation that the striatum and hippocampus interact noncompetitively during information integration category learning. In the current experiment, an event related design allowed for the decoupling of factors such as difficulty and examining changes across learning to further lend support to the hypothesis that parallel processing in the BG and MTL contributes to overall learning during probabilistic paradigms (irrespective of the type of learning). Furthermore, we observed a significant positive correlation between the BG (caudate nucleus) and MTL (hippocampus) during the observation learning session and a trend towards a positive correlation during the feedback learning session. Whereas negative correlations may imply competition between memory systems (Poldrack et al., 2001), positive correlations may suggest noncompetitive, perhaps even synergistic interactions, leading to the interpretation that one system may inform the other in specific contexts to facilitate probabilistic learning. This functional connectivity is illustrated by Granger causality maps using both the hippocampus and the caudate nucleus as seed regions, which highlighted that areas within the striatum and the hippocampus were correlated at simultaneous time points during the learning phase.

It should be noted that our hippocampus region of activation from the main effect of cue difficulty analysis is reported at p<0.005, uncorrected for multiple comparisons. Additionally, two regions from the test phase analysis, the caudate nucleus and a region adjacent to the hippocampus, are also reported at p<0.005 uncorrected. Although this information could be useful for future studies, regions reported at
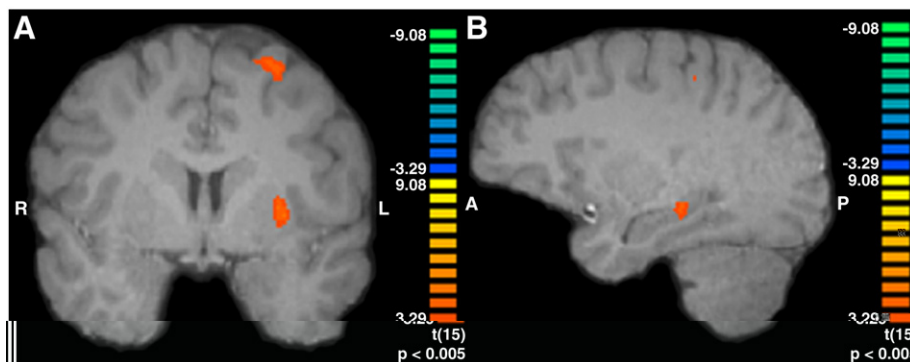


**Fig. 5.** Activation in both (A) left putamen (x, y, z = −30, 2, 4) and (B) right hippocampus (x, y, z = 27, −28, −14) were found to positively correlate with a prediction error signal during probabilistic feedback learning.

an uncorrected level should be regarded with caution when interpreting the results due to the increased likelihood of producing a Type I error (Poldrack et al., 2008).

The major differences between the feedback and observation versions of our task were outlined the Materials and methods section. Despite their differences, however, the two learning sessions share the common goal of learning the value of probabilistic cues. Thus, participants may engage in a variety of cognitive strategies in order to facilitate successful performance. As learning progresses over time in the feedback session, for instance, it is possible that participants employ a more declarative based cued recall strategy during the cue phase. Participants may also use verbal rehearsal strategies during the learning phase, irrespective of the task version. Research examining how participants solve another probabilistic learning task, the Weather Prediction Task (WPT), may shed some insight into possible declarative and nondeclarative components of category learning tasks as well as the knowledge that participants may have during these types of learning tasks (Gluck et al., 2002; Meeter et al., 2006). A relatively recent study by Newell et al. (2007) found that participants had comparable declarative knowledge on a feedback and observation version of the WPT. The authors argue therefore, that the feedback

differences were also apparent. First, only regions within the basal ganglia were modified by a main effect of learning type, while no voxels were identified in the MTL showing such differentiation. Several neuroimaging papers have shown that feedback and reward processing recruit ventromedial regions of the striatum (for review see Delgado, 2007), thus, it is not surprising that this region was recruited more strongly during the feedback learning trials. It may have been expected that the MTL would be selectively modulated by the observation version given previous results (Poldrack et al., 2001); however we did not observe this. While a null result in neuroimaging is not indicative of any particular finding per se and the context and details of our paradigm differ from previous probabilistic learning studies, it is possible that MTL BOLD signals within our paradigm are recruited during both feedback and observation learning–as suggested by the main effect of difficulty analysis–to contribute to overall learning. A second difference which emerged between the hippocampus and caudate nucleus was that the hippocampus showed a main effect of time (early × late learning) during the learning session, whereas caudate nucleus responses were not significant. This effect was driven primarily by activity during feedback learning (primarily for the hard cues), which was greater during late compared to early stages of learning in the hippocampus. This result may suggest that the involvement of the hippocampus in feedback learning happens later on during the learning process.

One finding to note is that the active voxels comprising regions of interest within the BG varied according to the type of analysis. It is perhaps not surprising, given the multifunctionality and connectivity of the striatum (Middleton and Strick, 2000a,b; Pennartz et al., 2009) that different voxels would be sensitive to the type of learning or level of difficulty. In the current experiment, dorsomedial regions of the striatum (caudate nucleus) were involved in initial learning and testing, whereas the dorsolateral regions (putamen) were involved in subsequent analyses with activity during learning positively correlating with prediction error signals. A functional connectivity analysis also revealed correlations between both caudate nucleus and putamen regions and the hippocampus. Interestingly, a ventromedial striatum region was found to be more responsive to feedback compared to observation learning, suggesting that some striatal subcomponents may distinguish between learning type (Poldrack et al., 2001) as supported by neuropsychological studies (Shohamy et al., 2004). In our design, however, this analysis is affected by the presence of affective feedback known to engage regions of the ventromedial striatum (for review see Delgado, 2007). Yet, this result is in accordance with previous probabilistic learning studies which report engagement of posterior parts of the caudate nucleus and putamen during learning, and more anterior parts of the caudate nucleus and ventral striatum linked with feedback processing (Cincotta and Seger, 2007; Seger and Cincotta, 2006).

In contrast to the striatum results, activity in the hippocampus was fairly consistent across analyses. There has also been recent research parsing out the distinct functional role of the subregions of the hippocampus. However, these studies have focused more on other issues such as anatomical distinctions between regions involved in memory encoding versus retrieval as well as pattern completion versus pattern separation (Bakker et al., 2008; Eldridge et al., 2005; Greicius et al., 2003). Future studies using high resolution fMRI may investigate how distinct subregions of MTL and BG interact during learning, to contribute to decision making processes underlying goal directed behaviors. To conclude, an investigation of the interactions between the MTL and BG during probabilistic learning suggests that these distinct memory regions may interact in a parallel manner to facilitate goal directed learning, acting synergistically during prediction error like learning scenarios.

during the gradual acquisition and expression of paired-associate memory. J. Neurosci. 25, 5720–5729.

Lee, A.S., Duman, R.S., Pittenger, C., 2008. A double dissociation revealing bidirectional competition between striatum and hippocampus during learning. Proc. Natl Acad. Sci. USA 105, 17163–17168.

Lisman, J.E., Grace, A.A., 2005. The hippocampal-VTA loop: controlling the entry of information into long-term memory. Neuron 46, 703–713.

Lynd-Balta, E., Haber, S.N., 1994. The organization of midbrain projections to the ventral striatum in the primate. Neuroscience 59, 609–623.

Mattfeld, A.T., Stark, C.E., 2010. Striatal and medial temporal lobe functional interactions during visuomotor associative learning. Cereb. Cortex.

McClure, S.M., Berns, G.S., Montague, P.R., 2003. Temporal prediction errors in a passive learning task activate human striatum. Neuron 38, 339–346.

Meeter, M., Myers, C.E., Shohamy, D., Hopkins, R.O., Gluck, M.A., 2006. Strategies in probabilistic categorization: results from a new way of analyzing performance. Learn. Mem. 13, 230–239.

Meeter, M., Radics, G., Myers, C.E., Gluck, M.A., Hopkins, R.O., 2008. Probabilistic categorization: how do normal participants and amnesic patients do it? Neurosci. Biobehav. Rev. 32, 237–248.

Middleton, F.A., Strick, P.L., 2000a. Basal ganglia and cerebellar loops: motor and cognitive circuits. Brain Res. Brain Res. Rev. 31, 236–250.

Middleton, F.A., Strick, P.L., 2000b. Basal ganglia output and cognition: evidence from anatomical, behavioral, and clinical studies. Brain Cogn. 42, 183–200.

Newell, B.R., Lagnado, D.A., Shanks, D.R., 2007. Challenging the role of implicit processes in probabilistic category learning. Psychon. Bull. Rev. 14, 505–511.

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

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.

Packard, M.G., Knowlton, B.J., 2002. Learning and memory functions of the Basal Ganglia. Annu. Rev. Neurosci. 25, 563–593.

Packard, M.G., McGaugh, J.L., 1996. Inactivation of hippocampus or caudate nucleus with lidocaine differentially affects expression of place and response learning. Neurobiol. Learn. Mem. 65, 65–72.

Packard, M.G., White, N.M., 1991. Dissociation of hippocampus and caudate nucleus memory systems by posttraining intracerebral injection of dopamine agonists. Behav. Neurosci. 105, 295–306.

Pagnoni, G., Zink, C.F., Montague, P.R., Berns, G.S., 2002. Activity in human ventral striatum locked to errors of reward prediction. Nat. Neurosci. 5, 97–98.

Pennartz, C.M., Berke, J.D., Graybiel, A.M., Ito, R., Lansink, C.S., van der Meer, M., Redish, A.D., Smith, K.S., Voorn, P., 2009. Corticostriatal interactions during learning, memory processing, and decision making. J. Neurosci. 29, 12831–12838.

Ploghaus, A., Tracey, I., Clare, S., Gati, J.S., Rawlins, J.N., Matthews, P.M., 2000. Learning about pain: the neural substrate of the prediction error for aversive events. Proc. Natl Acad. Sci. USA 97, 9281–9286.

Poldrack, R.A., Packard, M.G., 2003. Competition among multiple memory systems: converging evidence from animal and human brain studies. Neuropsychologia 41, 245–251.

Poldrack, R.A., Clark, J., Pare-Blagoev, E.J., Shohamy, D., Creso Moyano, J., Myers, C., Gluck, M.A., 2001. Interactive memory systems in the human brain. Nature 414, 546–550.

Poldrack, R.A., Fletcher, P.C., Henson, R.N., Worsley, K.J., Brett, M., Nichols, T.E., 2008. Guidelines for reporting an fMRI study. Neuroimage 40, 409–414.

Rescorla, R.A., Wagner, A.R., 1972. A theory of Pavlovian conditioning: variations in the effectiveness of reinforcement and nonreinforcement. In: Black, A.H., Prokasy, W.F. (Eds.), Classical Conditioning II: Current Research and Theory. Appleton Century Crofts, New York, pp. 64–99.

Rice, W.R., 1989. Analyzing tables of statistical tests. Evolution 43, 223–225.

Roebroeck, A., Formisano, E., Goebel, R., 2005. Mapping directed influence over the brain using Granger causality and fMRI. Neuroimage 25, 230–242.

Scatton, B., Simon, H., Le Moal, M., Bischoff, S., 1980. Origin of dopaminergic innervation of the rat hippocampal formation. Neurosci. Lett. 18, 125–131.

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

Schultz, W., Dayan, P., Montague, P.R., 1997. A neural substrate of prediction and reward. Science 275, 1593–1599.

Seger, C.A., 2008. How do the basal ganglia contribute to categorization? Their roles in generalization, response selection, and learning via feedback. Neurosci. Biobehav. Rev. 32, 265–278.

Seger, C.A., Cincotta, C.M., 2005. The roles of the caudate nucleus in human classification learning. J. Neurosci. 25, 2941–2951.

Seger, C.A., Cincotta, C.M., 2006. Dynamics of frontal, striatal, and hippocampal systems during rule learning. Cereb. Cortex 16, 1546–1555.

Sherry, D.F., Schacter, D.L., 1987. The evolution of multiple memory systems. Psychol. Rev. 94, 439–454.

Shohamy, D., Adcock, R.A., 2010. Dopamine and adaptive memory. Trends Cogn. Sci. 14.

Shohamy, D., Wagner, A.D., 2008. Integrating memories in the human brain: hippocampal–midbrain encoding of overlapping events. Neuron 60, 378–389.

Shohamy, D., Myers, C.E., Grossman, S., Sage, J., Gluck, M.A., Poldrack, R.A., 2004. Cortico-striatal contributions to feedback-based learning: converging data from neuroimaging and neuropsychology. Brain 127, 851–859.

Shohamy, D., Myers, C.E., Kalanithi, J., Gluck, M.A., 2008. Basal ganglia and dopamine contributions to probabilistic category learning. Neurosci. Biobehav. Rev. 32, 219–236.

Squire, L.R., 1992. Declarative and nondeclarative memory: multiple brain systems supporting learning and memory. J. Cogn. Neurosci. 4, 232–243.

Squire, L.R., Zola, S.M., 1996. Structure and function of declarative and nondeclarative memory systems. Proc. Natl Acad. Sci. USA 93, 13515–13522.

Swanson, L.W., 1982. The projections of the ventral tegmental area and adjacent regions: a combined fluorescent retrograde tracer and immunofluorescence study in the rat. Brain Res. Bull. 9, 321–353.

Talairach, J., Tournoux, P., 1988. Co-Planar Stereotaxic Atlas of the Human Brain. Thieme Medical Publishers, Inc., New York.

Toni, I., Ramnani, N., Josephs, O., Ashburner, J., Passingham, R.E., 2001. Learning arbitrary visuomotor associations: temporal dynamic of brain activity. Neuroimage 14, 1048–1057.

Tricomi, E., Fiez, J.A., 2008. Feedback abackuorea9(J.,TD[(arbitrary)-235.4(visuomo8TD[(effec5332(for