SPECIAL ISSUE/REWARD SYSTEMS, COGNITION, AND EMOTION



# Hippocampal contributions to value-based learning: Converging evidence from fMRI and amnesia

Daniela J. Palombo<sup>1,2,3</sup> · Scott M. Hayes<sup>1,2,4</sup> · Allison G. Reid<sup>1</sup> · Mieke Verfaellie<sup>1,2</sup>

© This is a U.S. government work and its text is not subject to copyright protection in the United States; however, its text may be subject to foreign copyright protection 2019

# Abstract

Recent evidence suggests that the human hippocampus—known primarily for its involvement in episodic memory—plays a role in a host of motivationally relevant behaviors, including some forms of value-based decision-making. However, less is known about the role of the hippocampus in value-based learning. Such learning is typically associated with a striatal system, yet a small number of studies, both in human and nonhuman species, suggest hippocampal engagement. It is not clear, however, whether this engagement is *necessary* for such learning. In the present study, we used both functional MRI (fMRI) and lesion-based neuropsychological methods to clarify hippocampal contributions to value-based learning. In Experiment 1, healthy participants were scanned while learning value-based contingencies (whether players in a "game" win money) in the context of a probabilistic learning task. Here, we observed recruitment of the hippocampus, in addition to the expected ventral striatal (nucleus accumbens) activation that typically accompanies such learning. In Experiment 2, we administered this task to amnesic patients with medial temporal lobe damage and to healthy controls. Amnesic patients, including those with damage circumscribed to the hippocampus, failed to acquire value-based contingencies, thus confirming that hippocampal engagement is necessary for task performance. Control experiments established that this impairment was not due to perceptual demands or memory load. Future research is needed to clarify the mechanisms by which the hippocampus contributes to value-based learning, but these findings point to a broader role for the hippocampus in goal-directed behaviors than previously appreciated.

Keywords hippocampus · reward · value · amnesia

A wealth of evidence suggests that episodic memories are augmented in the presence of reward. This reward-based memory enhancement is demonstrated across a range of stimuli and paradigms (e.g., Callan & Schweighofer, 2008; Castel,

**Electronic supplementary material** The online version of this article (https://doi.org/10.3758/s13415-018-00687-8) contains supplementary material, which is available to authorized users.

Daniela J. Palombo Daniela.palombo@psych.ubc.ca

- <sup>1</sup> VA Boston Healthcare System, Brockton, MA, USA
- <sup>2</sup> Department of Psychiatry, Boston University School of Medicine, Boston, MA, USA
- <sup>3</sup> Department of Psychology, University of British Columbia, Vancouver, BC, Canada
- <sup>4</sup> Department of Psychology, The Ohio State University, Columbus, OH, USA

Farb, & Craik, 2007; Madan, Fujiwara, Gerson, & Caplan, 2012; Mather & Schoeke, 2011; Spaniol, Schain, & Bowen, 2014). Although the mechanism for this enhancement is not fully understood, strong evidence is accumulating for involvement of dopaminergic-rich mesolimbic (i.e., midbrain and basal ganglia) systems implicated in reward anticipation, in conjunction with the hippocampus—a region known for its role in episodic memory (see Shohamy & Adcock, 2010, for review). Such studies suggest that mesolimbic regions induce motivational brain states that augment long-term memory processes and highlight interactive synergy between these brain systems (Adcock, Thangavel, Whitfield-Gabrieli, Knutson, & Gabrieli, 2006; Loh et al., 2016; Mather & Schoeke, 2011; Murty & Adcock, 2014; Murty, LaBar, & Adcock, 2016; Wittmann, Bunzeck, Dolan, & Duzel, 2007).

Notably, a similar neural synergy has also been observed during tasks that involve reinforcement-based reward learning. Unlike episodic learning, which involves the rapid acquisition of single-instance events, in typical reinforcementbased tasks, learning occurs over many instances based on trial and error; this type of learning has historically been conceptualized as habitual or incremental in nature. In such cases, the role of the hippocampus is more surprising, as such stimulus-response learning has been thought of as a canonical form of striatally based learning, according to a classic memory systems view (Squire, 2004). For example, using fMRI, Li, Delgado, and Phelps (2011) showed BOLD response both in the striatum and hippocampus when comparing monetary wins to losses during simple feedback-based value learning (also see Delgado, Nystrom, Fissell, Noll, & Fiez, 2000; Dickerson & Delgado, 2015; Preuschoff, Bossaerts, & Quartz, 2006). Further, some work has shown prediction error signaling (which represents the difference between expected and actual outcomes) both in the striatum and the hippocampus using a similar task (Dickerson, Li, & Delgado, 2011; Schonberg et al., 2010; also see Lee, Ghim, Kim, Lee, & Jung, 2012, for related work in rodents),<sup>1</sup> although this has not always been observed (Li et al., 2011).

What is the nature of hippocampal involvement in valuebased learning? One possibility is that the hippocampal recruitment observed in these studies is epiphenomenal to the task at hand, akin to theoretical models of parallel hippocampal processing of stimulus-response contingencies in aspects of conditioning (Gluck, Ermita, Oliver, & Myers, 1997). On the other hand, more recent work suggests that the hippocampal signal observed during such learning may actually contribute to performance. For example, Dickerson and Delgado (2015) showed that accuracy on a value-based learning task was adversely affected in a condition involving a competing hippocampally mediated task (i.e., a concurrent scenerecognition task) and, critically, whereas learning accuracy correlated with hippocampal activity in the standard feedback version of the task, this correlation was significantly reduced when participants performed the concurrent hippocampally based task. Together, these findings provide more substantive evidence that the hippocampal signal observed during learning may be relevant to performance-an idea that potentially challenges a classic memory system's view that regards hippocampal and striatal regions as supporting dissociable aspects of learning and memory (Squire, 2004). Nonetheless, the latter finding is correlational, and as of yet it is unknown whether the hippocampus is *necessary* for this form of learning.

The strongest test of the hypothesis that hippocampal fMRI activation during value-based learning tasks actually contributes to learning would be to examine whether performance on the very same task is adversely affected in amnesic patients who have hippocampal damage. To fill this gap in the literature, in the present study we used a combined neuroimaging (Experiment 1) and lesion (Experiments 2a and 2b) approach and a novel value-based reinforcement learning task. To date, little is known about the consequences of hippocampal lesions on value-based learning. On the one hand, work on reinforcement learning (without a value component) has shown normal performance in amnesic patients, suggesting that this form of learning may not require the hippocampus (Foerde, Race, Verfaellie, & Shohamy, 2013; Shohamy, Myers, Hopkins, Sage, & Gluck, 2009). On the other hand, a study in amnesic patients by Hopkins, Myers, Shohamy, Grossman, and Gluck (2004) suggests that reinforcement learning with a value component may be hippocampal dependent. In that study, amnesic patients and controls were required to learn ice cream preferences for Mr. Potato Head<sup>TM</sup> characters, and on correct trials feedback was accompanied by the sound of coins in a tip jar. Amnesic patients were impaired on this task. However, it should be noted that in this task, optimal learning depended on a combination of cues (i.e., multiple facial features of the Mr. Potato Head<sup>TM</sup> characters). The complexity of cues in itself may have been responsible for the impairment in amnesia, as amnesic patients are also impaired in reinforcement learning without a value component when learning depends on a combination of cues (e.g., the weather prediction task; Hopkins et al., 2004; Knowlton, Squire, & Gluck, 1994). Thus, it is still difficult to ascertain whether hippocampal lesions necessarily interfere with value-based learning.

To examine the role of the hippocampus in value-based learning, in the present study, participants learned the valuebased contingencies of single-cue stimuli: Participants were asked whether single players in a "game" would win or lose money. As in many reinforcement-learning tasks (that typically focus on striatal involvement), the contingencies were probabilistic, such that different outcomes were provided as feedback for a given stimulus (e.g., a "winning" player would win only 75% of the time). In Experiment 1, we administered this task to a group of healthy adults during fMRI scanning to establish that the task successfully recruits the hippocampus (along with the expected activation in the ventral striatum; see below). Notably, given recent interest in hippocampal long-axis specialization of function, and given some evidence for a greater role of the anterior (vs. posterior) hippocampus in motivational behaviors, we tested the hypothesis that the anterior hippocampus would show stronger engagement during value-based learning, acknowledging that the literature is not fully consistent on this matter (see Poppenk, Evensmoen, Moscovitch, & Nadel, 2013; Strange, Witter, Lein, & Moser, 2014).

We next administered the same value-based learning task to a group of amnesic patients with damage to the medial temporal lobes (MTL), including a subset of patients with damage thought to be circumscribed to the hippocampus proper (Experiment 2a). If the hippocampal recruitment observed in Experiment 1 is simply epiphenomenal to the task, then

<sup>&</sup>lt;sup>1</sup> We note that prediction error signaling has also been observed in the hippocampus under other conditions of reinforcement learning (i.e., when there is no reward component; Davidow, Foerde, Galvan, & Shohamy, 2016; Foerde & Shohamy, 2011; Lighthall, Pearson, Huettel, & Cabeza, 2018).

patients should acquire normal stimulus–response contingencies. If, instead, the observed hippocampal activation is required for task performance, patients with hippocampal damage should perform poorly on this value-based learning task. Experiment 2b sought to replicate the findings from Experiment 2a under conditions of reduced memory load.

# **Experiment 1**

# Materials and methods

# Participants

Thirty healthy, right-handed, native English speakers (15 female), with a mean age of 19.6 (SD = 1.0) years and a mean education of 13.2 (SD = 1.1) years participated in the study. Participants were recruited from Boston University through online postings. Participants were given a detailed phone screen prior to participating in the study and were excluded from participation if they had any MRI contraindications or major psychiatric or neurological conditions. The session lasted approximately 2.5 hours (approximately 1 hour in the scanner), and participants were paid \$60 for their participation. The VA Boston Healthcare System and the Boston University School of Medicine institutional review boards approved all experimental procedures, and all participants provided informed consent.

## Task paradigm and procedure

As shown in Fig. 1, participants learned reward-based contingencies-namely, whether distinct players in a "game" would win money or not, in the context of a probabilistic learning task (75% majority outcome status). The players were distinguished based on the color pattern depicted on their jumpsuits. To bias participants away from an explicit rule-forming strategy, for each player we used fractal-like color patterns, which are more difficult to verbalize. Participants were shown an image of the player along with, "Does the man win money?" printed on the screen (2,134 ms). Participants were instructed to press the "yes" button during that time if they believed the player would win, and the "no" button if they believed the player would not win. After a choice was made, a short delay, which displayed the player in isolation (400 ms) was followed by the actual outcome for the player (1,067 ms). If the player won, then a dollar bill was shown above the player, along with, "The man wins \$1.00!" If the player did not win, an opaque gray rectangle (displaying "\$0.00") was shown along with, "The man does not win money!" If the participant failed to make a response, "Too late!" was displayed on the screen. A jittered interstimulus interval preceded the next trial (M =2,801 ms; range: 667-9,203 ms). In a control condition,

randomly intermixed with the abovementioned experimental condition trials, participants made responses for players wherein no learning was required. For such trials, the outcome of the trial ("yes" or "no") was displayed on the face of the player, and the contingencies were consistent for each player (100% rewarded or not rewarded).<sup>2</sup> The exact instructions given to participants are provided in the Supplementary Materials.

Participants performed the task over four runs. The first two runs ("Set 1") involved a set of six experimental (three rewarded, three nonrewarded) and two control players (one rewarded, one nonrewarded), intermixed, and the last two runs ("Set 2") involved a different set of six experimental and two control players, intermixed. Within a run, each experimental player repeated eight times (majority outcome status for six trials; 75%), for a total of 48 experimental trials per run. Each control player also repeated eight times, for a total of 16 control trials per run. Accordingly, across the four runs, there were 192 experimental trials (majority outcome status for 144 trials; 75%) and 64 control trials. The presentation order of the runs was quasirandomized for each participant, keeping pairs of runs that formed sets together (i.e., set1a, set1b, set2a, set2b; set2b, set2a, set1b, set1a, etc., with the letters referring to the stimulus order). The assignment of a given player as rewarded or nonrewarded was counterbalanced across participants.

All stimuli were presented using a PC computer (Lenovo ThinkPad) with E-Prime (Version 2.0) and an MRI-compatible projector and screen. Participants made their responses using an MRI-compatible box placed in their right hand.

In order to familiarize participants with the materials and procedure, immediately prior to the scan, participants were provided with the task instructions and completed practice trials, using a regular keyboard, with a different set of stimuli, in a private testing room. Participants completed additional practice trials with these same practice stimuli in the MRI scanner (during the MP-RAGE scan) to help them acclimate to the scanning environment and the button box used to make responses.

**Debrief** Finally, participants were debriefed about the task, which ensured that no participants had difficulty seeing the screen, using the button box, or felt rushed during the task. Participants were also asked about the strategies they used in the task and how they felt they performed on the task.<sup>3</sup>

 $<sup>^2</sup>$  It should be noted that inclusion of control trials in this task increases the stimulus load, which may have made the task more difficult. Although the outcome for the player is provided on the control trials—hence, there is nothing new to be learned from the feedback per se—some observational learning may nonetheless have taken place, again, potentially resulting in greater task difficulty.

<sup>&</sup>lt;sup>3</sup> After the scan, participants also completed a test phase, wherein the experimental players from the learning phase were presented side by side and participants made responses with no feedback provided. These data are not presented in this paper and will not be discussed further.



Fig. 1 Schematic of the value-based learning paradigm. In the actual experiment, the stimuli were presented in color (Color figure online)

#### Image acquisition

Images were collected on a 3.0 Tesla Siemens Prisma scanner equipped with a 64-channel head coil and located at the Jamaica Plain campus of the VA Boston Healthcare System. A high-resolution T1-weighted magnetization-prepared rapid gradient-echo (MP-RAGE) sequence was acquired in the sagittal plane (TR = 2,530 ms, TE = 3.35 ms, TI = 1,100 ms, flip angle =  $7^{\circ}$ , sections = 176, slice thickness = 1 mm, matrix =  $256^2$ , FOV = 256 mm, voxel size = 1 mm<sup>3</sup>). Four whole-brain task-based functional scans were acquired parallel to the anterior-posterior commissural plane using a multiband echo-planar imaging (EPI) sequence sensitive to the blood oxygenation level-dependent (BOLD) signal (e.g., Moeller et al., 2010): multiband = 6; TR = 1,067 ms, TE = 34.80 ms, flip angle =  $65^{\circ}$ , slices = 72, slice thickness = 2 mm, FOV = 208, matrix =  $104^2$ , voxel size = 2 mm<sup>3</sup>, volumes = 388, phase encoding = anterior-posterior). To correct for image distortion, a brief scan using the same parameters was also acquired, although the phase encoding direction was inverted (posterior-anterior). Two additional whole-brain resting-state scans (before and after the task-based runs) were collected but were not analyzed and are not discussed further.

#### Data processing and analyses

All analyses discussed below (behavioral and fMRI) pertain to experimental trials only; the control task will not be discussed further. Behavioral For analysis of accuracy, as in other papers (e.g., Shohamy, Myers, Kalanithi, & Gluck, 2008), we considered a response correct based on the *majority* outcome status for a given player. That is, if a participant offered the majority outcome response on a minority trial, that trial was scored as correct. Accordingly, the maximum score for a given participant is 100%. Although missed trials were somewhat rare (see Results), the denominator for the accuracy calculation was based on valid trials (i.e., the total number of trials on which the participant responded). Our main fMRI analyses (discussed below) pertain to overall learning (i.e., collapsed across runs), but we also report learning as a function of stage (early vs. late runs), as one of our fMRI analyses pertained to this comparison. For each individual participant, we used a binomial test to calculate whether his or her performance was above chance (50%), based on valid trials. Mean reaction-time data for correct and incorrect responses are also reported.

**FMRI** Functional imaging data were preprocessed and analyzed using FEAT (FMRI Expert Analysis Tool) Version 6.00, part of FSL (FMRIB's Software Library, www.fmrib. ox.ac.uk/fsl). FSL's topup tool was used to estimate susceptibility fields. Images were motion corrected using MCFLIRT (Jenkinson, Bannister, Brady, & Smith, 2002). Next, an estimated susceptibility field correction was applied to the functional time series using applytopup. The BOLD time series was skull stripped using FSL's Brain Extraction tool (BET) and bias-field corrected using FMRIB's Automated Segmentation Tool (FAST). Subsequent fMRI

data processing was carried out using the following prestatistics: Spatial smoothing was performed using a Gaussian kernel of FWHM 5 mm and grand-mean intensity normalization of the entire 4D data set by a single multiplicative factor. ICA-AROMA (a robust ICA-based strategy for removing motion artifacts from fMRI data; Pruim et al., 2015) was used to identify and remove additional motion components. The data were then high-pass temporal filtered (Gaussian-weighted least-squares straight line fitting, with sigma = 30.0 s).

Next, in a two-step registration process, each functional image was coregistered to the participant's same-session T1weighted structural image using FMRIB Linear Image Registration Tool (FLIRT). Between-subject registration was accomplished by alignment of functional images to the MNI152 standard-space template and further refined using the FMRIB Nonlinear Image Registration Tool (FNIRT). Images for each run for each participant were visually inspected to confirm proper registration to MNI space. Time-series statistical analysis was carried out using FILM with local autocorrelation correction (Woolrich, Ripley, Brady, & Smith, 2001). Trial onset times were convolved with a double gamma hemodynamic response function, modeled with the entire trial duration (3.6 s, which included the total time for player onset, response, and outcome for each trial).<sup>4</sup> Subject-level analysis was carried out using a fixed-effects model in FLAME (FMRIB's Local Analysis of Mixed Effects; Beckmann, Jenkinson, & Smith, 2003; Woolrich, Behrens, Beckmann, Jenkinson, & Smith, 2004). The general linear model (GLM) consisted of task regressors for each level of the experimental condition (i.e., correct and incorrect responses for rewarded and nonrewarded stimuli) and additional regressors of no interest, which included control trials and trials in which no response was made.

At the third level, a series of whole-brain and ancillary region-of-interest group-level analyses were carried out using FLAME Stage 1. The resulting statistical images were compared using paired *t* tests and using a cluster-defining threshold of Z > 3.09 (i.e., p < .001) and a corrected cluster significance threshold of p = .05 (Eklund, Nichols, & Knutsson, 2016). Given the known uncertainty about regional specificity when a given cluster comprises multiple regions (Woo, Krishnan, & Wager, 2014), and given our a priori interest in the hippocampus and ventral striatum (particularly the nucleus accumbens; NAcc), we performed, when relevant, follow-up

targeted analyses that included only a single binarized regions-of-interest (ROI) mask of the bilateral hippocampus and NAcc in the analysis (Harvard-Oxford Subcortical Structural Atlas, 50% threshold; see Fig. 2), using a familywise error (FWE) voxel-wise correction of p = .05. Although we observed very similar results using a cluster-based correction approach with this ROI analysis, for small volumes, a voxel-wise threshold is thought to be advantageous over a cluster-based approach, as clusters often extend beyond ROI boundaries (Roiser et al., 2016). The purpose of this ROI analysis was to firmly localize our whole-brain effects to these hypothesized regions. The statistical approach and use of a single ROI mask were based on recent recommendations from Roiser et al. (2016). Notably, the ROI mask included the whole hippocampus proper, but excluded the MTL cortices and amygdala. An additional targeted analysis, described below, statistically examined whether our observed effects localized to the anterior (versus posterior) hippocampus per se.

Following the literature on reinforcement learning more broadly (Davidow et al., 2016; Foerde & Shohamy, 2011; Li et al., 2011), the primary contrast of interest was correct versus incorrect, which we hypothesized would elicit activation in anterior hippocampus and the NAcc. The opposite contrast was also examined (incorrect vs. correct). Notably, here, correct and incorrect trials are calculated from the point of view of the feedback provided to a participant, not based on the player's majority outcome status as per above (e.g., if the participant responded "yes" to a typically rewarded player, but, on that particular trial, the player did not win, then that trial would be coded as "incorrect").

To formally implement the comparison of anterior versus posterior hippocampus, we next split the Harvard-Oxford hippocampal anatomical mask at the level of the uncal apex (i.e., at y = -21 mm) into anterior and posterior parts (Poppenk et al., 2013). Then, for each hemisphere, we extracted averaged parameter estimates (i.e., averaged across all voxels in the mask) from the relevant contrast of parameter estimate (COPE) images for each participant at the second level; these data were inputted into a 2 (hemisphere [left, right]) × 2 (region [anterior, posterior]) repeated-measures ANOVA in SPSS, with the threshold set to p < .05.

In an exploratory fashion, we also examined whether neural responses in the NAcc and hippocampus for correct versus incorrect differed as a function of whether the player was typically rewarded or not rewarded; we performed a 2 (correct, incorrect)  $\times$  2 (rewarded, nonrewarded) *F* test at the third level in FSL, using the abovementioned ROI mask.

Having established robust activation in the NAcc and hippocampus for the correct versus incorrect contrast (see below), we next examined whether this pattern of activity differed as a function of learning phase, by comparing activation for correct versus incorrect as a function of early versus late learning runs using the abovementioned ROI mask. This comparison

<sup>&</sup>lt;sup>4</sup> To separate these phases would require a jittered epoch between the response and outcome, which would impose a necessary delay in the arrival of the outcome. However, given a number of studies showing hippocampal involvement in delayed reinforcement learning (see Palombo & Verfaellie, 2017, for review), we did not opt for such a design, as we wanted to observe whether hippocampal effects occur independently of delay in feedback. Our approach deviates from that of Foerde & Shohamy (2011; described in the Discussion), in which the authors analyzed data from the feedback epoch.



**Fig. 2** Results from Experiment 1a. **a** Brain images depicting activation in the bilateral hippocampus (HPC; top) and nucleus accumbens (NAcc; bottom) for the whole-brain contrast of correct versus incorrect. For display purposes only, percentage of signal change is shown for the left (pink) and right (blue) hippocampus and NAcc. Percentage of signal change was calculated by extracting the peak from each structure for correct and incorrect responses using the COPE images from the second level and a corrected scale factor (i.e., 100\*baseline-to-max range). Using

was first implemented in FSL at the second level by coding early and late runs as 1 and -1, respectively. Although some literature suggests greater contribution of the hippocampus early in learning (and vice versa for the striatum; Dickerson et al., 2011; Fera et al., 2014; Poldrack et al., 2001; Poldrack & Packard, 2003; Shohamy et al., 2008), other work does not support this notion (see Delgado, Miller, Inati, & Phelps, 2005; Shohamy et al., 2008). Accordingly, we did not make specific predictions about the nature of changes across learning but perform such analyses only to align our work with that of others in the literature.

# Results

#### **Behavioral**

On average, participants responded on 96.0% of trials (SD = 4.9%). Mean accuracy was 65.3% (SD = 7.9%). A paired *t* test comparing performance in early versus late runs showed a significant increase in accuracy across learning (early: 63.6%, SD = 7.8%; late: 67.1%, SD = 9.6%), t(29) = 2.58, p = .015, Cohen's *d* (using pooled variance) = 0.40. Three participants performed at or below chance level, but the pattern

an isolated 3-s long double-gamma hemodynamic response function, the baseline-to-max range was set at 0.587. A 3-D render of the Harvard-Oxford masks used to extract data from these structures is shown. **b** Brain images depicting activation in the bilateral insula and medial prefrontal cortex for the whole-brain contrast of incorrect versus correct. For the analyses depicted in **a** and **b**, a cluster-defining threshold of Z > 3.09 (i.e., p < .001) and a corrected cluster significance threshold of p = .05 was used. (Color figure online)

did not change when these three participants were removed from the analysis (p = .02).

On average, participants took 993.4 ms (SD = 137.4) to respond on trials in which they were correct, and 1,075.8 ms (SD,=,134.2) to respond on trials in which they were incorrect; the difference in reaction time on incorrect versus correct trials (M = 82.4 ms) was statistically significant, t(29) = 7.0, p < .0001, Cohen's *d* (using pooled variance) = 0.61, though negligible in terms of differences in fMRI signal.

#### FMRI

**Correct versus incorrect** The results for the contrast of correct versus incorrect (and vice versa) are displayed in Fig. 2a and 2b and Table 2. For the correct versus incorrect contrast, BOLD response differences were observed in expected regions, including striatal and MTL structures, as well as the ventromedial prefrontal cortex. These results did not change when we excluded the three participants who did not perform above chance level.

Critically, ancillary ROI analyses confirmed strong BOLD response localized to the bilateral nucleus accumbens (left peak: -10, 6, -8; right peak: 12, 8, -10), bilateral anterior hippocampus (left peak: -28, -16, -16; right peak: 22, -12,

-20), and left posterior hippocampus (left peak: -30, -32, -8). Our exploratory analysis examining whether BOLD response differences for correct versus incorrect varied as a function of whether the player's majority outcome status was rewarded or nonrewarded failed to reveal any significant interaction effect. Notably, no main effect of reward versus nonreward was observed in either of these ROIs.<sup>5</sup>

The contrast of incorrect versus correct showed no activation anywhere in the basal ganglia or MTL, but showed robust activation in other brain regions, including the bilateral insula and a more dorsal part of the medial prefrontal cortex (see Table 1 for the full list of regions; also see Fig. 2b).

**Correct versus incorrect across learning (early vs. late)** We did not observe any significant differences in the patterns of activation as a function of learning block (early vs. late). Still, we urge caution in interpreting this null effect, as this may be due to low power (i.e., because the data are split in half) or due to a minimal increase in performance from early to late learning (see Behavioral Results).

**Correct versus incorrect across the long axis** A comparison of differences along the long axis of the hippocampus in correct versus incorrect BOLD response revealed stronger activation in the anterior portion of the hippocampus, bilaterally, relative to the posterior, F(1, 29) = 16.5, p = .0003,  $\eta^2 = .36$ , with no main effect of hemisphere (p = .81,  $\eta^2 = .002$ ) or interaction with hemisphere (p = .17,  $\eta^2 = .064$ ; see Fig. S1 in the Supplementary Materials).

# **Experiments 2a and 2b**

## Methods

Having established hippocampal involvement in the task used in Experiment 1, we next administered a very similar task that involved learning the reward contingencies for six players to amnesic patients and well-matched healthy controls (Experiment 2a). To examine performance under conditions of reduced memory load, in a separate session, we administered another version of the task to amnesics and a new set of healthy controls (Experiment 2b), with four players instead of six. The decision to reduce the load to four players was motivated by prior work, in which *intact* performance was observed in amnesic patients in a trial-and-error learning task that involved acquiring the contingencies for only four stimuli (Foerde et al., 2013); in light of those results, it follows that any observed deficit in the present study in the four-player version would unlikely be due to memory load per se.

## Participants

Patients In Experiment 2a, seven patients with amnesia (one female) secondary to MTL damage participated (see Table 2 for demographic and neuropsychological data). An eighth amnesic patient was excluded from all analyses because this patient had a substantial number of missed responses (17%), resulting in less overall exposure to the player outcomes. The neuropsychological profile for each patient indicated severe impairment that was limited to the domain of memory. Etiology of amnesia included hypoxic-ischemic injury secondary to either cardiac or respiratory arrest (n = 3), stroke (n = 2), encephalitis (n = 1), and status epilepticus followed by left temporal lobectomy (n = 1). Lesions for six of the seven patients are presented in Fig. 3, either on MRI or CT images. P4, who had suffered from cardiac arrest, could not be scanned due to medical contraindications and is thus not included in the figure. MTL pathology for this patient was inferred based on etiology and neuropsychological profile. Of the patients with available scans, two patients (P3, P5) had lesions that were restricted to the hippocampus, one patient (P7) had a lesion that included the hippocampus as well as the amygdala (see below), one patient (P1) had a lesion that included the hippocampus and MTL cortices, and one patient (P2) had a lesion that extended well beyond the medial portion of the temporal lobes into the anterolateral temporal neocortex (due to the temporal lobectomy). For the patient whose etiology was encephalitis (P6), clinical MRI was acquired, but only in the acute phase of the illness, with no visible lesions observed on T1-weighted images. However, T2-flair images demonstrated bilateral hyperintensities in the hippocampus and MTL cortices as well as the anterior insula. Hence, across all patients with available information, the hippocampus was the only area of overlap. As shown in Table 2, volumetric data for the hippocampus and MTL cortices was available for four of the seven patients (P2, P3, P5, P7), using methodology reported elsewhere (see Kan, Giovanello, Schnyer, Makris, & Verfaellie, 2007 for methodology).

Due to the known involvement of the amygdala and basal ganglia structures in motivational processes, for patients P3, P5, and P7, for whom reliable extra-hippocampal subcortical volumetric data could be obtained (see Supplementary Materials), we quantified the volume of the amygdala, caudate, putamen, pallidum, and nucleus accumbens using an automated pipeline (FreeSurfer) that has been employed in amnesic patients in other studies (Baker et al., 2016; Sheldon, Romero, & Moscovitch, 2013). No significant volume loss was observed in any of these structures, with the exception of the right amygdala, which was significantly

<sup>&</sup>lt;sup>5</sup> For completeness, we also compared rewarded versus nonrewarded trials at the whole-brain level; this contrast revealed activation only in the fusiform gyrus, bilaterally (left peak: -26, -68, -12; right peak: 24, -72, -8). This pattern of activation is to be expected, given the greater sensory input associated with the dollar bill image (vs. the gray rectangle image; see Methods); the opposite contrast (nonrewarded vs. rewarded) failed to reveal any significant effects.

Table 1	Regions of fMRI	Activation	in Experiment 1	l
---------	-----------------	------------	-----------------	---

Extent (voxels)	z score	MNI coordinates (mm)			Side	Region	Brodmann area	
		x	У	z				
Correct > Incorrect								
1,119	5.50	-8	32	-10	L	Ventromedial prefrontal cortex	24	
875	5.54	16	6	-10	R	Putamen		
						*Includes bilateral NAcc, R anterior	hippocampus	
850	5.01	14	-28	66	R	Postcentral gyrus	3	
486	4.74	62	-2	-6	R	Superior temporal gyrus	22	
397	5.00	-16	-4	-24	L	Parahippocampal gyrus	34	
						*Includes L anterior hippocampus		
376	4.00	38	-22	48	R	Postcentral gyrus	3	
215	4.42	-66	-34	10	L	Superior temporal gyrus	22	
147	3.96	34	-48	64	R	Superior parietal lobule	7	
Incorrect > Correct								
3,402	6.32	8	20	62	R	Superior medial prefrontal cortex	6	
820	4.97	34	28	0	R	Insula	13	
762	5.36	-38	20	-12	L	Inferior frontal gyrus and insula	47	
625	5.39	44	10	52	R	Middle frontal gyrus	6	
448	4.94	6	-24	0	R	Thalamus		
350	4.73	34	-52	-34	R	Cerebellum		
236	4.64	-34	-58	-28	L	Cerebellum		
192	4.73	-42	20	48	L	Middle frontal gyrus	6	
136	3.96	48	-52	48	R	Inferior parietal lobule	40	
122	4.21	-30	50	24	L	Superior frontal gyrus	9	
118	4.50	50	-34	-4	R	Middle temporal gyrus	21	
112	3.80	40	28	20	R	Middle frontal gyrus	9	

Note. MNI = Montreal Neurological Institute; L = left; R = right; NAcc = nucleus accumbens

#### Table 2 Patient Information for Experiment 2

Experiment	Patient	Etiology	Age	Edu	WAI	WAIS-III		WMS-III		Volume loss (%)	
			(years)	(years)	VIQ	WMI	GM	VD	AD	Hippocampal	Subhippocampal
2a, 2b	P1	Hypoxic ischemic	67	12	88	75	52	56	55	N/A	N/A
2a, 2b	P2	Status epilepticus + Left temporal Lobectomy	53	16	93	94	49	53	52	63%	60% <sup>a</sup>
2a, 2b	Р3	Hypoxic ischemic	61	14	106	115	59	72	52	22%	_
2a, 2b	P4	Hypoxic ischemic	65	17	131	126	86	78	86	N/A	N/A
2a	Р5	Stroke	64	18	117	88	67	75	55	62%	_
2a, 2b	P6	Encephalitis	75	13	99	104	49	56	58	N/A	N/A
2a, 2b	P7	Stroke	53	20	111	99	60	65	58	43%	-

*Note.* WAIS-III = Wechsler Adult Intelligence Scale, Third Edition; WMS-III = Wechsler Memory Scale, Third Edition; VIQ = verbal intelligence quotient; WMI = working memory index; GM = general memory; VD = visual delayed; AD = auditory delayed; N/A = not available. Age is represented at the time of Experiment 2a; Experiments 2a and 2b were completed within approximately 1 year of each other (minimum time between sessions = 167 days).

<sup>a</sup> Volume loss in left anterior parahippocampal gyrus (i.e., entorhinal cortex, medial portion of the temporal pole, and the medial portion of perirhinal cortex; see Kan et al., 2007, for methodology)



**Fig. 3** Structural MRI and CT scans depicting medial temporal lobe (MTL) lesions for six of the seven amnesic participants. The left side of the brain is displayed on the right side of the image. CT slices show lesion location for P1 in the axial plane. T1-weighted MRI images depict lesions for P2, P3, P5, and P7 in the coronal and axial plane. T2-flair MRI images depict lesion locations for P6 in the axial plane

smaller in P7, as noted above. (Given the size of P2's lesion, which would likely deem the automated segmentation unreliable, we opted not to include his data in this analysis.)

For Experiment 2b, six of the seven amnesic patients from Experiment 2a participated and are indicated in Table 2. Patient P5 was not available due to long-term personal commitments.

**Healthy controls** For Experiment 2a, 16 healthy control participants (eight female) were matched to the patient group in age (60.9 years, SD = 10.5), education (15.8 years, SD = 2.4 years), and verbal IQ (110.4, SD = 16.2), which was assessed

with the Wechsler Adult Intelligence Scale–Third Edition (Wechsler, 1997).

For Experiment 2b, a new group of 12 healthy control participants (three female) were matched to the patient group in age ( $60.8 \pm 7.61$  years), education ( $15.4 \pm 2.5$  years), and verbal IQ ( $112.1 \pm 13.4$ ).

All participants provided informed consent in accordance with the Institutional Review Board at the VA Boston Healthcare System.

# Materials and procedure

For Experiment 2a, the task was modeled after the one used in Experiment 1, with the following modifications for behavioral testing of amnesic patients (also see Supplementary Materials for task instructions): Participants were given more time to make a response (4,000 ms); a fixed, rather than jittered, intertrial interval (2,667 ms) was used; and the control condition was eliminated. Finally, participants were given only one set of six players (three rewarded, three nonrewarded), which were administered over three learning blocks, providing more overall repetitions of the players relative to Experiment 1 (i.e., a greater opportunity to learn the contingencies for a given player; with a total of 24 presentations of each player).

As in Experiment 1, within a block, each player was presented eight times (majority outcome status for six trials; 75%), for a total of 48 intermixed trials per block. Accordingly, across the three blocks, there were 144 trials. There were three presentation orders of blocks (a-b-c; b-c-a; c-a-b), which were randomly assigned to participants in each group so that each counterbalance order was represented approximately equally in the two groups. Moreover, the assignment of a given player to the rewarded or nonrewarded condition was counterbalanced across participants. As in Experiment 1, the task was preceded by a practice phase consisting of six trials (with separate stimuli) and was followed by a test phase (not discussed) and debriefing. To determine whether patients could acquire the stimulus contingencies by the end of learning, we compared performance between amnesic patients and controls in the last learning block.

To ensure that amnesic patients had no trouble distinguishing the players from each other, in a separate session we performed a perceptual discrimination control task using the players from Experiment 2a, for which amnesic patients performed very well (see Supplementary Materials).

For Experiment 2b, the methods (including counterbalancing) were identical to those of Experiment 2a, except this version included only four players (two rewarded, two nonrewarded) and was administered over two learning blocks of 48 trials each (with a total of 12 presentations of each player). A new set of stimuli was used in Experiment 2b (see Fig. S2 in the Supplementary Materials). As in

Experiment 2a, we compared performance between amnesic patients and controls in the last learning block.

#### Results

In Experiment 2a, amnesic patients responded on average on 97.8% of trials (SD = 0.8%) and control participants on 99.5% of trials (SD = 0.8%), suggesting that participants had sufficient time to make a response. As in Experiment 1, accuracy was calculated according to majority outcome status, and the mean accuracy for each group across the last learning block is shown in Fig. 4; the figure also shows performance for each individual amnesic patient (also see Table S1 in the Supplementary Materials). Patients showed a significant impairment in learning, t(21) = 3.77 p = .001, Cohen's d = 1.91. Notably, at the individual level, all seven of the patients (100%) were at or below chance, whereas only three (18.8%) control participants were at or below chance.

In Experiment 2b, amnesic patients responded on average on 98.1% of trials (SD = 0.8%) and controls on 99.7% of trials (SD = 0.5%). The mean accuracy for each group is shown in Fig. 4 (also see Table S1 in the Supplementary Materials). Patients showed a significant impairment in learning, t(16) =2.31 p = .035, Cohen's d = 1.12. At the individual level, six out of seven patients (86%) were at or below chance; the remaining patient (P6) performed quite well (83%) and was significantly above chance. By contrast, three (25%) of the control participants were at or below chance.



**Fig. 4** Results from Experiment 2a and 2b. The plot depicts mean accuracy (with standard error of the mean) for amnesic patients (filled circle) and healthy controls (filled square). Each individual patient is shown with an open circle. Accuracy was defined according to majority outcome status of each player (see Method section). Note that in experiment 2b, one patient performed above chance (P6)

Ancillary analyses for Experiment 2a and 2b examining performance as a function of reward outcome are presented in the Supplementary Materials. It is worth noting that although this critical final block involved the same number of stimuli (N= 48) across the two experiments (which allowed us to set chance at the same level in the two experiments using the binomial distribution test), by necessity, the number of exposures to each player differed (eight exposures per player in Experiment 2a, and 12 exposures per player in Experiment 2b). Nonetheless, when approximately matching across experiments, the number of exposures to each player in the final block (i.e., by increasing the number of stimuli included in the block analyzed in Experiment 2a to 72 trials), the same pattern of results was observed.

# **General discussion**

The goal of the present study was to clarify the role of the hippocampus in value-based learning. First, using fMRI, we showed strong engagement of bilateral hippocampus, alongside the expected recruitment of striatal regions (e.g., NAcc) and ventromedial prefrontal cortex (Experiment 1). The hippocampal finding was a prerequisite for asking next whether the hippocampus is critical for value-based learning. The latter was demonstrated in Experiment 2a, in which we showed that amnesic patients with MTL lesions, and some with lesions limited to the hippocampus, failed to learn the value-based contingencies in this task. We replicated the effect in these same amnesic patients under conditions of reduced memory load (Experiment 2b). Taken together, the current results provide compelling converging evidence that the hippocampus is required for value-based learning.

Our findings align well with prior fMRI work demonstrating that the hippocampus is engaged during various forms of reward learning (see Introduction) and with converging evidence from rodent work showing strong modulation of hippocampal neurons by reward information during learning (Lee et al., 2012). This modulation is likely supported through a dynamic interplay of dopamine projections between midbrain, striatum, and hippocampus (Groenewegen, Vermeulen-Van der Zee, te Kortschot, & Witter, 1987; Kelley & Domesick, 1982; Lisman & Grace, 2005). This interplay may be similar to that responsible for effects of value on episodic memory (e.g., Adcock et al., 2006).

In considering a role for the hippocampus in value-based learning, it is interesting to compare our findings to prior work by Foerde et al. (2013) that examined reinforcement learning. In that study, amnesic patients (some of whom participated in the present study) were asked to determine, through trial and error, which flower each of four butterflies preferred. As in the present study, the contingencies were probabilistic, and participants received feedback (correct vs. incorrect) for their choices. Thus, the task demands were quite similar, particularly to the four-player version we used in Experiment 2b. Foerde and colleagues showed that patients performed as well as healthy controls under conditions in which the feedback was delivered immediately (as in our task); moreover, in an fMRI version of the task, the hippocampus was not engaged under such conditions (Foerde & Shohamy, 2011). An intriguing difference between our task and that of Foerde et al. is that whereas in our task learning involved mapping stimulusvalue contingencies, in Foerde et al., learning involved mapping stimulus-stimulus contingencies. That is, in their study, there was no value component. Notwithstanding the limitations of cross-study comparisons, this bolsters the idea that reward information per se may be relevant to eliciting hippocampal engagement and may be a critical mechanistic feature underlying our results. Our findings also raise the possibility that prior findings showing impaired reinforcement learning in amnesic patients (see Introduction) may have been due not only to the complexity of the stimuli but also to the inclusion of a value component (Hopkins et al., 2004).

Notably, our fMRI data showed stronger recruitment of the anterior relative to the posterior portion of the hippocampus, a finding that aligns well with the notion that the anterior hippocampus (ventral hippocampus in rodents) is more critical for motivational, affective, or value-based aspects of cognition, likely due to stronger anterior relative to posterior hippocampal projections with the NAcc (Groenewegen et al., 1987; Kelley & Domesick, 1982), as well as the amygdala and ventromedial prefrontal cortex (reviewed in Poppenk et al., 2013).<sup>6</sup> Altogether, these findings and the existing literature provide support for the idea that value learning per se may be a factor that elicits hippocampal involvement.

In the present task, we examined hippocampal involvement when participants learned about the *value of stimuli* (in this case, whether the stimulus player wins money or does not), whereas in other tasks, participants learn what types of choices lead to a *valuable response* (i.e., participants are rewarded for their correct choices about stimuli that themselves do not have value attached to them). Our focus on the learning of stimulus value allowed us to orthogonalize reward outcome from accuracy of the participant's response. That is, in our task, the feedback provided to the participant emphasized the outcome for the player and was thus independent of whether the participant made a correct response. Here, we showed that the hippocampus was not sensitive to the presence of valuable stimuli per se (i.e., rewarded vs. nonrewarded trials), but rather, was sensitive to learning in the context of value-based stimuli (i.e., correct vs. incorrect trials). In apparent contrast to our findings, Delgado et al. (2000) demonstrated that the MTL is modulated by value-based information ("wins" vs. "losses"), even in a task that does not have explicit learning demands-suggesting, contrary to our findings, that the MTL is sensitive to the mere presence of reward. Yet it is important to note that the task used in Delgado et al.'s study was not completely devoid of learning, in that participants could still acquire information about long-term probabilities of value over time.

It is nonetheless important to note that our study design rendered a more complex feedback prescription relative to other paradigms used in prior work. Is it possible that the hippocampus was needed to resolve the ambiguity in our task between player and participant outcome? Relevant to this issue are the fMRI results: If this were the case, one would expect an interaction between the reward status of the player and participant outcome in the hippocampus. That is, one would expect the hippocampus to be most strongly engaged in these "incongruent" scenarios (i.e., when the player is rewarded but the participant gets the trial incorrect, and when the player is not rewarded but the participant gets the trial correct). However, no such interaction was observed. These findings fail to provide supporting evidence that our main hippocampal effects are driven by task complexity.

An alternative account, recently put forth in the literature, is that the hippocampus performs a more domain-general computation that is not specific to reward. Relevant to this idea, Ballard, Wagner, and McClure (2018) have suggested that hippocampal-based pattern separation mechanisms (Leutgeb, Leutgeb, Moser, & Moser, 2007) may support conjunctive coding in tandem to a more basic reinforcement learning system that is striatal (Ballard et al. 2018; also see e.g., Floresco, 2007, for a discussion of related ideas). To test this idea, the authors examined hippocampal and striatal engagement, via fMRI, during a probabilistic stimulus-value learning task that involved stimuli that have overlapping features (e.g., AB+, B -, AC-, C+). Based on hippocampal similarity patterns, the authors showed that the hippocampus formed conjunctive representations that facilitated value-based learning by influencing striatal-based prediction errors-a finding that fits with the conceptualization that the hippocampus entrains the striatum (Bornstein, Khaw, Shohamy, & Daw, 2017). Notably, other recent work suggests that such conjunctive coding also occurs in non-value-based feedback learning: Duncan, Doll, Daw, and Shohamy (2018) showed hippocampal engagement associated with the use of configural information during

<sup>&</sup>lt;sup>6</sup> Such long-axis considerations have not been addressed in prior fMRI studies of value-based learning, although they have received attention in studies of episodic and spatial learning. In this literature, the focus has been on gradient-based differences across the long axis in terms of mnemonic specificity, with the anterior and posterior hippocampi implicated in gist-level and detail-level processing, respectively (reviewed in Poppenk et al., 2013; Sheldon & Levine, 2016). Although these models are not necessarily mutually exclusive to a motivational one (see Sheldon & Levine, 2016), consistent with this alternative view, it is possible that probabilistic trial-and-error learning, wherein information is accrued over repeated trials, is more likely to recruit gist-based anterior hippocampal processes to facilitate extraction of the global regularities of the contingencies (i.e., which players win *most of the time*), whereas the posterior hippocampus may be more engaged when specific details from discrete episodes are more relevant.

reinforcement learning. Such hippocampal involvement was observed even when configural processing was not required for learning per se (Duncan et al., 2018).

Can such an explanation account for our findings? In contrast to the abovementioned Hopkins et al. (2004) amnesia study, our study used one-to-one stimulus-value mappings (i.e., there was no requirement to incorporate multiple stimuli into learning, hence limiting the demands on configural processing). Nonetheless, it is possible that in the absence of explicit conjunctive-coding demands, the use of fractal stimuli (which can share shape- and color-based features with one another) augmented the involvement of hippocampal-based pattern separation processes in our task. This explanation could help explain why an intact striatal system was insufficient to support what appeared to be basic stimulus-response learning in amnesic patients. It also provides an alternative explanation for the divergent results of the present study from those of Foerde et al. (2013)-namely, that it is the complexity of the stimuli (fractal patterns in the present study vs. plain colors in Foerde et al.) and the resulting pattern separation demands that drive hippocampal engagement, as opposed to value information per se.

Although this post hoc explanation is appealing, it does not conform with some observations that the hippocampus is activated in value-based learning tasks that use one-to-one stimulus-value mappings that include simple stimuli, such as monotone shapes (Dickerson et al., 2011; also see Li et al., 2011). Based on the findings to date, it is possible that multiple mechanisms are at play-namely, that the hippocampus is engaged when the task draws on pattern separation mechanisms, and it is also sensitive to learning about value-based information above and beyond its role in pattern separation. The precise mechanism by which the hippocampus contributes to value-based learning, and how its contribution differs from the striatum, remains to be further elucidated. Relevant to this topic, it will be important for future research to ascertain whether value signals are computed in house in the hippocampus versus propagated from elsewhere (also see Lee et al., 2012).

In interpreting our results, we also considered whether hippocampal involvement in this task might simply be due to the influence of declarative memory. Our task was probabilistic and involved learning from feedback—conditions thought to maximize nondeclarative learning (as this learning is historically considered incremental [habitual] in nature—i.e., learning without awareness; Squire, 2004). Yet it is important to consider that in healthy individuals no task is process pure, and we cannot rule out the possibility that participants had explicit knowledge about stimulus contingencies during learning (see Gluck, Shohamy, & Myers, 2002, for further discussion). Related to this idea is the possibility for involvement of episodic or relational processes in influencing performance—processes that are known to depend on the hippocampus (Cohen, Poldrack, & Eichenbaum, 1997; Eichenbaum, Yonelinas, & Ranganath, 2007). The prevailing idea in prominent reinforcement learning models is that participants create a running average of rewards accrued for a given action, and that this average is updated incrementally as learning ensues. Yet accumulating evidence suggests that episodic or relational processes play a role in value-based reinforcement learning, even when there is no explicit task demand to use such processes, and even when participants are unaware of the use of these processes (Bornstein et al., 2017; Wimmer, Daw, & Shohamy, 2012). For example, Bornstein et al. (2017) recently showed that an episodic memory model (one in which participants sample individual trial memories) better fit choices in a probabilistic value-based learning task than did a classic incremental learning model. Other work suggests that participants incidentally incorporate relational structure into their choice behavior-a phenomenon supported by functional coupling between the striatum and hippocampus (Wimmer et al., 2012).

Still, an important piece of evidence that speaks against either a declarative or an episodic or relational explanation comes from the Foerde et al. (2013) findings described above. Given the similarities between our tasks, there is no obvious reason that the demands on declarative or episodic/relational memory would be larger in our study as compared with Foerde et al. On the surface, the demands on explicit memory should, if anything, be greater in Foerde et al., as the stimuli were more easily verbalizable (i.e., they involved solid, basic colors such as "blue," whereas we used fractal-like patterns; see Fig. S2 in the Supplementary Materials), yet in such a case the hippocampus was neither necessary (Foerde et al., 2013) nor engaged (Foerde & Shohamy, 2011).

Other domain-general accounts of hippocampal contributions to value-based learning have also been proposed namely, that the hippocampus provides a temporal context signal (Howard & Eichenbaum, 2015; Palombo, Di Lascio, Howard, & Verfaellie, 2018; see Palombo & Verfaellie, 2017) or an internal model (Stachenfeld, Botvinick, & Gershman, 2017; also see Shohamy & Turk-Browne, 2013). Evidence suggests that the former is more relevant under conditions where feedback is delayed and the latter under conditions of multistep learning. Because neither of these conditions apply to the current task, it is not obvious how they provide an explanation of the hippocampal contribution observed here, although they may help explain task dissociations in other work (e.g., see Foerde & Shohamy, 2011; Foerde et al., 2013).

Although the precise mechanism is unclear, the observation that hippocampal and striatal systems were both engaged in our task provides another instance in which these systems may cooperate during learning. Such dual engagement calls for refinement of existing theoretical memory system models that postulate that these brain regions support dissociable aspects of learning and memory or even compete during learning. Future research is needed to determine the boundary conditions of hippocampal versus striatal involvement in such value-based learning and, crucially, the precise nature of their contributions to such learning. Nonetheless, the present findings highlight a broader role of the hippocampus in cognition than previously appreciated and may elucidate how the hippocampus contributes to goal-directed behaviors more broadly (Palombo, Keane, & Verfaellie, 2015; Shohamy & Turk-Browne, 2013).

Acknowlegements M.V. is supported by a Senior Research Career Scientist Award and Merit Award (I01CX000925) from the Clinical Science Research and Development Service, Department of Veterans Affairs, and a grant from NIH (RO1 MH093431). D.J.P. was supported by a postdoctoral fellowship from the Canadian Institutes of Health Research. D.J.P. is currently supported by start-up funds from the University of British Columbia. S.M.H. was supported by the Boston University Spivack Emerging Leaders in Neurosciences Award and The Ohio State University Discovery Themes Chronic Brain Injury Initiative. This work was further supported with resources and use of facilities at the Neuroimaging Research for Veterans Center, VA Boston Healthcare System. We thank Renee Hunsberger for research assistance. The content is solely the responsibility of the authors and does not necessarily represent the views of the U.S. Department of Veterans Affairs, the National Institutes of Health, or the United States Government. The authors have no conflicts of interest to report.

**Publisher's note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

# References

- Adcock, R. A., Thangavel, A., Whitfield-Gabrieli, S., Knutson, B., & Gabrieli, J. D. (2006). Reward-motivated learning: Mesolimbic activation precedes memory formation. *Neuron*, 50(3), 507–517.
- Baker, S., Vieweg, P., Gao, F., Gilboa, A., Wolbers, T., Black, S. E., & Rosenbaum, R. S. (2016). The human dentate gyrus plays a necessary role in discriminating new memories. *Current Biology*, 26(19), 2629–2634.
- Ballard, I. C., Wagner, A. D., & McClure, S. M. (2018). Hippocampal pattern separation supports reinforcement learning. https://doi.org/ 10.1101/293332
- Beckmann, C. F., Jenkinson, M., & Smith, S. M. (2003). General multilevel linear modeling for group analysis in fMRI. *NeuroImage*, 20(2), 1052–1063.
- Bornstein, A. M., Khaw, M. W., Shohamy, D., & Daw, N. D. (2017). Reminders of past choices bias decisions for reward in humans. *Nature Communications*, 8, 15958. https://doi.org/10.1038/ ncomms15958
- Callan, D. E., & Schweighofer, N. (2008). Positive and negative modulation of word learning by reward anticipation. *Human Brain Mapping*, 29(2), 237–249.

- Castel, A. D., Farb, N. A., & Craik, F. I. (2007). Memory for general and specific value information in younger and older adults: Measuring the limits of strategic control. *Memory & Cognition*, 35(4), 689–700.
- Cohen, N. J., Poldrack, R. A., & Eichenbaum, H. (1997). Memory for items and memory for relations in the procedural/declarative memory framework. *Memory*, 5(1/2), 131–178.
- Davidow, J. Y., Foerde, K., Galvan, A., & Shohamy, D. (2016). An upside to reward sensitivity: The hippocampus supports enhanced reinforcement learning in adolescence. *Neuron*, 92(1), 93–99.
- Delgado, M. R., Miller, M. M., Inati, S., & Phelps, E. A. (2005). An fMRI study of reward-related probability learning. *NeuroImage*, 24(3), 862–873.
- Delgado, M. R., Nystrom, L. E., Fissell, C., Noll, D. C., & Fiez, J. A. (2000). Tracking the hemodynamic responses to reward and punishment in the striatum. *Journal of Neurophysiology*, 84(6), 3072–3077.
- Dickerson, K. C., & Delgado, M. R. (2015). Contributions of the hippocampus to feedback learning. *Cognitive, Affective, & Behavioral Neuroscience*, 15(4), 861–877.
- Dickerson, K. C., Li, J., & Delgado, M. R. (2011). Parallel contributions of distinct human memory systems during probabilistic learning. *NeuroImage*, 55(1), 266–276.
- Duncan, K., Doll, B. B., Daw, N. D., & Shohamy, D. (2018). More than the sum of its parts: A role for the hippocampus in configural reinforcement learning. *Neuron*, 98(3), 645–657.
- Eichenbaum, H., Yonelinas, A. P., & Ranganath, C. (2007). The medial temporal lobe and recognition memory. *Annual Reviews of Neuroscience*, 30, 123–152.
- Eklund, A., Nichols, T. E., & Knutsson, H. (2016). Cluster failure: Why fMRI inferences for spatial extent have inflated false-positive rates. *Proceedings of the National Academy of Sciences*, 113(28), 7900–7905.
- Fera, F., Passamonti, L., Herzallah, M. M., Myers, C. E., Veltri, P., Morganti, G., ... Gluck, M. A. (2014). Hippocampal BOLD response during category learning predicts subsequent performance on transfer generalization. *Human Brain Mapping*, 35(7), 3122– 3131.
- Floresco, S. B. (2007). Dopaminergic regulation of limbic-striatal interplay. Journal of Psychiatry & Neuroscience 32(6), 400–411.
- Foerde, K., Race, E., Verfaellie, M., & Shohamy, D. (2013). A role for the medial temporal lobe in feedback-driven learning: Evidence from amnesia. *Journal of Neuroscience*, 33(13), 5698–5704.
- Foerde, K., & Shohamy, D. (2011). Feedback timing modulates brain systems for learning in humans. *Journal of Neuroscience*, 31(37), 13157-13167.
- Gluck, M. A., Ermita, B. R., Oliver, L. M., & Myers, C. E. (1997). Extending models of hippocampal function in animal conditioning to human amnesia. *Memory*, 5(1/2), 179–212.
- Gluck, M. A., Shohamy, D., & Myers, C. (2002). How do people solve the "weather prediction" task?: Individual variability in strategies for probabilistic category learning. *Learning & Memory*, 9(6), 408– 418.
- Groenewegen, H. J., Vermeulen-Van der Zee, E., te Kortschot, A., & Witter, M. P. (1987). Organization of the projections from the subiculum to the ventral striatum in the rat: A study using anterograde transport of Phaseolus vulgaris leucoagglutinin. *Neuroscience*, 23(1), 103–120.
- Hopkins, R. O., Myers, C. E., Shohamy, D., Grossman, S., & Gluck, M. (2004). Impaired probabilistic category learning in hypoxic subjects with hippocampal damage. *Neuropsychologia*, 42(4), 524–535.
- Howard, M. W., & Eichenbaum, H. (2015). Time and space in the hippocampus. *Brain Research*, *1621*, 345–354.
- Jenkinson, M., Bannister, P., Brady, M., & Smith, S. (2002). Improved optimization for the robust and accurate linear registration and motion correction of brain images. *NeuroImage*, 17(2), 825–841.
- Kan, I. P., Giovanello, K. S., Schnyer, D. M., Makris, N., & Verfaellie, M. (2007). Role of the medial temporal lobes in relational memory:

Neuropsychological evidence from a cued recognition paradigm. *Neuropsychologia*, *45*(11), 2589-2597.

- Kelley, A. E., & Domesick, V. B. (1982). The distribution of the projection from the hippocampal formation to the nucleus accumbens in the rat: An anterograde- and retrograde-horseradish peroxidase study. *Neuroscience*, 7(10), 2321–2335.
- Knowlton, B. J., Squire, L. R., & Gluck, M. A. (1994). Probabilistic classification learning in amnesia. *Learning & Memory*, 1(2), 106–120.
- Lee, H., Ghim, J. W., Kim, H., Lee, D., & Jung, M. (2012). Hippocampal neural correlates for values of experienced events. *Journal of Neuroscience*, 32(43), 15053–15065.
- Leutgeb, J. K., Leutgeb, S., Moser, M. B., & Moser, E. I. (2007). Pattern separation in the dentate gyrus and CA3 of the hippocampus. *Science*, 315(5814), 961–966.
- Li, J., Delgado, M. R., & Phelps, E. A. (2011). How instructed knowledge modulates the neural systems of reward learning. *Proceedings of the National Academy of Sciences*, 108(1), 55–60.
- Lighthall, N. R., Pearson, J. M., Huettel, S. A., & Cabeza, R. (2018). Feedback-based learning in aging: Contributions and trajectories of change in striatal and hippocampal systems. *Journal of Neuroscience*, 38(39), 8453–8462.
- Lisman, J. E., & Grace, A. A. (2005). The hippocampal-VTA loop: Controlling the entry of information into long-term memory. *Neuron*, 46(5), 703–713.
- Loh, E., Kumaran, D., Koster, R., Berron, D., Dolan, R., & Duzel, E. (2016). Context-specific activation of hippocampus and SN/VTA by reward is related to enhanced long-term memory for embedded objects. *Neurobiology of Learning and Memory*, 134(Pt. A), 65–77.
- Madan, C. R., Fujiwara, E., Gerson, B. C., & Caplan, J. B. (2012). High reward makes items easier to remember, but harder to bind to a new temporal context. *Frontiers in Integrative Neuroscience*, *6*, 61.
- Mather, M., & Schoeke, A. (2011). Positive outcomes enhance incidental learning for both younger and older adults. *Frontiers in Neuroscience*, 5, 129.
- Moeller, S., Yacoub, E., Olman, C. A., Auerbach, E., Strupp, J., Harel, N., & Uğurbil K. (2010). Multiband multislice GE-EPI at 7 tesla, with 16-fold acceleration using partial parallel imaging with application to high spatial and temporal whole-brain fMRI. *Magnetic Resonance in Medicine*, 63(5), 1144–1153.
- Murty, V. P., & Adcock, R. A. (2014). Enriched encoding: Reward motivation organizes cortical networks for hippocampal detection of unexpected events. *Cerebral Cortex*, 24(8), 2160–2168.
- Murty, V. P., LaBar, K. S., & Adcock, R. A. (2016). Distinct medial temporal networks encode surprise during motivation by reward versus punishment. *Neurobiology of Learning and Memory*, 134(Pt. A), 55–64.
- Palombo, D. J., Di Lascio, J. M., Howard, M. W., & Verfaellie, M. (2019, ). Medial temporal lobe amnesia is associated with a deficit in recovering temporal context. *Journal of Cognitive Neuroscience*, *31*(2), 236-248.
- Palombo, D. J., Keane, M. M., & Verfaellie, M. (2015). How does the hippocampus shape decisions? *Neurobiology of Learning and Memory*, 125, 93–97.
- Palombo, D. J., & Verfaellie, M. (2017). Hippocampal contributions to memory for time: Evidence from neuropsychological studies. *Current Opinion in Behavioral Sciences*, 17, 107–113.
- Poldrack, R. A., Clark, J., Pare-Blagoev, E. J., Shohamy, D., Creso Moyano, J., Myers, C., & Cluck, M. A. (2001). Interactive memory systems in the human brain. *Nature*, 414(6863), 546–550.
- Poldrack, R. A., & Packard, M. G. (2003). Competition among multiple memory systems: Converging evidence from animal and human brain studies. *Neuropsychologia*, 41(3), 245–251.
- Poppenk, J., Evensmoen, H. R., Moscovitch, M., & Nadel, L. (2013). Long-axis specialization of the human hippocampus. *Trends in Cognitive Sciences*, 17(5), 230–240.

- Preuschoff, K., Bossaerts, P., & Quartz, S. R. (2006). Neural differentiation of expected reward and risk in human subcortical structures. *Neuron*, 51(3), 381–390.
- Pruim, R. H., Mennes, M., van Rooij, D., Llera, A., Buitelaar, J. K., & Beckmann, C. F. (2015). ICA-AROMA: A robust ICA-based strategy for removing motion artifacts from fMRI data. *NeuroImage*, *112*, 267–277.
- Roiser, J. P., Linden, D. E., Gorno-Tempinin, M. L., Moran, R. J., Dickerson, B. C., & Grafton, S. T. (2016). Minimum statistical standards for submissions to Neuroimage: Clinical. *NeuroImage: Clinical*, 12, 1045–1047.
- Schonberg, T., O'Doherty, J. P., Joel, D., Inzelberg, R., Segev, Y., & Daw, N. D. (2010). Selective impairment of prediction error signaling in human dorsolateral but not ventral striatum in Parkinson's disease patients: Evidence from a model-based fMRI study. *NeuroImage*, 49(1), 772–781.
- Sheldon, S., & Levine, B. (2016). The role of the hippocampus in memory and mental construction. Annals of the New York Academy of Sciences, 1369(1), 76–92.
- Sheldon, S., Romero, K., & Moscovitch, M. (2013). Medial temporal lobe amnesia impairs performance on a free association task. *Hippocampus*, 23(5), 405–412.
- Shohamy, D., & Adcock, R. A. (2010). Dopamine and adaptive memory. *Trends in Cognitive Sciences*, 14(10), 464–472.
- Shohamy, D., Myers, C. E., Hopkins, R. O., Sage, J., & Gluck, M. A. (2009). Distinct hippocampal and basal ganglia contributions to probabilistic learning and reversal. *Journal of Cognitive Neuroscience*, 21(9), 1821–1833.
- Shohamy, D., Myers, C. E., Kalanithi, J., & Gluck, M. A. (2008). Basal ganglia and dopamine contributions to probabilistic category learning. *Neuroscience and Biobehavioral Reviews*, 32(2), 219–236.
- Shohamy, D., & Turk-Browne, N. B. (2013). Mechanisms for widespread hippocampal involvement in cognition. *Journal of Experimental Psychology: General*, 142(4), 1159–1170.
- Spaniol, J., Schain, C., & Bowen, H. J. (2014). Reward-enhanced memory in younger and older adults. *The Journals of Gerontology. Series B, Psychological Sciences and Social Sciences*, 69(5), 730–740.
- Squire, L. R. (2004). Memory systems of the brain: a brief history and current perspective. *Neurobiology of Learning and Memory*, 82(3), 171–177.
- Stachenfeld, K. L., Botvinick, M. M., & Gershman, S. J. (2017). The hippocampus as a predictive map. *Nature Neuroscience*, 20(11), 1643–1653.
- Strange, B. A., Witter, M. P., Lein, E. S., & Moser, E. I. (2014). Functional organization of the hippocampal longitudinal axis. *Nature Reviews Neuroscience*, 15(10), 655–669.
- Wechsler, D. (1997). Wechsler Adult Intelligence Scale–Third Edition: Administration and scoring manual. San Antonio, TX: Harcourt Assessment.
- Wimmer, G. E., Daw, N. D., & Shohamy, D. (2012). Generalization of value in reinforcement learning by humans. *Eur J Neurosci*, 35(7), 1092-1104.
- Wittmann, B. C., Bunzeck, N., Dolan, R. J., & Duzel, E. (2007). Anticipation of novelty recruits reward system and hippocampus while promoting recollection. *NeuroImage*, 38(1), 194–202.
- Woo, C. W., Krishnan, A., & Wager, T. D. (2014). Cluster-extent based thresholding in fMRI analyses: Pitfalls and recommendations. *NeuroImage*, 91, 412–419.
- Woolrich, M. W., Behrens, T. E., Beckmann, C. F., Jenkinson, M., & Smith, S. M. (2004). Multilevel linear modelling for fMRI group analysis using Bayesian inference. *Neuroimage*, 21(4), 1732-1747.
- Woolrich, M. W., Ripley, B. D., Brady, M., & Smith, S. M. (2001). Temporal autocorrelation in univariate linear modeling of fMRI data. *NeuroImage*, 14(6), 1370–1386.