

# Modulating the Use of Multiple Memory Systems in Value-based Decisions with Contextual Novelty

Katherine Duncan<sup>1</sup>, Annika Semmler<sup>2</sup>, and Daphna Shohamy<sup>3</sup>

## Abstract

■ With multiple learning and memory systems at its disposal, the human brain can represent the past in many ways, from extracting regularities across similar experiences (incremental learning) to storing rich, idiosyncratic details of individual events (episodic memory). The unique information carried by these neurologically distinct forms of memory can bias our behavior in different directions, raising crucial questions about how these memory systems interact to guide choice and the factors that cause one to dominate. Here, we devised a new approach to estimate how decisions are independently influenced by episodic memories and incremental learning. Furthermore, we identified a biologically motivated factor that

biases the use of different memory types—the detection of novelty versus familiarity. Consistent with computational models of cholinergic memory modulation, we find that choices are more influenced by episodic memories following the recognition of an unrelated familiar image but more influenced by incrementally learned values after the detection of a novel image. Together this work provides a new behavioral tool enabling the disambiguation of key memory behaviors thought to be supported by distinct neural systems while also identifying a theoretically important and broadly applicable manipulation to bias the arbitration between these two sources of memories. ■

## INTRODUCTION

Decades of research have demonstrated that experiences are encoded within neurally distinct learning and memory systems (Poldrack et al., 2001; Squire & Zola, 1996; Knowlton, Squire, & Gluck, 1994; McDonald & White, 1993), each of which can bias future choices. For example, after reading a compelling article, you could later recall your experience reading a particular finding—an episodic memory (Tulving, 1983)—as well as increase your valuation of the journal itself and, consequently, the likelihood of your searching for new articles within it—incremental reward learning (Sutton & Barto, 1998). This multiple memory system framework has proven powerful, providing a structure for understanding learning and memory disorders (Small, Schobel, Buxton, Witter, & Barnes, 2011; Yassa, Mattfeld, Stark, & Stark, 2011; Frank, Seeberger, & O'reilly, 2004; Knowlton et al., 1994) and driving recent investigation into the distinct learning mechanisms that underlie each separate system. Understanding the interactions between these forms of learning has important implications for how memory guides later behavior. However, most research so far has focused on

studying each system in isolation, leaving open important questions about how they interact, how they guide decisions, and which factors bias their use.

Incremental learning of value is often studied by characterizing how values are updated through experience (Daw, 2011)—a form of reinforcement learning thought to depend on dopamine release in the striatum (Schultz, 1998, 2016; Schonberg et al., 2010; Pessiglione, Seymour, Flandin, Dolan, & Frith, 2006; Barto, 1995). Experiments in this field tend to present the same choice options across hundreds of trials. Each choice is reinforced with a reward or loss. The difference between the obtained outcome and the choice's expected value (reward prediction error) is thought to be signaled by dopamine release (Schultz, 1998; Sutton & Barto, 1998), so that unexpected outcomes incrementally nudge the values attributed to each option toward the experienced outcome. In this way, incremental learning extracts the decontextualized value of each choice option by averaging across experiences and discarding the individual episodes.

A separate line of research has focused on episodic memory and its dependence on hippocampal processes (Tulving & Markowitsch, 1998; Cohen & Eichenbaum, 1993; Squire, 1992). This research tends to assess the influence of past experiences on behavior by asking participants to directly reflect on them. Experiments often begin with an encoding phase in which participants are presented with a unique stimulus on each trial. This is followed by a retrieval phase in which new stimuli are

---

This paper is part of a Special Focus deriving from a symposium at the 2017 Annual Meeting of the Cognitive Neuroscience Society entitled “Memory Neuromodulation: Influences of Learning States on Episodic Memory.”

<sup>1</sup>University of Toronto, <sup>2</sup>Vrije Universiteit Amsterdam, <sup>3</sup>Columbia University

intermixed with those from the encoding session (old stimuli). Because old stimuli were only presented once before, they unambiguously cue memory for a single episode, which participants are asked to recall. Damage to the hippocampus robustly impairs both the encoding of these episodic memories and their retrieval (Gilboa et al., 2006; Clark, Broadbent, Zola, & Squire, 2002; Nadel & Moscovitch, 1997; Scoville & Milner, 1957). Analogously, hippocampal fMRI activity has been related to the successful encoding (Ranganath et al., 2004; Davachi, Mitchell, & Wagner, 2003) and retrieval (Eldridge, Knowlton, Furmanski, Bookheimer, & Engel, 2000) of episodic memory in healthy populations, providing converging evidence for the hippocampus' involvement in episodic memory.

Unlike these experimental paradigms, in our everyday experiences, we are constantly tasked with choosing between options that could trigger the retrieval and use of distinct episodic memories as well as incrementally learned values. Which system will dominate in these less prescriptive contexts? Not only is this quandary frequently encountered, it also has important implications. For example, breaking persistent maladaptive behaviors, such as addiction, often requires overriding ingrained habitual responses, presumed to depend on striatal incremental learning (Volkow, Wang, Fowler, Tomasi, & Telang, 2011), with episodic memories of newly learned strategies, presumed to depend on the hippocampus (Tulving & Markowitsch, 1998). Thus, the challenge facing adaptive memory is often not just selecting the right memories but instead biasing behavior toward the most appropriate memory system.

At a neural level, one factor that may arbitrate between these systems is cholinergic regulation of hippocampal-dependent memory. Neuromodulation plays a central role in theories of incremental learning, with dopamine regulating when new learning should occur (Barto, 1995). Cholinergic neuromodulation has also been hypothesized to play a somewhat analogous role in hippocampal memory; it has been argued to bias when the hippocampus engages in the encoding versus the retrieval of episodic memories (Easton, Douchamps, Eacott, & Lever, 2012; Meeter, Murre, & Talamini, 2004; Hasselmo, Wyble, & Wallenstein, 1996). Specifically, neurocomputational models (Meeter et al., 2004; Hasselmo et al., 1996) and empirical findings (Vandecasteele et al., 2014; Douchamps, Jeewajee, Blundell, Burgess, & Lever, 2013) suggest that acetylcholine levels maintain prolonged hippocampal "states," which last for a few seconds at a time (Pabst et al., 2016; Hasselmo & Fehrlau, 2001). The value of these states lies in their ability to accommodate the competing computational demands of memory formation and retrieval (O'Reilly & McClelland, 1994), optimally timing each memory process based on contextual factors. Specifically, the recognition of familiar contexts decreases cholinergic transmission to favor further episodic memory retrieval in the presence of potential memory cues.

Conversely, the detection of novel contexts increases cholinergic transmission (Giovannini et al., 2001) to favor episodic memory encoding in the presence of new and unexpected information. We have previously explored the consequences of this hypothesis for human behavior, demonstrating that recent exposure to familiar images increases the ability to later recall unrelated associations (Patil & Duncan, 2018) and increases the likelihood that people use episodic memory cues to make decisions (Duncan & Shohamy, 2016). By contrast, recent exposure to novel images increases the subsequent formation of episodic memories, which can later be used to guide future actions (Duncan & Shohamy, 2016).

Our earlier research, however, only assessed the impact of familiarity- and novelty-evoked states on episodic memories; memory cues were unambiguously associated with single past events. Thus, the important question remains—can novelty bias the type of memory used by individuals when they have multiple memory sources at their disposal? The cholinergic framework would predict that episodic memories are more likely to dominate over competing incrementally learned values to guide choice in familiar context, because recognizing one familiar cue primes episodic memory to take advantage of other familiar episodic memory cues.

To test this prediction, we modified our episodic decision-making paradigm (Duncan & Shohamy, 2016) so that people are now free to make choices using both individual episodic memories and incrementally learned values. The few established paradigms that assess both hippocampal- and striatal-mediated learning do so by pitting the systems against each other, reinforcing either the hippocampal- or striatal-dependent behavior at different points (Packard, 2009). This approach, however, offers limited insights into how people arbitrate between two sources of memories when both are viable sources of information. By contrast, here we combine elements of incremental reward learning and episodic memory tasks into a single paradigm in which both types of learning are simultaneously reinforced. Specifically, participants chose one of two cards dealt in a computerized card game. Each card had two dimensions: (1) a distinctive object, which was repeated at most once during the experiment, and (2) a deck color, which was present on every trial and probabilistically related to reward distributions. Thus, the participant could either use the distinctive objects to cue episodic memories of a specific card's values or they could use the colored decks by incrementally updating and refining their values across trials. Importantly, we controlled trial sequences and outcomes to decorrelate the object and deck values, allowing us to independently estimate the influence of each on participants' choices.

We used this paradigm to assess whether recent novelty and familiarity can arbitrate between the use of episodic memories and incrementally learned values. First, we demonstrated that people are more likely to

select cards using episodic object memories following the successful retrieval and use of other object memories. Conversely, the use of object memories negatively impacted the subsequent use of incrementally learned deck values. These biases were robustly observed in both reward (Experiment 1A) and loss experiments (Experiment 1B), underscoring the power of the manipulation. Second, we further challenged the underlying mechanisms driving the bias using a contextual novelty manipulation; we displayed an unrelated novel or familiar scene image before dealing cards on each trial (Experiment 2). Despite the incidental relationship between the scene images and the task, familiar scenes increased the use of episodic memories of object values at the cost of incrementally learned deck values. This combination of results supports a neurally inspired model of memory in which accessing the contents of one system increases the subsequent accessibility of other unrelated memories stored within the same system.

## EXPERIMENT 1

In Experiments 1A and 1B we designed a new task to disambiguate the contributions of episodic and incremental learning to behavior. On each of a series of trials, participants chose between two cards for the chance to win a monetary reward. Each card had two features: a distinctive object and the color of the deck from which the card was drawn. We designed these features to tap distinct forms of memory. As in our previous work, each object repeated at most once in the task so that using memory for the objects required the rapid acquisition of an object value association (Duncan & Shohamy, 2016). Here, we added a second dimension to this choice problem—two decks (red and blue) that were available on every trial and were probabilistically associated with reward outcomes. This noisy and nonstationary relationship between decks and outcomes could be efficiently extracted by incrementally updating the values of decks across trials, forming a recency-weighted average of value across experience. We first confirmed that participants could employ both types of learning without experiencing trade-offs or interactions between them in both reward (Experiment 1A) and loss (Experiment 1B) experiments. We then tested our prediction that, when people have multiple types of memory at their disposal, recognizing familiar cards would increase the use of episodic memory on the next trial, potentially at the cost of incremental value learning.

### Methods

#### Participants

Sixty-one members of the Columbia University community (30 women, mean age = 20.2 years) participated in the study for pay (\$12/hr + bonus earnings). Participants

were divided between Experiment 1A (31 participants) and Experiment 1B (30 participants). All participants provided informed consent at the beginning of the study, and all procedures were approved by the Columbia Morningside ethics board.

#### Stimuli

Five hundred forty-three color images of different commonplace objects were used as stimuli. Additionally, two decks of virtual playing cards (red and blue decks) were generated. At any point in the experiment, one deck drew outcomes from a lucky distribution (average outcome = 63¢), whereas the other drew from an unlucky distribution (average outcome = 37¢; Table 1). The left versus right position of the card was randomly assigned for every trial.

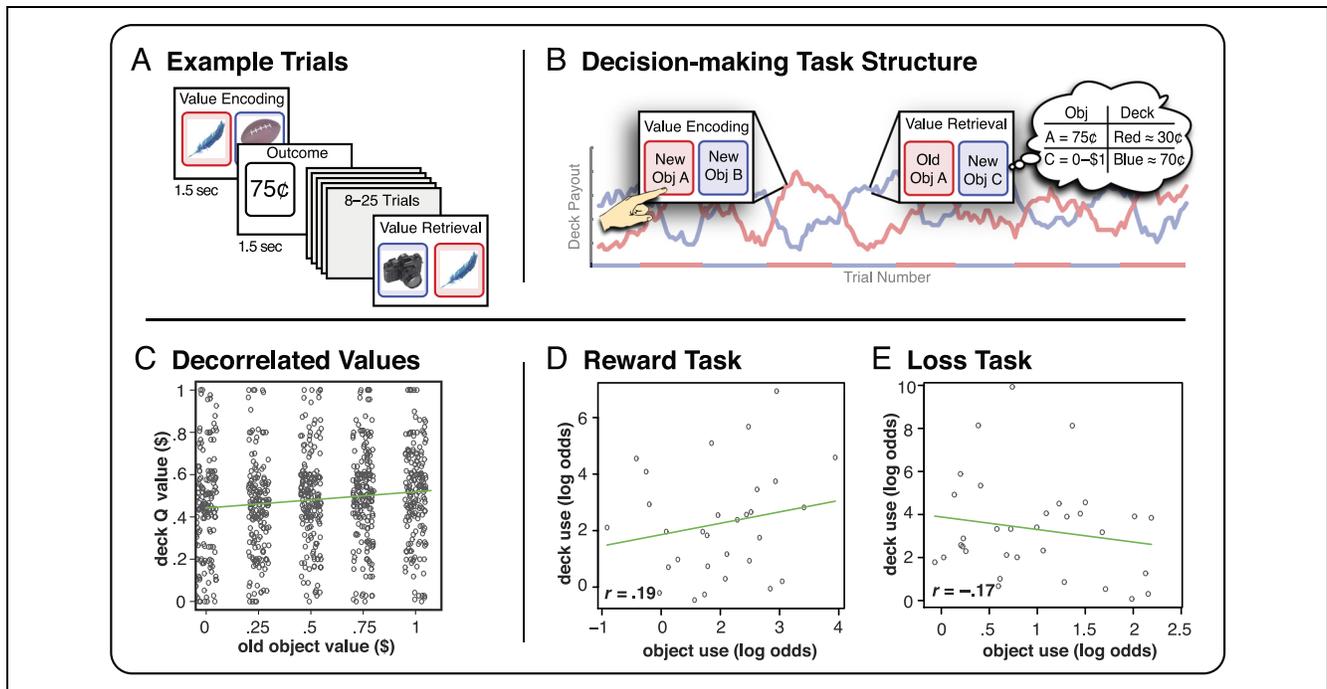
#### Procedure

Participants performed a series of trials in which they chose between cards for the chance to earn a monetary reward (Figure 1A). On each trial, participants were presented with two cards and were given 1.5 sec to choose one using the “j” and “k” keys on a standard keyboard. They were then shown the outcome of their choice (i.e., the selected card’s value) for 1.5 sec, followed by a 1.5-sec fixation cross between trials. On every trial, one of the cards was drawn from a red deck and one from a blue deck. Each card additionally had an object on it; the identity of the objects varied across trials. Outcomes ranged between 0 cents and 1 dollar. As detailed below, a key feature of the design is that both the color of the deck (red or blue) and the specific object were related to the outcome.

Following many prior studies of reinforcement learning, the relationship between the color of the deck (red or blue) and the likelihood of reward was probabilistic and varied over the course of the experiment. In particular, the likelihood of each colored deck being the better choice (the “luckier” deck) reversed frequently (every  $20 \pm 5$  trials), such that participants had to continuously use the outcomes to update deck values to inform their future choices. Learning such probabilistic associations between decks and outcomes unfolds over time because no single outcome can be used to infer which deck is “luckier” at a given phase of the experiment. Learning this sort of task is well fit by reinforcement learning

**Table 1.** Distribution of Old Object Card Values by Deck Luckiness

	0¢	20¢	40¢	60¢	80¢	\$1	Mean
Lucky deck	16%	18%	16%	18%	17%	15%	49¢
Unlucky deck	17%	16%	17%	16%	16%	16%	49¢



**Figure 1.** Task schematic and structure for independent assessment of incremental versus episodic memory. (A) Example trials illustrating cards' distinctive object and repeating deck features. We use "Value Encoding" to designate trials on which both cards have new objects (new–new). Conversely, "Value Retrieval" refers to trials when one card has a previously selected object (new–old), permitting episodic retrieval of a specific card's value. Incremental updating of red and blue deck values can occur on every trial, as can the use of deck values. (B) Schematic illustrating the logic employed in the decorrelation of object and deck values. The graph plots the running average of recent deck outcomes experienced by a participant. The two trials illustrate how periodically reversing deck values can result in incongruent object and deck values. Specifically, if a participant chose a new object card from the red deck during a phase when the red deck was lucky and received a good outcome, then the next time the same card appeared it might be during a phase when the blue deck is luckier. (C) Incrementally learned deck Q values (estimated by a reinforcement learning model) are plotted against episodically learned object values to visualize their independence. (D–E). The use of each value type is not significantly correlated across participants in either the reward or loss experiments.

models that assume that participants learn the average value of each deck color across trials.

Here, in contrast to previous studies, we included a second feature, which also predicted outcomes—the distinctive object. On roughly half of the trials, both objects had never been seen before ("new"). On the other trials, an object card that was selected on one earlier trial was re-presented ("old") alongside a new card. These old objects were worth the same amount as they were the first time they appeared. Therefore, participants could use their memory for an old object's value to predict the outcome of selecting it.

Critically, the values of the decks versus the objects were decorrelated to permit the independent estimation of how much participants used each to make their choices. This decorrelation was enabled by the reversals in deck luckiness; objects from a particular deck could be repeated at points when that deck had the same or reversed luckiness (Figure 1B). As can be seen in Table 1, this approach succeeded in making object and deck values independent—all old object card values were similarly likely to occur when selected from a deck that was currently lucky or unlucky. Additionally, by carefully controlling card presentations, we also ensured that old

object cards and new object cards were similarly valuable on average. We also used a reinforcement learning model (see below) to derive personalized estimates of incrementally learned deck values on a trial-by-trial basis. As illustrated in Figure 1C, these model-derived deck values were not significantly correlated with object values (mean  $r = .09$ ,  $SD = .09$ ). Altogether, this design allowed us to use participants' choices to infer how much each choice was driven by incremental value versus episodic memory on a trial-by-trial basis.

In Experiment 1A, participants made choices in a reward context and received 7% of their total winnings at the end of the experiment. In Experiment 1B, participants made choices in a loss context; they received a \$20 endowment at the beginning of the experiment and selected cards to minimize their losses (7% of total outcomes). In both experiments, participants performed 348 trials divided equally into three blocks.

#### Reinforcement Learning Model

We modified a temporal difference reinforcement learning model (Sutton & Barto, 1998) to derive tailored estimates of each participant's learning of both deck and

object values based on their personal history of choices and experienced outcomes. This model assumes that the value,  $Q$ , of the chosen deck,  $d_c$ , is updated on every trial,  $t$ , based on the difference between the expected value of the chosen deck and the obtained reward,  $r$ . This difference is termed the prediction error,  $\delta$

$$\delta_t = r_t - Q_{t-1}(d_c)$$

The degree to which this prediction error updates the value of the chosen deck depends on the learning rate parameter,  $\alpha$

$$Q_t(d_c) = Q_{t-1}(d_c) + \alpha\delta_t$$

As  $\alpha$  approaches 1, the value of the chosen deck will approach the most recently received outcome. As  $\alpha$  approaches 0, the value of the chosen deck will be minimally updated. Intermediate values of  $\alpha$  will result in deck values that integrate across outcomes, reflecting an incremental learning process. We fit  $\alpha$  to participants' behavior, restricting its range to be between 0 and 1.

The value of the unchosen deck,  $d_u$ , was not updated:

$$Q_t(d_u) = Q_{t-1}(d_u)$$

The deck values and the participant's history with the object on the card were used to compute the probability,  $P$ , of selecting the card from the red deck using a softmax (logistic) choice rule:

$$P(d_{\text{red}}) = \frac{1}{1 + \exp(-(\beta_d(Q_{\text{red}} - Q_{\text{blue}}) + \beta_o V_o + \beta_r O))}$$

Here, the inverse temperature parameter,  $\beta_d$ , controls how closely the difference in deck values tracks choices.  $\beta_o$  controls the influence of the old object card's value,  $V_o$ .  $V_o$  is positive when the old card is in the red deck, negative when the old card is in the blue deck, and 0 when there is no old card.  $\beta_r$  controls how much participants prefer cards with old objects,  $O$  ( $0 =$  red deck card is new;  $1 =$  red deck card is old). Each  $\beta$  parameter was fit to participants' choice with a lower bound of 0.

We estimated the four free parameters ( $\alpha$ ,  $\beta_d$ ,  $\beta_o$ , and  $\beta_r$ ) for each participant by minimizing the sum of the negative log likelihoods of choices given the estimated probability,  $P$ , of each choice using constrained nonlinear optimization (*fmincon*, MATLAB). We repeated the search with all parameters set at five random starting points to avoid local optima and selected the iteration, which resulted in the lowest negative log likelihoods. To generate  $Q$ -value regressors, we used the average  $\alpha$  across participants to generate trial-by-trial  $Q$ -value estimates based on each participant's personal trial sequence. This approach reduces overfitting and the noisiness in individual participants' estimates (Schönberg, Daw, Joel, & O'Doherty, 2007; Daw, O'Doherty, Dayan, Seymour, & Dolan, 2006).

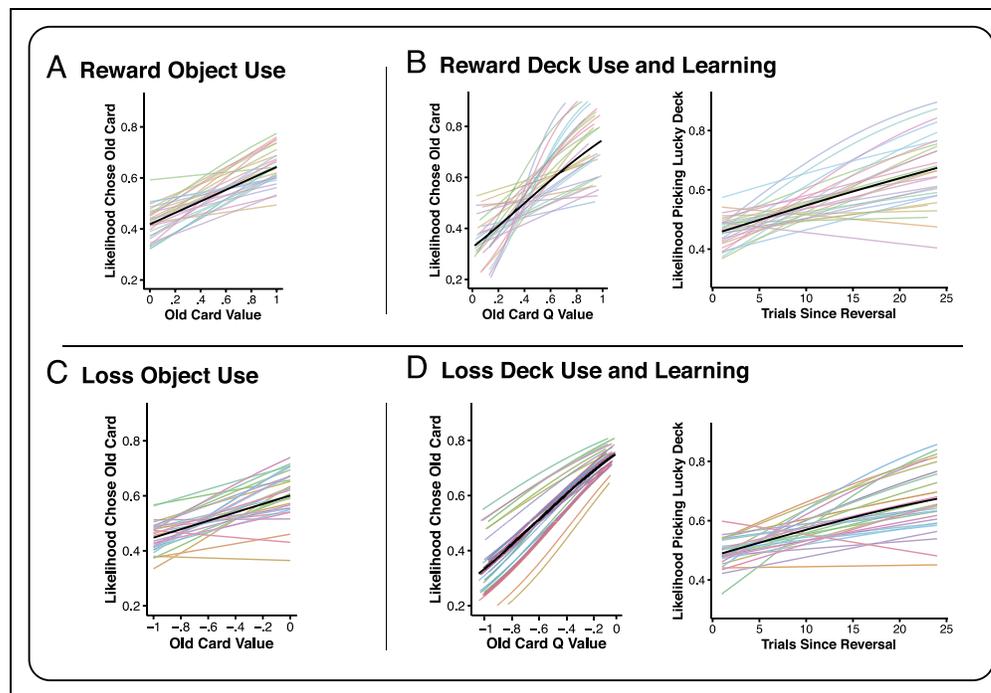
### Mixed Generalized Linear Models

We also used mixed generalized linear models with a logistic linking function to predict choices based on the object values, deck values, and their interaction with recent mnemonic processes. Specifically, we identified object recognition trials as the trials on which choices were optimally consistent with previously experienced object values (choosing old cards worth  $> \$0.50$  and avoiding old cards worth  $< \$0.50$ ). In separate models, we also classified object recognition using experiment-tailored decision thresholds that took biases toward selecting old over new cards into account. In these models, trials were labeled as recognized if participants selected old cards worth  $> \$0.34$  (Experiment 1A) and  $> \$0.32$  (Experiment 1B) or avoiding old cards worth less than these thresholds. It should be noted that this approach is quite similar to the optimal threshold as only cards worth  $\$0.40$  were affected. Recognized trials were coded with 1, whereas trials which did not contain an old card or on which choices did not reflect object memory were coded with 0. We then assessed the degree to which expressing object memory on the preceding trial modulated the use of object memory on the subsequent trial by predicting choice (old or new card) with the value of the old card ( $\$0$ – $\$1$ ), the preceding object memory, and the interaction between these variables. Similarly, we assessed the impact of preceding object memory on the use of deck values by predicting the likelihood of selecting the same deck as on the preceding trial (stay response) based on the outcome of the preceding choice ( $\$0$ – $\$1$ ), the preceding object use, and the interaction between these factors. All models were estimated using the *lme4* package (Bates, Maechler, Bolker, & Walker, 2015) in the R programming language and were estimated by optimizing restricted maximum likelihood. Each fixed effect term was also included as a random slope, grouped by participant.

### Results and Discussion

We first assessed whether participants could independently use both episodic object memories and incrementally learned deck values to make choices. We used a reinforcement learning model (Sutton & Barto, 1998) to derive personalized trial-by-trial estimates of participants' deck values ( $Q$  values) based on their unique sequence of choices and outcomes. Notably, model-derived deck  $Q$  values reflect the outcome of incremental value updating. By contrast, remembered object values could be simply coded as the previously experienced outcome, as participants only had one relevant past experience. We then quantified each participants' use of both types of learned values by estimating the increased log odds of selecting an old card based on its object value and the difference deck  $Q$  values. As shown in Figure 2, both types of value memories were strong predictors of choice in both the reward and loss experiments (Reward: object:

**Figure 2.** Use of episodically learned object values and incrementally learned deck values. (A) Use of object values in the reward experiment (Experiment 1A). Participants are consistently more likely to pick an old object card that resulted in high outcomes in the past and avoid those that resulted in poor outcomes. (B) Deck Q values (estimated using a reinforcement learning model) also influence choices about old object cards (left graph). Participants are more likely to select the old card when it is dealt from a deck with a higher Q value and avoid it when dealt from a lower Q-value deck. Participants also slowly increase their selection of the currently lucky deck with repeated experience (right graph). A similar pattern of object (C) and deck (D) use was observed in the loss experiment (Experiment 1B). Black lines indicate the group fixed effect and colored lines plot individual participants.



$\beta = 0.77, z = 4.61, p < .0001$ ; deck:  $\beta = 1.86, z = 4.69, p < .0001$ , Figure 2A and B; Loss: object:  $\beta = 0.48, z = 3.16, p = .001$ ; deck:  $\beta = 1.67, z = 7.06, p < .0001$ , Figure 2C and D). Participants also had a slight preference for old over new cards in the reward context ( $\beta = 0.15, z = 2.49, p = .01$ ) and a trend toward this same bias in the loss context ( $\beta = 0.12, z = 1.86, p = .06$ ). Together, this indicates that participants were more likely to select decks that had recently resulted in higher outcomes while avoiding decks that had recently resulted in lower outcomes. Additionally, participants were more likely to select old object cards that yielded higher outcomes when previously selected while avoiding old object cards that yielded lower outcomes.

Importantly, the use of these values was independent, both across and within participants. Across participants, we did not find a significant correlation between the use of deck and object value (reward,  $r = .19, p = .31$ ; loss  $r = -.17, p = .33$ ; Figure 1D and E). This suggests that participants did not strategically rely on one memory type throughout the experiment, but instead that they had to arbitrate between two competing memory types on a trial-by-trial basis. Within participants, deck and object value memories did not significantly interact when guiding choices (reward:  $\beta = 0.34, z = 0.92, p = .36$ ; loss:  $\beta = -0.22, z = -0.65, p = .51$ ), suggesting that participants could independently access and use both types of values. Moreover, there was no systematic shift in the type of memory used across the two experimental runs (reward object:  $\beta = 0.05, z = 0.23, p = .82$ ; reward

1deck:  $\beta = 0.36, z = 1.22, p = .22$ ; loss object:  $\beta = 0.13, z = 0.58, p = .56$ ; loss deck:  $\beta = 0.19, z = 0.52, p = .60$ ) or in the 10 trials following a reversal in deck “luckiness” (reward object:  $\beta = 0.10, z = 0.34, p = .73$ ; reward deck:  $\beta = 0.14, z = 0.56, p = .57$ ; loss object:  $\beta = -0.10, z = -0.43, p = .67$ ; loss deck:  $\beta = 0.19, z = 0.88, p = .38$ ).

We additionally used the reinforcement learning model to confirm that participants used incremental learning to acquire deck values. We specifically assessed the distribution of learning rates estimated across participants. A learning rate of 0 reflects no updating of deck Q values based on outcomes, whereas a learning rate of 1 reflects complete updating of deck Q values, which effectively erases all memory of prior outcomes; learning rates between 0 and 1 reflect incremental updating of values, aggregating outcomes across trials. The lower the learning rate, the greater the impact of more distant experiences. We found a mean learning rate of 0.54 and an interquartile range of 0.29–0.89 in Experiment 1A (reward) and a mean learning rate of 0.64 and an interquartile range of 0.36–.97 in Experiment 1B (loss), with no significant difference between the reward and loss experiments,  $t(59) = -1.01, p = .32$ . These moderate model-derived learning rates suggest that participants incrementally aggregated information across trials to guide their choices about the decks.

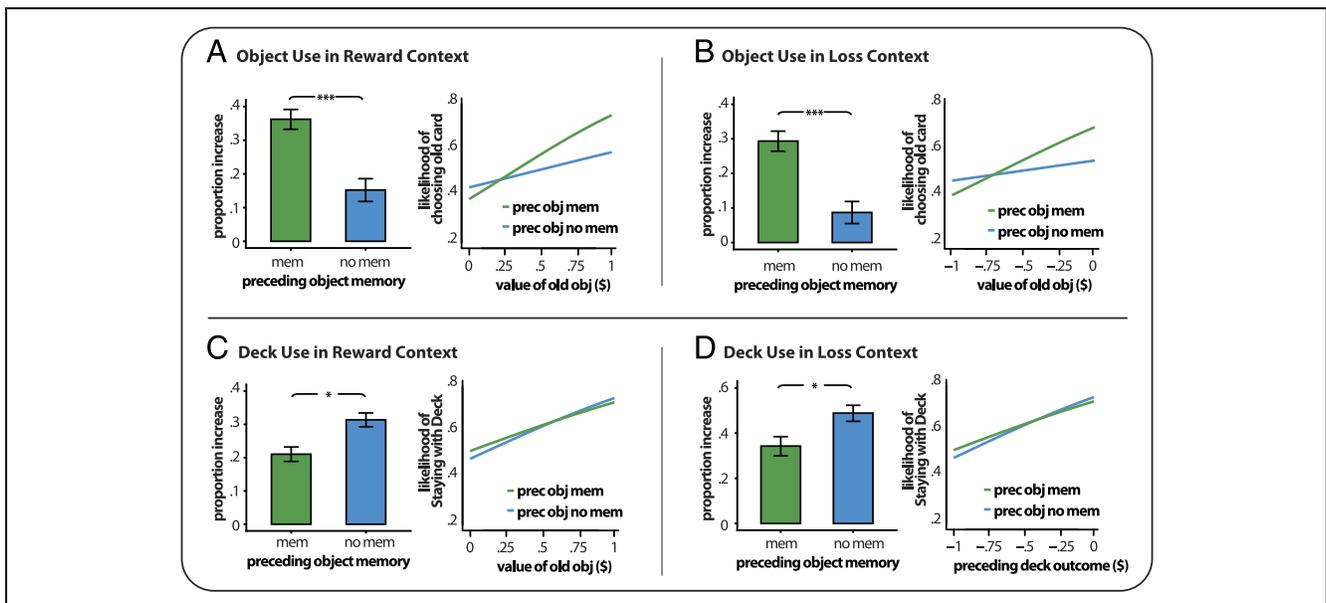
Given that both episodic object memories and incrementally learned deck memories were found to independently drive choices, we turned to our primary question:

Can the recent detection of novelty versus familiarity bias the subsequent use of different memory systems in decision-making? We first asked whether participants' use of episodic memories increased after their successful retrieval and use of other episodic memories on the preceding trial, as would be predicted by our memory state framework. In both reward and loss experiments, participants were nearly three times more likely to use an object memory if they had used one on the preceding trial (reward:  $\beta = 0.94, z = 3.40, p = .0007$ ; loss:  $\beta = 0.89, z = 3.40, p = .0006$ ; Figure 3A and B). Similar results were obtained when using experiment-tailored decision thresholds to infer recognition on the prior trial (reward:  $\beta = 0.97, z = 3.24, p = .001$ ; loss:  $\beta = 0.71, z = 2.79, p = .005$ ). Preceding episodic memory use also increased participants' preference for old over new cards, regardless of their values (reward:  $\beta = 0.26, z = 2.73, p = .006$ ; loss:  $\beta = 0.18, z = 2.39, p = .02$ , Figure 3A and B); in fact, this bias toward familiar options was not observed when the preceding trial was new or not recognized (reward:  $\beta = 0.06, z = 0.85, p = .39$ ; loss:  $\beta = 0.06, z = 0.83, p = .41$ ). Our unique task allowed us also to assess how this bias toward episodic memory use influenced the use of incrementally learned values. We found that participants were less likely to use prior deck outcomes to guide choices (i.e., staying with a deck that resulted in high outcomes) following the use of an episodic memory as compared with following trials on which episodic memories were not used (reward:  $\beta = -0.46, z = -2.37, p = .02$ ; loss:  $\beta = -0.86, z = 4.32, p < .0001$ ;

Figure 3C and D). Thus, using episodic object memories orients participants toward using other object memories on the subsequent trial, but at the cost of using deck memories.

## EXPERIMENT 2

Experiments 1A and 1B developed a new approach to disambiguating the contributions of episodic and incrementally learned memories to behavior. They also identified a crucial factor that biases the arbitration between these two types of memory—the recent retrieval and use of episodic memories. This bias could be driven by two possible mechanisms: The first is the generation of an episodic retrieval state, akin to those identified by our prior research, which studied episodic memory in isolation (Patil & Duncan, 2018; Duncan & Shohamy, 2016; Duncan, Sadanand, & Davachi, 2012) and which is predicted by neurocomputational models of hippocampal function (Meeter et al., 2004; Hasselmo et al., 1996). Alternatively, this bias may be driven by fluctuations in how participants attend to the cues; that is, orienting to the object on one trial may be related to a tendency to orient toward objects on the subsequent trial. These accounts make different predictions about the breadth of choices that should be influenced by recent episodic retrieval. The attentional mechanism would only apply when episodic memories are consistently cued in similar ways across experiences—as was the case when episodic memories were cued only by objects. By contrast, the



**Figure 3.** Choices depend on the use of episodic memories in the immediately preceding trial. The value of old objects influenced choice more strongly on trials that were preceded by the retrieval of another, unrelated object card memory. This bias was observed in reward (A) and loss (B) experiments. The left graph plots how well the old card's value predicted participants' choices (old vs. new card) for decisions that were or were not made following episodic memory retrieval. Statistical comparisons were performed by testing the interaction between old card value and preceding object memory. The right graph plots the model estimates of the likelihood of choosing old cards of different values on trials following or not following object memory retrieval. Use of incrementally learned deck Q values was reduced on trials that were preceded by object card memory retrieval. This bias was also observed in reward (C) and loss (D) experiments. \* $p < .05$ , \*\*\* $p < .001$ .

state-based mechanism implies that memory systems can be primed by incidental contextual manipulations—a phenomenon that could result in broad applications. To tease apart these two explanations, we introduced a novelty/familiarity manipulation, which was incidental to the card game itself and which has been shown previously to modulate the tendency to encode versus retrieve episodic memories (Duncan & Shohamy, 2016). Specifically, cards were dealt on “decorative mats,” which were either novel or familiar images of scenes. These mats did not predict the content of the to-be-dealt cards or their outcomes and, thus, could only influence the type of memory used by evoking a general bias toward episodically retrieved information. Their influence on value-based decision-making would, thus, strongly support the hypothesis that episodic memory states bias the arbitration between episodically and incrementally learned values.

## Methods

### *Participants*

Thirty-eight members of the Columbia University community (17 women, mean age = 23.2 years) who had not participated in Experiment 1 participated in the study for pay (\$12/hr + bonus earnings). All participants provided informed consent at the beginning of the study, and all procedures were approved by the Columbia Morningside ethics board.

### *Stimuli*

One hundred ninety-five scenes and 460 objects were used as stimuli. Five scenes were randomly assigned to the “familiar” condition, and participants were preexposed to them before starting the experiment. Additionally, two decks of virtual playing cards (red and blue decks) were generated. These decks drew from the same lucky and unlucky distributions as were used in Experiments 1A and 1B (Table 1).

### *Procedure*

Participants performed the same card task as in Experiment 1A. The only modifications were the insertion of novel and familiar scenes and a reduction in the total number of trials. Specifically, each choice was preceded by the 1-sec presentation of a novel or familiar scene (referred to for the participants as a “decorative mat”). Participants were told that the purpose of the mat was to prepare them for the upcoming cards. The scene remained on the screen for the subsequent 1.5 sec decision period. Participants were preexposed to the five familiar scenes in a brief task immediately preceding the card task; each scene was presented five times (randomly ordered), and participants were asked to indicate whether each image displayed an indoor or outdoor scene.

We designed the experiment to determine whether incidental contextual novelty influences which type of memory is used to make value-based decisions. A challenge for this aim is that incremental learning, by definition occurs over repeated experiences; thus, manipulations to incremental learning are best assessed across trials. With this in mind, we divided the experiment into two blocks, which were designed to manipulate the use of episodically learned values, but which could be used to measure how our manipulation impacts both episodic and incremental memory use. Specifically, we used the results of our related prior experiments to position contextual scenes such that their combined influence during object value encoding (trials with two new cards) and retrieval (trials with one old card) would either drive participants toward or away from using episodic memory.

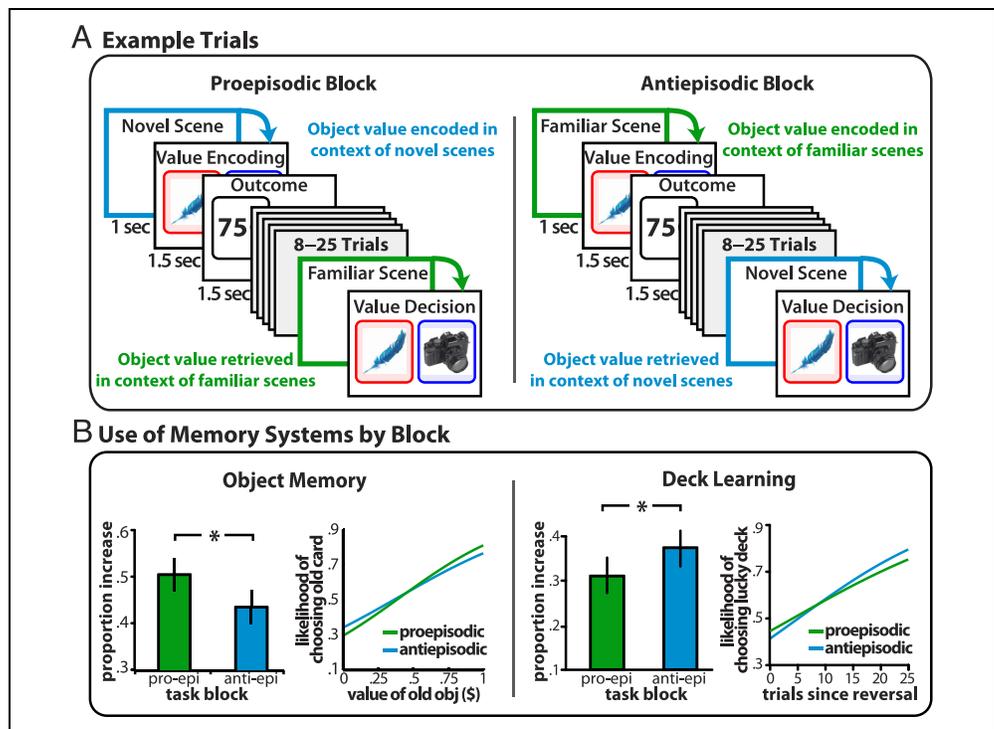
We previously found that novel images enhance object value encoding, whereas familiar images enhance object value retrieval. We used this manipulation to create blocks that enhance either encoding or retrieval of episodic memories to guide choice. We created “proepisodic” blocks by having novel scenes always precede new–new trials, with the aim of enhancing episodic encoding of the new cards (Figure 4A) and having familiar scenes precede old–new trials, with the aim of enhancing the retrieval and use of object value memories. In “antiepisodic” blocks, we did the reverse: Familiar scenes always preceded new–new trials, whereas novel scenes preceded old–new trials. Participants performed one block of each (approximately 150 trials), and the order of blocks was counterbalanced across participants. Because each scene was presented immediately before and concurrently with specific cards, the scenes could additionally become associated with cards in memory, and thus, their content could also prime memories for card values. To avoid this possibility, the trial sequences were designed such that the same scene was never present during both the learning of a card’s value and the subsequent retrieval and use of that value. Additionally, when possible, novel/familiar scene status was reversed for pairs of participants such that similar choices for one participant would be made in the opposite experimental condition as another participant.

We hypothesized that the novel/familiar quality of the scene itself, regardless of the actual content of previous associations with the scene, would influence how people use memory to make choices, with novel scenes enhancing encoding of episodic value and familiar scenes enhancing retrieval of episodic value. In line with the results of Experiments 1A and 1B, we additionally hypothesized that modulating the encoding and use of episodic memory would impact the extent to which participants incrementally learn about deck values.

## Results and Discussion

As with Experiment 1, participants’ choices were influenced by both the previously experienced values of

**Figure 4.** Design and results of Experiment 2. (A) Trials were divided into blocks designed to manipulate the use of episodically learned values. In proepisodic blocks, object values were always encoded in the context of novel scenes. These objects were then later repeated in the context of familiar scenes. The antiepisodic blocks had the reverse scene contingencies. Scenes were described as “decorative mats” and were not predictive of card content or value. (B) Object values were more often used in pro- as compared with antiepisodic blocks, whereas deck learning was superior in anti- as compared with proepisodic blocks. The graphs on the left panel plot how well choices (old vs. new object) were predicted by the old object values in each memory block. The graphs in the right panel plot how likely the lucky deck was to be chosen based on the number of trials that elapsed since a reversal in deck luckiness. \* $p < .05$ .



object cards ( $\beta = 2.05$ , 95% CI = [1.64, 2.47],  $p < .00001$ ) and by deck luckiness ( $\beta = 0.18$ , 95% CI = [0.10, 0.26],  $p < .0001$ ). Additionally, the use of object and deck values was not significantly correlated across participants,  $r(36) = .08$ ,  $p = .62$ , suggesting that participants were able to independently use both types of memories to guide their choices. Object value was a nonsignificant negative predictor of choice in 2 of 38 participants, and these participants were removed from subsequent analyses. Including these participants does not change the pattern of results presented below.

As predicted, the degree to which participants' choices relied on episodic versus incremental learning differed as a function of the experimental block manipulation. A direct comparison of choices in each of the two conditions revealed that participants were significantly more likely to base their choices on memory for distinct object value experiences in the proepisodic block as compared with the antiepisodic block ( $\beta = 0.50$ ,  $p = .03$ ; Figure 4B). Conversely, in the antiepisodic block, participants were more likely to use incrementally learned deck values to guide their choices ( $\beta = -0.013$ , 95% CI = [-0.03, 0],  $p = .05$ ; Figure 4B), as measured by a more shallow learning curve in the proepisodic block as compared with the antiepisodic block. Thus, when choices can be driven by either episodic or incremental learning, familiar contexts enhance the likelihood that people's choices will

be guided by episodically rather than incrementally learned value. Importantly, familiarity was manipulated in a contextual feature, which was incidental to the primary card game; and thus, these results unambiguously support a general state mechanism over a selective attention mechanism.

## GENERAL DISCUSSION

With multiple distinct learning and memory systems at its disposal (Poldrack et al., 2001; Squire & Zola, 1996; Knowlton et al., 1994; McDonald & White, 1993), the human brain has the capacity to represent past experiences in several ways, from extracting regularities across similar experiences and distilling decontextualized values to storing rich idiosyncratic details of individual events. Far from redundant, these different types of memory representations have the potential to drive behavior in different directions, raising crucial questions about how and when different memory systems guide choices. Here, we devised a new approach to estimate the independent influence of episodic memories from incremental learning on people's choices. Designing a task in which both sources of value learning could adaptively guide behavior allowed us to also identify a factor that arbitrates between the use of these different types of memory—recent mnemonic processing. Specifically, people were more likely to use episodic memories on the trials following retrieval of an

episodic memory, even when the recently retrieved memory had no bearing on the task at hand.

Our results suggest that the simple act of episodic retrieval primes our brains to retrieve other, ostensibly unrelated episodic memories within the ensuing seconds. The demonstration of this behavioral phenomenon confirms the predictions of cholinergic models of hippocampal function (Easton et al., 2012; Meeter et al., 2004; Hasselmo et al., 1996). According to these models, cholinergic modulation of the hippocampus establishes sustained biases, which shape hippocampal processing for seconds (Pabst et al., 2016; Hasselmo & Fehrlau, 2001). Higher cholinergic levels, evoked by detecting novel contexts (Giovannini et al., 2001), are thought to bias the hippocampus toward forming distinctive memories by suppressing the reactivation of related associations. Conversely, lower cholinergic levels, evoked by recognizing familiar contexts, are thought to bias the hippocampus toward memory reactivation, such as the value associated with a particular object card. Thus, this framework predicts that use of episodic memories should be increased following familiarity or in the presence of familiar contexts. Consistent with this prediction, our prior research has demonstrated that recent recognition of familiar images improves people's ability to recall other associations (Patil & Duncan, 2018) and use memories to guide decisions (Duncan & Shohamy, 2016). Our prior research, however, was restricted to testing situations where only episodic memory cues are available to guide behavior. We, thus, provide an important extension here by demonstrating that these same contextual familiarity manipulations have the power to bias the competition between different types of memory in more realistic (and complex) situations where both memory systems vie for behavior control.

We demonstrated this in two ways: First, people were more likely to use episodic object memories on trials following the successful use of another episodic object memory (Experiments 1A and 1B), and second, these memories were also more often used following the incidental presentation of a familiar image (Experiment 2). The second finding provides stronger support for our memory state framework, because multiple factors could contribute to the within-task autocorrelation in object memory use observed in the first experiments. For example, successfully using an object memory may bias attention toward the objects on cards on the next trial (hence, away from deck color). Of note, the magnitude of memory modulation observed in Experiments 1A and 1B was substantially larger than that observed in Experiment 2. This difference suggests that attention-based biases may also contribute to the original effects. However, it should be noted that we manipulated familiarity in very different ways across experiments and that the within-task manipulations employed in the first experiments may have been particularly potent at evoking memory states. Specifically, in this manipulation, preceding trials were coded as

“familiar” only when there was evidence that participants recognized and used the object memories. Conversely, in the incidental manipulation, contextual images were coded as “familiar” whenever they were repeated—a procedure that does not account for participants' recognition of the familiarity. Our previous work suggests that the objective presence of familiarity does not elicit episodic memory states as strongly as the subjective recognition of images (Patil & Duncan, 2018; Duncan & Shohamy, 2016; Duncan et al., 2012). Thus, the effects observed in the within-task manipulation may have also been strengthened by identifying the subjective mnemonic experience associated with our manipulation, a step that we could not take with the incidental manipulation because we had no behavioral index of the mnemonic experience triggered by the scene image.

Inspired by the cholinergic framework, we yoked recent familiarity manipulations to the episodic dimension of the task (e.g., proepisodic block), treating its impact on incremental learning use as byproduct of episodic memory modulation. Novelty detection, though, triggers the release of multiple interacting neurochemicals (Avery & Krichmar, 2017; Schomaker & Meeter, 2015; Patel, Rossignol, Rice, & Machold, 2012; Mena-Segovia, Winn, & Bolam, 2008). Of particular relevance to incremental learning, salient, novel stimuli have been shown to elicit dopamine release (Lisman & Grace, 2005; Ljungberg, Apicella, & Schultz, 1992), and in the context of reinforcement learning, new cues evoke fast and robust phasic responses in putative dopamine neurons (Lak, Stauffer, & Schultz, 2016; Stauffer, Lak, Kobayashi, & Schultz, 2016). These early responses are not modulated by the cue's reinforcement history, unlike the smaller value-dependent responses observed a fraction of a second later (Lak et al., 2016). Given dopamine's role in reinforcement learning, value-insensitive novelty/salience responses have sparked hypotheses that novelty could either directly modify value learning by inflating the predicted value of new cues or more generally promote the exploration of these cues, without distorting prediction error computations (Guitart-Masip, Bunzeck, Stephan, Dolan, & Duzel, 2010; Wittmann, Daw, Seymour, & Dolan, 2008; Kakade & Dayan, 2002). Although the current study is not positioned to address the former hypothesis, it is notable that participants were less likely to explore new cards following the recognition of an old card in Experiments 1A and 1B. At first blush, this seems consistent with the hypothesis that novelty promotes exploration, but this interpretation is complicated by participants' general reluctance to explore novel options in these tasks. Furthermore, increases in postnovelty exploration were not observed in the more controlled Experiment 2, suggesting that contextual novelty is not always sufficient to bias exploration.

Our demonstration that multiple forms of memory contribute to choice also has important implications for understanding basic mechanisms of decision-making. Prompted by the discovery that striatal dopamine release

instantiates many of the properties of prediction errors (Bayer & Glimcher, 2005; Schultz, 1998), striatal learning mechanisms have been the primary target of reinforcement learning and decision-making research. More recently, however, there has been growing interest in hippocampal and episodic memory contributions to decision-making (Bornstein, Khaw, Shohamy, & Daw, 2017; Bornstein & Norman, 2017; Murty, FeldmanHall, Hunter, Phelps, & Davachi, 2016; Shadlen & Shohamy, 2016; Peters & Büchel, 2010), often envisioned either as a key component of model-based decision systems (Doll, Shohamy, & Daw, 2015; Wimmer, Daw, & Shohamy, 2012) or as a distinct control system (Gershman & Daw, 2017; Lengyel & Dayan, 2008). The deck dimension of our task was designed after the sort of “two-armed bandit” tasks routinely used to assess striatal contributions to decision-making. Of note, when participants were given the option of using this source of information alongside episodic memories, they readily used episodic memories. In fact, object values were nearly as reliable a predictor of choice as deck values, despite participants only having a single prior experience with each object. This underscores the importance of incorporating episodic, one-shot learning mechanisms into theories of reinforcement learning. It also raises interesting questions about the balance between these different forms of learning across species. Are humans uniquely predisposed to use episodic over incrementally learned values, and if so, what benefits might be conveyed by this predisposition?

In summary, we used neural models of learning and memory to develop a new memory modulation approach, an approach that modulates which memory system guides behavior. The approach we identified—eliciting memory recognition—was found to be contextual in nature and, as such, has the potential to impact behavior broadly. It is additionally notable that our approach appears to enhance rather than inhibit the use of episodic memories. Striatal-dependent incremental learning is thought to be particularly robust (Schwabe & Wolf, 2013), requiring less attentional or cognitive resources than episodic memory (Foerde, Knowlton, & Poldrack, 2006). Thus, identifying a factor that could increase the more fragile use of episodic memories has important implications for the more common need to overcome striatal habits. Lastly, identifying factors that impact the use of different memory systems is a much-needed complement to recent research investigating the arbitration of these memory systems during learning (Lee, O’Doherty, & Shimojo, 2015). Combined, these learning and use factors will provide new insights into how people adaptively make use of multiple forms of memory and how this arbitration can be brought back into registration when it breaks down.

Reprint requests should be sent to Katherine Duncan, 100 St George St., 4th Floor, Sidney Smith Building, Toronto, Ontario M5S 3G3, Canada, or via e-mail: duncan@psych.utoronto.ca.

## REFERENCES

- Avery, M. C., & Krichmar, J. L. (2017). Neuromodulatory systems and their interactions: A review of models, theories, and experiments. *Frontiers in Neural Circuits*, *11*, 108.
- Barto, A. (1995). Adaptive critic and the basal ganglia. In J. C. Houk, J. L. Davis, & D. Beiser (Eds.), *Models of information processing of the basal ganglia* (pp. 215–232). Cambridge, MA: MIT Press.
- Bayer, H. M., & Glimcher, P. W. (2005). Midbrain dopamine neurons encode a quantitative reward prediction error signal. *Neuron*, *47*, 129–141.
- Bates, D., Maechler, M., Bolker, B., & Walker, S. (2015). Fitting linear mixed-effects models using lme4. *Journal of Statistical Software*, *67*, 1–48.
- 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.
- Bornstein, A. M., & Norman, K. A. (2017). Reinstated episodic context guides sampling-based decisions for reward. *Nature Neuroscience*, *20*, 997–1003.
- Clark, R. E., Broadbent, N. J., Zola, S. M., & Squire, L. R. (2002). Anterograde amnesia and temporally graded retrograde amnesia for a nonspatial memory task after lesions of hippocampus and subiculum. *Journal of Neuroscience*, *22*, 4663–4669.
- Cohen, N. J., & Eichenbaum, H. (1993). *Memory, amnesia, and the hippocampal system*. Cambridge, MA: MIT Press.
- Davachi, L., Mitchell, J. P., & Wagner, A. D. (2003). Multiple routes to memory: Distinct medial temporal lobe processes build item and source memories. *Proceedings of the National Academy of Sciences, U.S.A.*, *100*, 2157–2162.
- Daw, N. D. (2011). Trial-by-trial data analysis using computational models. In *Decision making, affect, and learning: Attention and performance XXIII (Vol. 23)*, pp. 3–38). Oxford: Oxford University Press.
- Daw, N. D., O’Doherty, J. P., Dayan, P., Seymour, B., & Dolan, R. J. (2006). Cortical substrates for exploratory decisions in humans. *Nature*, *441*, 876–879.
- Doll, B. B., Shohamy, D., & Daw, N. D. (2015). Multiple memory systems as substrates for multiple decision systems. *Neurobiology of Learning and Memory*, *117*, 4–13.
- Douchamps, V., Jeewajee, A., Blundell, P., Burgess, N., & Lever, C. (2013). Evidence for encoding versus retrieval scheduling in the hippocampus by theta phase and acetylcholine. *Journal of Neuroscience*, *33*, 8689–8704.
- Duncan, K., Sadanand, A., & Davachi, L. (2012). Memory’s penumbra: Episodic memory decisions induce lingering mnemonic biases. *Science*, *337*, 485–487.
- Duncan, K. D., & Shohamy, D. (2016). Memory states influence value-based decisions. *Journal of Experimental Psychology: General*, *145*, 1420–1426.
- Easton, A., Douchamps, V., Eacott, M., & Lever, C. (2012). A specific role for septohippocampal acetylcholine in memory? *Neuropsychologia*, *50*, 3156–3168.
- Eldridge, L. L., Knowlton, B. J., Furmanski, C. S., Bookheimer, S. Y., & Engel, S. A. (2000). Remembering episodes: A selective role for the hippocampus during retrieval. *Nature Neuroscience*, *3*, 1149–1152.
- Foerde, K., Knowlton, B. J., & Poldrack, R. A. (2006). Modulation of competing memory systems by distraction. *Proceedings of the National Academy of Sciences, U.S.A.*, *103*, 11778–11783.
- Frank, M. J., Seeberger, L. C., & O’Reilly, R. C. (2004). By carrot or by stick: Cognitive reinforcement learning in Parkinsonism. *Science*, *306*, 1940–1943.
- Gershman, S. J., & Daw, N. D. (2017). Reinforcement learning and episodic memory in humans and animals:

- An integrative framework. *Annual Review of Psychology*, 68, 101–128.
- Gilboa, A., Winocur, G., Rosenbaum, R. S., Poreh, A., Gao, F., Black, S. E., et al. (2006). Hippocampal contributions to recollection in retrograde and anterograde amnesia. *Hippocampus*, 16, 966–980.
- Giovannini, M. G., Rakovska, A., Benton, R. S., Pazzagli, M., Bianchi, L., & Pepeu, G. (2001). Effects of novelty and habituation on acetylcholine, GABA, and glutamate release from the frontal cortex and hippocampus of freely moving rats. *Neuroscience*, 106, 43–53.
- Guitart-Masip, M., Bunzeck, N., Stephan, K. E., Dolan, R. J., & Duzel, E. (2010). Contextual novelty changes reward representations in the striatum. *Journal of Neuroscience*, 30, 1721–1726.
- Hasselmo, M. E., & Fehrlau, B. P. (2001). Differences in time course of ACh and GABA modulation of excitatory synaptic potentials in slices of rat hippocampus. *Journal of Neurophysiology*, 86, 1792–1802.
- Hasselmo, M. E., Wyble, B. P., & Wallenstein, G. V. (1996). Encoding and retrieval of episodic memories: Role of cholinergic and GABAergic modulation in the hippocampus. *Hippocampus*, 6, 693–708.
- Kakade, S., & Dayan, P. (2002). Dopamine: Generalization and bonuses. *Neural Networks*, 15, 549–559.
- Knowlton, B. J., Squire, L. R., & Gluck, M. A. (1994). Probabilistic classification learning in amnesia. *Learning & Memory*, 1, 106–120.
- Lak, A., Stauffer, W. R., & Schultz, W. (2016). Dopamine neurons learn relative chosen value from probabilistic rewards. *eLife*, 5, e18044.
- Lee, S. W., O'Doherty, J. P., & Shimojo, S. (2015). Neural computations mediating one-shot learning in the human brain. *PLOS Biology*, 13, e1002137.
- Lengyel, M., & Dayan, P. (2008). Hippocampal contributions to control: The third way—supporting material. *Advances in Neural Information Processing Systems*, 20, 889–896.
- Lisman, J. E., & Grace, A. A. (2005). The hippocampal-VTA loop: Controlling the entry of information into long-term memory. *Neuron*, 46, 703–713.
- Ljungberg, T., Apicella, P., & Schultz, W. (1992). Responses of monkey dopamine neurons during learning of behavioral reactions. *Journal of Neurophysiology*, 67, 145–163.
- McDonald, R. J., & White, N. M. (1993). A triple dissociation of memory systems: Hippocampus, amygdala and dorsal striatum. *Behavioral Neuroscience*, 107, 3–22.
- Meeter, M., Murre, J. M., & Talamini, L. M. (2004). Mode shifting between storage and recall based on novelty detection in oscillating hippocampal circuits. *Hippocampus*, 14, 722–741.
- Mena-Segovia, J., Winn, P., & Bolam, J. P. (2008). Cholinergic modulation of midbrain dopaminergic systems. *Brain Research Reviews*, 58, 265–271.
- Murty, V. P., FeldmanHall, O., Hunter, L. E., Phelps, E. A., & Davachi, L. (2016). Episodic memories predict adaptive value-based decision-making. *Journal of Experimental Psychology: General*, 145, 548–558.
- Nadel, L., & Moscovitch, M. (1997). Memory consolidation, retrograde amnesia and the hippocampal complex. *Current Opinion in Neurobiology*, 7, 217–227.
- O'Reilly, R. C., & McClelland, J. L. (1994). Hippocampal conjunctive encoding, storage, and recall: Avoiding a trade-off. *Hippocampus*, 4, 661–682.
- Pabst, M., Braganza, O., Dannenberg, H., Hu, W., Pothmann, L., Rosen, J., et al. (2016). Astrocyte intermediaries of septal cholinergic modulation in the hippocampus. *Neuron*, 90, 853–865.
- Packard, M. G. (2009). Exhumed from thought: Basal ganglia and response learning in the plus-maze. *Behavioural Brain Research*, 199, 24–31.
- Patel, J. C., Rossignol, E., Rice, M. E., & Machold, R. P. (2012). Opposing regulation of dopaminergic activity and exploratory motor behavior by forebrain and brainstem cholinergic circuits. *Nature Communications*, 3, 1172.
- Patil, A., & Duncan, K. (2018). Linger cognitive states shape fundamental mnemonic abilities. *Psychological Science*, 29, 45–55.
- Pessiglione, M., Seymour, B., Flandin, G., Dolan, R. J., & Frith, C. D. (2006). Dopamine-dependent prediction errors underpin reward-seeking behaviour in humans. *Nature*, 442, 1042–1045.
- Peters, J., & Büchel, C. (2010). Episodic future thinking reduces reward delay discounting through an enhancement of prefrontal-mediocaudal interactions. *Neuron*, 66, 138–148.
- Poldrack, R. A., Clark, J., Paré-Blagoev, E. J., Shohamy, D., Creso Moyano, J., Myers, C., et al. (2001). Interactive memory systems in the human brain. *Nature*, 414, 546–550.
- Ranganath, C., Yonelinas, A. P., Cohen, M. X., Dy, C. J., Tom, S. M., & D'Esposito, M. (2004). Dissociable correlates of recollection and familiarity within the medial temporal lobes. *Neuropsychologia*, 42, 2–13.
- Schomaker, J., & Meeter, M. (2015). Short- and long-lasting consequences of novelty, deviance and surprise on brain and cognition. *Neuroscience & Biobehavioral Reviews*, 55, 268–279.
- Schönberg, T., Daw, N. D., Joel, D., & O'Doherty, J. P. (2007). Reinforcement learning signals in the human striatum distinguish learners from nonlearners during reward-based decision making. *Journal of Neuroscience*, 27, 12860–12867.
- 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, 772–781.
- Schultz, W. (1998). Predictive reward signal of dopamine neurons. *Journal of Neurophysiology*, 80, 1–27.
- Schultz, W. (2016). Dopamine reward prediction-error signalling: A two-component response. *Nature Reviews Neuroscience*, 17, 183–195.
- Schwabe, L., & Wolf, O. T. (2013). Stress and multiple memory systems: From “thinking” to “doing”. *Trends in Cognitive Sciences*, 17, 60–68.
- Scoville, W. B., & Milner, B. (1957). Loss of recent memory after bilateral hippocampal lesions. *Journal of Neurology, Neurosurgery, and Psychiatry*, 20, 11–21.
- Shadlen, M. N., & Shohamy, D. (2016). Decision making and sequential sampling from memory. *Neuron*, 90, 927–939.
- Small, S. A., Schobel, S. A., Buxton, R. B., Witter, M. P., & Barnes, C. A. (2011). A pathophysiological framework of hippocampal dysfunction in ageing and disease. *Nature Reviews Neuroscience*, 12, 585–601.
- Squire, L. R. (1992). Memory and the hippocampus: A synthesis from findings with rats, monkeys, and humans. *Psychological Review*, 99, 195–231.
- Squire, L. R., & Zola, S. M. (1996). Structure and function of declarative and nondeclarative memory systems. *Proceedings of the National Academy of Sciences, U.S.A.*, 93, 13515–13522.
- Stauffer, W. R., Lak, A., Kobayashi, S., & Schultz, W. (2016). Components and characteristics of the dopamine reward utility signal. *Journal of Comparative Neurology*, 524, 1699–1711.

- Sutton, R. S., & Barto, A. G. (1998). *Reinforcement learning: An Introduction (Vol. 1)*. Cambridge, MA: MIT Press.
- Tulving, E. (1983). Elements of episodic memory. *Canadian Psychology, 26*, 351.
- Tulving, E., & Markowitsch, H. J. (1998). Episodic and declarative memory: Role of the hippocampus. *Hippocampus, 8*, 198–204.
- Vandecasteele, M., Varga, V., Berényi, A., Papp, E., Barthó, P., Venance, L., et al. (2014). Optogenetic activation of septal cholinergic neurons suppresses sharp wave ripples and enhances theta oscillations in the hippocampus. *Proceedings of the National Academy of Sciences, U.S.A., 111*, 13535–13540.
- Volkow, N. D., Wang, G. J., Fowler, J. S., Tomasi, D., & Telang, F. (2011). Addiction: Beyond dopamine reward circuitry. *Proceedings of the National Academy of Sciences, U.S.A., 108*, 15037–15042.
- Wimmer, G. E., Daw, N. D., & Shohamy, D. (2012). Generalization of value in reinforcement learning by humans. *European Journal of Neuroscience, 35*, 1092–1104.
- Wittmann, B. C., Daw, N. D., Seymour, B., & Dolan, R. J. (2008). Striatal activity underlies novelty-based choice in humans. *Neuron, 58*, 967–973.
- Yassa, M. A., Mattfeld, A. T., Stark, S. M., & Stark, C. E. L. (2011). Age-related memory deficits linked to circuit-specific disruptions in the hippocampus. *Proceedings of the National Academy of Sciences, U.S.A., 108*, 8873–8878.