Published in final edited form as:

Neuropsychologia. 2018 June ; 114: 214–230. doi:10.1016/j.neuropsychologia.2018.04.023.

Manipulating memory efficacy affects the behavioral and neural profiles of deterministic learning and decision-making

Joshua J. Tremel^{a,b,c,*}, Daniella M. Ortiz^b, and Julie A. Fiez^{a,b,c,d}

^aDepartment of Psychology, University of Pittsburgh, Pittsburgh, PA, USA

^bLearning Research and Development Center, University of Pittsburgh, Pittsburgh, PA, USA

^cCenter for the Neural Basis of Cognition, University of Pittsburgh, Pittsburgh, PA, USA

^dCenter for Neuroscience, University of Pittsburgh, Pittsburgh, PA, USA

Abstract

When making a decision, we have to identify, collect, and evaluate relevant bits of information to ensure an optimal outcome. How we approach a given choice can be influenced by prior experience. Contextual factors and structural elements of these past decisions can cause a shift in how information is encoded and can in turn influence later decision-making. In this two-experiment study, we sought to manipulate declarative memory efficacy and decision-making in a concurrent discrimination learning task by altering the amount of information to be learned. Subjects learned correct responses to pairs of items across several repetitions of a 50- or 100-pair set and were tested for memory retention. In one experiment, this memory test interrupted learning after an initial encoding experience in order to test for early encoding differences and associate those differences with changes in decision-making. In a second experiment, we used fMRI to probe neural differences between the two list-length groups related to decision-making across learning and assessed subsequent memory retention. We found that a striatum-based system was associated with decision-making patterns when learning a longer list of items, while a medial cortical network was associated with patterns when learning a shorter list. Additionally, the hippocampus was exclusively active for the shorter list group. Altogether, these behavioral, computational, and imaging results provide evidence that multiple types of mnemonic representations contribute to experienced-based decision-making. Moreover, contextual and structural factors of the task and of prior decisions can influence what types of evidence are drawn upon during decision-making.

* Address correspondence to: Joshua J. Tremel (j.tremel@pitt.edu), University of Pittsburgh, 615 Learning Research and Development Center, 3939 O'hara Street, Pittsburgh, PA 15260.

Conflict of interest

The authors declare no competing conflicts of interest.

Publisher's Disclaimer: This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final citable form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

Keywords

Drift-diffusion model; Hippocampus; Memory; Feedback; Basal ganglia; Striatum

Prior experience is a rich resource for decision-making. By learning about outcomes and retrieving memories of prior decision experiences, decisions in the past can inform decisions in the present. These past experiences can be encoded in different ways, reflecting that different types of mnemonic representations and their associated neural systems (e.g., the medial temporal lobe or the striatum) can influence later behavior. However, disentangling how these systems work together or in opposition to utilize past decision experiences has been challenging. The current study manipulates the strength of declarative memories for past decision experiences and measures how this impacts the behavioral and neural profile of subsequent decisions.

The study examines decision-making in the context of a concurrent discrimination learning task. In this task, subjects learn to choose an arbitrarily predetermined correct item for each of a list of item pairs. In the first round, subjects guess at which item in each pair is correct and use this experience to guide their choices on subsequent rounds. In particular, the goal is to learn to choose items associated with positive feedback while avoiding those associated with negative feedback. Thus, the concurrent discrimination task offers a framework to investigate how improvements in deterministic decision-making are supported by the learning and memory systems of the human brain.

Prior neuropsychological findings suggest that multiple memory systems, including separate systems centering in the medial temporal lobe (MTL) and the striatum, contribute to decision-making in concurrent discrimination. Damage to the MTL declarative memory system often impairs concurrent discrimination learning in both monkeys (Zola-Morgan et al., 1989; Zola-Morgan et al., 1994; Buffalo et al., 1999) and humans (Squire et al., 1988; Squire and Zola, 1996; Corkin, 2002; Hood et al., 1999). However, factors such as long retention intervals or larger sets of items can help preserve learning (Malamut et al., 1984; Phillips et al., 1988; Suzuki et al., 1993; Gaffan & Murray, 1992; Buffalo et al., 1998; Hood et al., 1999; Bayley et al., 2005; Chudasama et al., 2008). This suggests that other, non-declarative memory systems can potentially compensate for damage to the MTL system. One such possible system is a striatal procedural memory system. For instance, lesions and pharmacological antagonism of the striatal system can impair concurrent discrimination learning in monkeys (Fernandez-Ruiz et al., 2001; Teng et al., 2000; Turchi et al., 2010). Similarly, in human participants with basal ganglia dysfunction, such as those with Parkinson's disease, intact concurrent discrimination learning can only be observed if participants also exhibit awareness of the task's deterministic structure, suggesting that declarative memory strategies can likewise compensate for striatal dysfunction (Moody et al., 2010). Taken together, it appears that both the MTL and striatal systems can support concurrent discrimination learning, though the conditions that govern their relative use remain poorly understood.

In the current study, functional magnetic resonance imaging (fMRI) is used to examine how the MTL and striatum contribute to decision-making in concurrent discrimination when both

of these key memory systems are intact. This study builds upon previous fMRI results reported by Tremel et al. (2016), in which computational models of feedback-based learning (model-free reinforcement learning) and decision-making (drift-diffusion) were used to model concurrent discrimination learning. Model parameters were used as regressors in an analysis of the hemodynamic signals underlying different aspects of learning and decision-making in concurrent discrimination. The study identified a network of brain regions associated with top-down decision control and a separate network of regions associated with bottom-up choice selection driven by the learned value of the stimuli. Of particular relevance, activity in the head of the caudate nucleus was associated with performance monitoring, wherein activity tracked with both decision thresholds and prediction errors in learning the value associated with stimuli. In contrast, activity in the putamen was associated with mnemonic information storage, wherein activity tracked with the accrual of stimulus value acquired from experiencing repeated outcomes. Across learning, decision-making shifted from controlled processing involving the caudate to stimulus-driven choice selection involving the putamen. Thus, decision-making during concurrent discrimination learning seemed to be driven in part by a striatum-based network whose role changed as the value of the stimuli became learned through repeated decision experiences.

Notably, Tremel et al. (2016) was unable to establish associations between the modeled parameters of concurrent discrimination decision performance and hemodynamic signal in the MTL. Given the neuropsychological evidence implicating the MTL in concurrent discrimination learning, this was an unexpected null result. The study therefore left unresolved the question of how declarative memory, acquired via an MTL system, helps to improve subsequent decision-making in concurrent discrimination. The current study examines this open question by creating differences in declarative memory between two groups of subjects via a list-length manipulation and then probing for group differences in the behavioral and neural profiles of concurrent discrimination memory and decision-making performance. The overarching hypothesis is that subjects with stronger declarative knowledge of the concurrent discrimination items will place greater reliance on a declarative memory strategy involving the MTL to perform the concurrent discrimination task, as compared to a procedural memory strategy involving the striatum.

The current study varies the number of item pairs (50 or 100) that subjects must learn as a way to manipulate declarative memory. Previous studies in monkeys have shown that subsequent memory for concurrent discrimination items tends to decrease with increasing list-lengths (Mishkin, 1982; Mahut et al., 1982). Beyond this, list-length manipulations have predictable and established effects on memory performance, including recollection (Wais et al., 2006) and free recall (Grasby et al., 1994). Moreover, these effects tend to impact MTL-based declarative memory more so than non-declarative memory such as implicit or procedural information (Jacoby, 1991; Yonelinas & Jacoby, 1994; Yonelinas, 2001). Thus, the current literature provides strong support for the hypothesis that performing the concurrent discrimination task with shorter list-lengths will result in stronger declarative memories for the items, as compared to performing the task with longer list-lengths.

The first goal is to examine how differences in declarative memories for task items influences the behavioral and neural profiles of concurrent discrimination learning.

Intuitively, stronger declarative memories for items should encourage further use of this knowledge to improve later choice selection. Indeed, given the deterministic structure of the concurrent discrimination task, perfect recall of the decision experiences in the first round could theoretically permit perfect performance in the subsequent rounds. Consistent with this prediction, there is some evidence that monkeys with MTL damage show poorer learning of the correct responses in early rounds of a concurrent discrimination task, as compared to monkeys without MTL damage (Malamut et al., 1984). Interestingly, however, both groups tend to require the same number of rounds to reach criterion levels of accuracy, suggesting that learning through different memory systems may have different but convergent trajectories. In humans with versus without MTL damage, however, the differences in learning trajectories are larger and more prolonged, though eventually criterion levels of performance can be achieved (Bayley et al., 2005). Thus, prior findings suggest that humans performing the concurrent discrimination task with a short versus long list of item pairs should exhibit a faster learning trajectory and greater engagement of the MTL system.

The second goal of the present study is to examine whether differences in declarative memories influence the behavioral and neural profile of concurrent discrimination decision-making. To address this, we again implement a drift-diffusion modeling approach, as we did in prior work (Tremel et al., 2016). This approach focuses on two key model parameters, decision threshold (a) and drift-rate (v). Changes in these model parameters can describe how subjects approach decision-making and how that approach changes across learning. Decision threshold captures the rigor of decision criteria, while drift-rate captures the mean rate of decision evidence accumulation (Ratcliff, 1978; Ratcliff and McKoon, 2008; Ratcliff et al., 2016).

Because the concurrent discrimination task has usually been studied as a learning task, the effects of list-length on concurrent discrimination decision-making have received little attention. However, the broader literature on declarative versus procedural learning strategies can be used to motivate some predictions. In general, declarative memory strategies are more effortful than procedural memory strategies (Mandler, 1980; Jacoby, 1991; Eichenbaum, 2001). This is why imposing fast response deadlines or dual-task demands can induce shifts from declarative to procedural memory strategies to support learning. In terms of the drift-diffusion model, one would expect higher decision thresholds to be associated with the short versus list conditions, capturing the greater degree of effort involved in retrieving and evaluating declarative information. Estimates of decision thresholds derived from these models can then serve as regressors for neuroimaging analyses to identify neural systems underlying the differences in decision-making due to the list-length manipulation. If differences in declarative memory create differences in decision strategies, then group differences would be expected in the fit of the model parameters to the activity of different brain regions. In other words, brain regions that support a declarative memory strategy should better reflect the patterns of parameter estimates from the short versus long list-length group, while regions that support a procedural memory strategy should better reflect the parameter estimates from the long versus short list-length group. Overall, if the list-length manipulation produces differences in decision-making strategies, then group

differences should be observed in drift-diffusion models of the behavioral and neural data associated with concurrent discrimination task performance.

To summarize, we report two experiments that test whether manipulations of list-length affect declarative memories, concurrent discrimination learning, and concurrent discrimination decision-making. The first experiment examines the behavioral effects of a list-length manipulation. The design features a manipulation of the number of item pairs (50 or 100) that subjects experience across three rounds of a concurrent discrimination task. Learning is interrupted after the first round of encoding to examine initial memory state before engaging in subsequent decision-making. Examining memory quality after an initial round allows us to test for the predicted effect of list-length on declarative memory without additional learning effects caused by repeated sampling of the list items. By examining the subsequent two rounds of concurrent discrimination performance, we are able to test whether the two groups also show differences in learning and decision-making strategies. The second experiment uses neuroimaging (fMRI) to investigate whether a list-length manipulation affects the neural substrates of concurrent discrimination performance. As in Experiment 1, subjects perform the concurrent discrimination task with 50 or 100 item pairs. In this experiment, however, subjects perform eight rounds of the task to provide a longer window of learning and perform a memory test at the end of the task to leave the learning trajectory uninterrupted. The primary questions are: (1) whether the list-length manipulation produces group differences in the engagement of the MTL and in the neural signatures of learning, and (2) whether the list-length manipulation produces group differences in the neural regions that contribute to decision-making in concurrent discrimination. To address the first question, we use a voxelwise analysis to test for group differences in the hemodynamic profile of overall concurrent discrimination task performance and changes in this profile across rounds. To address the second question, we use a drift-diffusion modeling approach to test for group differences in decision thresholds and the fit of modeled parameters to the acquired hemodynamic signals.

Experiment 1: Memory after initial encoding

Materials and Methods, Experiment 1

Subjects—Forty healthy, native English speakers recruited from the University of Pittsburgh's Psychology Subject Pool participated in a one- to two-hour behavioral session. One subject was excluded due to a hardware failure, leaving 39 total subjects. Subjects were split into two groups and performed a concurrent discrimination learning task with either 50 pairs of words ($N = 19$, 16 female) or 100 pairs ($N = 20$, 15 female). Subjects ranged in age from 18–20 (mean 18.53 years) and 18–21 (mean 18.50 years) in the 50-pair and 100-pair groups, respectively. Informed consent was obtained from all subjects according to procedures approved by the University of Pittsburgh Institutional Review Board. Subjects received course credit for their participation.

Task—For each trial of the concurrent discrimination task, subjects were presented with a pair of words at the center of the screen in a single column for three seconds. One of the words in each pair was designated ahead of time as the correct (positive) choice, while the

other was incorrect (negative). After selecting one of the words with a key press during the display period, subjects received feedback about their decision. Feedback was displayed for 1.5 s, indicating a correct choice with three green checkmarks or an error with three red X's. Trials were separated with a 1.5 s inter-stimulus interval.

After all word pairs in the set (50 or 100) were presented for the first time (Round 1), subjects were then presented with a surprise memory test. Subjects were presented with words either from the concurrent discrimination task or from a list of distractor items (i.e., items that did not appear in the learning task). One word was displayed at a time, and subjects indicated with a key press whether they recognized the displayed word from the learning task (old) or not (new). Trials with response times (RT) faster than 0.3 s or slower than 6 s were removed from analysis. If subjects identified a word as "old," regardless of their accuracy, they were separately prompted to identify more detailed information about whether the word was associated with positive (a correct word) or negative (an incorrect word) feedback in the learning task (value memory probe).

Following the memory test, subjects returned to the concurrent discrimination task for two additional rounds, to examine whether the early trajectory of the learning curves differed by list-length. Subjects saw the same set of word pairs as in Round 1, but in a randomized sequence. Subjects again made choices for each pair and received feedback. Over the course of these two additional rounds, subjects learned to associate an item in each pair with positive or negative feedback (i.e., which choice was correct or incorrect). The PsychoPy software package was used for task presentation and data collection (Peirce, 2007; Peirce, 2008).

In behavioral pilot work, we tested several list-lengths (10, 12, 24, 50, 100, 200) and found that subsequent memory performance was consistently at ceiling for the shortest lists, but decreased for 50+ items. We also found that the learning trajectory was statistically consistent across all list-lengths, so we chose to use 50-pair and 100-pair lists to test both learning and subsequent memory effects under differing encoding burdens. Additionally, using 50- and 100-pair lists as opposed to shorter ones ensured that trial counts would be sufficient enough to detect memory effects when splitting by conditions such as selection and item value.

Stimulus materials—Word stimuli were acquired from the MRC Psycholinguistic Database (Coltheart, 1981). All words were one-syllable, three to five letters, three or four phonemes, and had a log(HAL) frequency of at least 7.0. Words were further constrained to have concreteness, imageability, and familiarity scores of at least 400. Additional word frequency information was obtained from the English Lexicon Project (Balota et al., 2007).

Each subject's stimulus list was generated by randomly selecting two (50-pair task) or four (100-pair task) lists from a set of eight pre-made 50-word lists each balanced and matched for word frequency and number of letters. For the 50-pair task, the two lists were individually randomized and paired, ensuring each subject had a unique set of word pairs, balanced for selected psycholinguistic criteria. For the 100-pair task, two pairs of lists were first concatenated (to form two 100-pair lists), then randomized and paired as in the 50-pair

task. Two additional lists were randomly selected from the remaining pre-made lists and used as distractor (new) items in the surprise memory test.

Behavioral analysis

Memory test: Surprise memory test data were analyzed by computing hit, miss, correct rejection (CR), and false alarm (FA) scores. Hits were defined as “old” items that subjects correctly identified as “old” items. False alarms were “new” items that subjects misidentified as “old” items. Hit scores were adjusted by subtracting the FA rate to adjust for response tendency (i.e., the propensity or response bias of a subject to respond “old”) (Morcom and Rugg, 2002). Judgments of detailed information about item value (i.e., explicitly identifying that an item was associated with positive or negative feedback) were analyzed separately as a proportion of the number of hits (subjects must have first reported recognizing the item before being able to identify further details).

Memory measures were assessed for differences by Item Value (selection of the item results in positive or negative feedback), List-length (50 items, 100 items), and Selection Status (item was selected or unselected in Round 1 of the concurrent discrimination task). Selection was included as a factor because it could potentially influence subsequent memory if items are encoded in different ways, such as on an individual (i.e., item-level) basis versus an event (i.e., pair-level) basis. If encoding occurs at an item level, selection may draw attention toward one particular item, leading to an increased ability to later recognize that item. However, if encoding occurs at a event-level, selection should have equal effect on recognition of both items in the pair. For item value memory, however, selection could indicate that action is essential for binding outcome information to stimulus and pair information. We expected to see evidence of a more item-focused encoding method for the longer list (i.e., assigning specific outcomes to specific items) and a more event-focused encoding method for the shorter list (i.e., encoding the pair, action, and outcome as a singular experience).

Mixed-effects repeated-measures ANOVAs were computed for corrected hit rates (recognition memory) and value memory accuracy including within-subjects factors of Item Value (positive, negative) and Selection Status (selected, unselected in Round 1), and a between-subjects factor of List-length (50, 100 items). These and all other statistical analyses were completed using the R environment for statistical computing (R Development Core Team, 2016) and the Companion to Applied Regression (CAR) package for R (Fox and Weisberg, 2011). For these and all other ANOVA analyses, the Greenhouse-Geiser sphericity correction was applied when Mauchly’s test indicated a violation of the sphericity assumption.

To assess response biases to the value memory probe, we computed the proportion of “positive” responses for the value memory test for FA trials. This measure reflects the baseline tendency of subjects to respond “positive” (or “correct”) to the value memory probe. Differences were assessed between list-length groups with a group t-test. Within each group, we also assessed whether there was a bias (i.e., whether the value was statistically different from 50%, which would indicate no bias) with one-sample t-tests.

Decision-making performance: Drift-diffusion models were fit separately for subjects who learned 50-pairs and those who learned 100-pairs. Models were fit to subject RT and choice data using the HDDM software package (Wiecki et al., 2013). HDDM uses a hierarchical Bayesian approach wherein subject-level and group-level parameters are estimated at the same time using Markov-chain Monte-Carlo (MCMC). In our models, mean drift-rate (v) and decision threshold (a) were free parameters, allowed to vary across the three levels of Round (ignoring the interrupting surprise memory test). This resulted in one value per Round for each of drift-rate and threshold, describing changes in behavior across learning in terms of these two parameters. The non-decision time (t_{er}) and starting point (z) parameters were estimated at the subject- and group-levels, while variance parameters (variances for drift, sv ; non-decision time, st ; and starting point, sz) were estimated only at the group level. Separate hierarchical models were fit to each of the List-length groups. Data were accuracy-coded such that the upper threshold (a) of the model corresponded to a correct choice (i.e., positive items) and the lower threshold (0) to errors (i.e., negative items).

Model parameters were estimated using three MCMC chains of 3,000 samples each, discarding the first 500 samples of each chain to ensure stabilization. After sampling, model convergence was evaluated both visually, by inspecting model posterior traces, and statistically with the Gelman-Rubin statistic, which compares between- and within-chain variance. This statistic was near 1.0 across all parameters for both List-lengths (maximum deviance from 1.0 was 0.008 for the 50-pair group and 0.005 for the 100-pair group), suggesting that the models reached proper convergence. For a third check, the Geweke statistic, which compares beginning- and end-chain variance, also indicated proper convergence.

Results, Experiment 1

Memory after initial round of encoding

Recognition Memory: We first sought to characterize whether subjects encountering different list-lengths encoded different types or qualities of mnemonic information after the first encoding experience (Table 1). After performing the first round of a concurrent discrimination task, subjects were interrupted with a surprise memory test. We tested for differences in corrected hit rates for recognition memory (Hit - FA) with a $2 \times 2 \times 2$ mixed-effects repeated-measures ANOVA including within-subjects factors of Item Value (whether selection of the item is assigned to result in correct or incorrect feedback) and Selection Status (whether the item was selected or not in Round 1) and a between-subjects factor of List-length (50, 100 items).

There was a main effect of Item Value ($F[1,37] = 21.54$, $p < 0.001$, $\eta_p^2 = 0.37$), indicating that recognition performance was better on average for positive (i.e., reinforced in the learning task) items than negative ones ($t = +4.51$). Likewise, a main effect of List-length ($F[1,37] = 4.50$, $p = 0.04$, $\eta_p^2 = 0.11$) indicated that subjects in the 50-pair group had better recognition memory on average compared to those in the 100-pair group ($t = +2.14$). These two factors did not interact ($F[1,37] = 0.02$, $p = 0.90$, $\eta_p^2 = 0.001$), indicating that the differences by Item Value did not depend on differences by List-length (Fig. 1a). Overall

false alarm rates for the recognition memory probe did not differ between the two List-length groups ($d[37] = -1.06$, $p = 0.29$).

Additionally, there was a main effect of Selection ($F[1,37] = 40.03$, $p < 0.001$, $\eta_p^2 = 0.92$), but it did not interact with List-length ($F[1,37] = 3.06$, $p = 0.09$, $\eta_p^2 = 0.08$), indicating that selecting an item in Round 1 generally benefits subsequent recognition (Fig. 2). This selection effect, however, did interact with List-length when Item Value is also considered (List-length $\times$ Outcome $\times$ Selection: $F[1,37] = 5.06$, $p = 0.03$, $\eta_p^2 = 0.12$). This three-way interaction illustrates that selection seems to be more important for the 100-pair list, wherein selected positive items are better recognized than unselected ones ($t[38] = 3.14$, $p < 0.01$). This difference is absent for the 50-pair group ($t[38] = 0.88$, $p = 0.38$). The difference score between selected and unselected positive items is statistically greater for the 100-pair group compared to the 50-pair group ($t[36.96] = 2.80$, $p < 0.01$) (Fig. 2). This suggests that selection is differentially important for each list-length group, but that this depends on whether the item was associated with positive or negative feedback. The 100-pair group exhibited sensitivity to selection status that subjects in the 50-pair group did not exhibit.

Item value memory: Subjects were also asked to explicitly report the value associated with the recognized item. For items that were selected in Round 1, the value could have been directly experienced in the form of positive (i.e., correct) feedback for the selection of positive items, and negative (i.e., incorrect) feedback for the selection of negative items. For items that were not selected in Round 1, the value could be determined by inference (i.e., positive feedback indicates that the unselected item has a negative value, whereas negative feedback indicates the unselected item has a positive value). There was a main effect of Item Value ($F[1,37] = 5.36$, $p = 0.03$, $\eta_p^2 = 0.13$), indicating that subjects were better at reporting item value associations about the positive items versus negative items (Fig. 1b). Furthermore, overall value memory was better for the 50-pair group compared to the 100-pair group (List-length, $F[1,37] = 4.52$, $p = 0.04$, $\eta_p^2 = 0.11$). However, these factors did not interact ($F[1,37] = 0.02$, $p = 0.89$, $\eta_p^2 = 0.00$), suggesting that the differences by list-length and outcome are general overall differences rather than dependent on each other. Additionally, there was no main effect of Selection ($F[1,37] = 0.11$, $p = 0.74$, $\eta_p^2 = 0.003$), nor did Selection interact with Outcome ($F[1,37] = 2.04$, $p = 0.16$, $\eta_p^2 = 0.05$), List-length ($F[1,37] = 1.49$, $p = 0.23$, $\eta_p^2 = 0.06$), or both factors together ($F[1,37] = 0.24$, $p = 0.63$, $\eta_p^2 = 0.006$). This indicates that while selection seems to differentially influence recognition memory, it does not similarly influence memory for feedback associations.

Response bias in item value memory: To evaluate whether the two list-length groups have different underlying biases to respond to the value memory probe, we computed the proportion of “positive” responses (i.e., number of “positive” responses divided by total number of false alarms) for false alarm trials (“new” items mistakenly identified as “old” items). There was an overall difference by list-length in this measure (Fig. 1c), wherein subjects who learned the 50-pair list had a tendency to respond “negative” more often than subjects who learned the 100-pair list ($d[37] = -2.03$, $p < 0.05$, $d = 0.65$). This reflects the fact that for the 100-pair condition, subjects were unbiased in their responses, responding “positive” and “negative” at roughly equal rates ($d[19] = 0.78$, $p = 0.45$, test value = 50%).

For the 50-pair condition, however, subjects were negatively biased in their responses, responding “negative” at a statistically higher rate than “positive” ($t[18] = 3.91, p < 0.005$). This suggests that participants in the two groups may evaluate their memories of item values in different ways after an initial round of learning.

Concurrent discrimination learning trajectory—Following the surprise memory test after Round 1, subjects resumed the concurrent discrimination task for two additional rounds (Fig. 3). We tested for differences in learning with a mixed-effects repeated-measures ANOVAs with factors of Round (rounds 1–3) and List-length (50-pairs, 100-pairs), and accuracy as the dependent measure. There was a main effect of Round ($F[1.73, 64.10] = 74.80, p < 0.001, \eta_p^2 = 0.67$), indicating that overall accuracy increased across repetitions of the item set. There was, however, no main effect of List-length ($F[1, 37] = 1.31, p = 0.26, \eta_p^2 = 0.03$) nor an interaction between factors ($F[1.73, 64.10] = 1.83, p = 0.17, \eta_p^2 = 0.05$). This pattern was also evident in response times (RT, Fig. 3b). There was a main effect of Round ($F[1.64, 60.53] = 26.08, p < 0.001, \eta_p^2 = 0.41$), but no main effect of List-length ($F[1, 37] = 0.14, p = 0.71, \eta_p^2 = 0.00$) nor an interaction ($F[1.64, 60.53] = 1.48, p = 0.24, \eta_p^2 = 0.04$). This result indicates that subjects in both the 50- and 100-pair groups learned to a similar degree and at a similar rate.

Concurrent discrimination decision-making—A drift-diffusion model analysis was used to evaluate the profile of decision-making processes across all three rounds of the concurrent discrimination task. Model parameters can characterize subjects’ approaches to decision-making. Thus, if List-length influences decision processes, the modeled parameter values across Rounds of learning should differ between list-length conditions. Seven-parameter diffusion models were built for each of the two List-length groups with threshold (a) and drift-rate (v) parameters free to vary across Round. Average values for the free parameters, which served as dependent measures in 2×3 ANOVAs that included factors of List-length and Round, are enumerated in Table 2.

Drift-rate was found to vary across Round ($F[1.79, 66.18] = 294.86, p < 0.001, \eta_p^2 = 0.89$), corresponding to a linear increase in drift-rates over learning. There was no main effect of List-length ($F[1, 37] = 0.04, p = 0.84, \eta_p^2 = 0.01$) or interaction between the factors for drift-rate ($F[1.79, 66.18] = 2.45, p = 0.10, \eta_p^2 = 0.06$). This suggests that the repeated decision experiences similarly improved the quality of decision evidence for both List length conditions.

For decision thresholds, there was also a main effect of Round ($F[1.55, 57.37] = 46.24, p < 0.001, \eta_p^2 = 0.56$), indicating a fluctuation in decision thresholds across learning, on average. There was no main effect List-length ($F[1, 37] = 0.99, p = 0.33, \eta_p^2 = 0.03$). However, there was an interaction between List-length and Round ($F[1.55, 57.37] = 5.30, p < 0.05, \eta_p^2 = 0.09$), indicating that the pattern of changes in decision thresholds across learning depended on the number of items that subjects learned in the task (Fig. 3). Subjects learning the 50-pair list raised their decision thresholds higher in Round 2 than subjects learning the 100-pair list, even though both groups had similar levels of evidence quality across learning (i.e., drift rate).

Discussion, Experiment 1

List-length effects on memory—The results from Experiment 1 confirm the predicted effects of List-length on overall recognition memory for the concurrent discrimination items. This overall memory difference, wherein subjects learning the shorter list performed better than those learning the longer list is consistent with prior findings using a list-length manipulation with discrimination learning tasks (Mishkin, 1982; Mahut et al., 1982) and with the broader memory literature (Wais et al., 2006; Grasby et al., 1994). A finer-grained examination of item recognition performance identified two additional interactions with the list-length manipulation that provide leverage in interpreting the nature of the encoded representations.

First, the action of selecting an item in the concurrent discrimination task appears to differentially impact recognition memory based on list-length. For a longer list, selecting an item improves later recognition when the item is associated with positive feedback. However, for the shorter list, this effect is absent. This suggests that the underlying encoding processes in Round 1 related to item-level recognition memory may differ between groups. For the 50-pair list, subjects may be more likely to build event-level representations that encode information about both items in the pair, what action was taken, and what outcome was observed. Regardless of which item was chosen, both items are encoded to a similar extent. In contrast, for the 100-pair list, subjects may be more likely to build item-level representations, such that selecting an item and receiving positive feedback provides the information needed to successfully guide subsequent selections of the same item. Information about the unselected item may be encoded to a lesser extent, if at all.

Second, subjects in the short list condition exhibited a propensity to confabulate a negative value for “new” items they mistakenly recognized, while subjects in the long list condition exhibited no such response bias. The negative response bias present in the 50-pair condition suggests that subjects learning this list may encode or evaluate their Round 1 experiences in qualitatively different ways compared to the 100-pair condition. One possible explanation is that subjects in the 50-pair condition have a greater awareness of their declarative knowledge for items and in particular, are aware that their memories for positive items are stronger than those for negative items. Assuming that subjects are less confident of their memories on false-alarm trials, they may then assume that this weakly remembered item was associated with negative feedback. However, despite this bias in item value memory, subjects in this shorter list condition are able to successfully overcome it and perform well (and better than the 100-pair group) in correctly identifying value information for positive trials.

List-length effects on learning—The examination of the accuracy data revealed that list-length did not affect the rate of concurrent discrimination learning, even though the memory tests suggest that subjects in the 100-pair condition entered Round 2 with poorer declarative memory for the items as compared to those in the 50-pair condition. This result runs counter to predictions from the literature, wherein humans with MTL damage typically require many more rounds than usual to learn the correct responses (Bayley et al., 2005; Hood et al., 1999). Thus, from Experiment 1 alone, it is unclear what benefit, if any, the MTL system provides to task performance in concurrent discrimination (though subsequent

memory benefits are clear). Experiment 2 further considers this apparent discrepancy by using fMRI to test for MTL involvement and by examining a longer learning trajectory that might better reveal effects of list-length.

List-length effects on decision-making—A valuable feature of the list-length manipulation is that it targets declarative memory efficacy without changing the fundamental structure of the concurrent discrimination task. Irrespective of list length, each trial of the concurrent discrimination task requires that one of the two items be selected as a response. This facilitates the ability to directly compare the task conditions and decision-making processes without the confounds associated with many other manipulations of control and memory efficacy, such as reaction time pressure or dual-task demands. Drift-diffusion modeling was used to facilitate this comparison. List-length effects were not observed for the drift-rate parameter, which is consistent, in part, with the absence of accuracy differences between the two groups. Since accuracy is the end result of a drift-diffusion process, it should reflect the mean rate of evidence accumulation (drift-rate) when the boundaries are accuracy-coded (i.e., positive drift rate means that a choice is more likely to be correct for that condition). This suggests that overall memory quality, as captured by the drift-rate parameter, was similar between the two condition in the context of performing the concurrent discrimination task.

Notably, group differences were observed for the decision threshold parameter. The 50-pair condition exhibited a large increase in decision threshold in Round 2. This increase could capture the idea that declarative memory engagement and retrieval requires a more effortful approach (Mandler, 1980; Jacoby, 1991; Eichenbaum, 2001). In contrast, the 100-pair group exhibits a smaller increase in decision threshold in Round 2, which is maintained into Round 3. This suggests that subjects learning this longer list evaluate decisions in Round 2 with less strict criteria than their short list counterparts. These threshold differences may reflect group differences in declarative memory for the task items. Since short-list subjects have stronger recognition and value memory compared to long-list subjects, this may encourage the former to rely upon a declarative memory retrieval strategy for subsequent decisions and set a high standard for the amount of evidence needed before reaching a decision in Round 2. In contrast, faced with poorer overall declarative memory, especially for negative items, long-list subjects may fall back to a less effortful “gut instinct” approach (e.g., relying on familiarity of procedural information) which focuses on the positive-valued item in each pair at the expense of retaining information about the negative item.

It is important to point out, however, that by measuring memory quality early on in learning, subjects may have acquired a heightened awareness to the quality of their memory. This could inadvertently prompt subjects to explicitly evaluate their memory during subsequent rounds of learning and therefore approach the task differently than they normally would. In Experiment 2, a declarative memory test is conducted after an extended eight round concurrent discrimination task, without interrupting learning. This provides an opportunity to see whether similar list-length manipulations can be observed even though explicit memory evaluation is deferred until after concurrent discrimination learning has been completed.

Experiment 2: Neural systems underlying learning and memory processes

In a second experiment, we examined a longer trajectory of learning using fMRI to identify neural substrates associated with the list-length manipulation and its impact on concurrent discrimination decision-making. Similar to the design of Experiment 1, subjects in this experiment performed an extended, eight-round concurrent discrimination task in an fMRI scanner, with either 50-pair or 100-pair lists. The imaging session was followed by a surprise memory test to assess memory after the completion of the concurrent discrimination task.

Materials and Methods, Experiment 2

Subjects—Forty-six healthy, right-handed native English speakers with normal or corrected-to-normal vision participated in a behavioral and fMRI session. Subjects were split into two groups to perform a concurrent discrimination learning task with 50 items ($N = 20$) or 100 items ($N = 27$). fMRI scans lasted 50 minutes for the 50-pair group and 85 minutes for the 100-pair group. Across these groups, 13 subjects were excluded from analysis due to incomplete data from technical problems ($N = 3$), excessive movement during scanning ($N = 7$), or withdrawing from the study before completion ($N = 3$). This left 16 (12 female) and 17 (14 female) subjects in the 50-pair and 100-pair groups, ranging in age from 19–31 (mean 22.1) and 18–33 (mean 22.6) years, respectively. Informed consent was obtained from all subjects according to procedures approved by the University of Pittsburgh Institutional Review Board. Subjects were compensated \$75 for their time. Data from the 50-pair group have been published previously (Tremel et al., 2016).

Task and stimulus materials—Subjects performed a 50-pair or 100-pair concurrent discrimination task during an fMRI scan session. The overall structure of the concurrent discrimination task was the same as that in Experiment 1. However, instead of a fixed 1.5 s inter-stimulus interval, trials were separated with a variable inter-stimulus interval ranging from 1.5–7.5 s, pulled from a positively skewed distribution (1.5 s increments, mean trial separation 2.91 s). This interval served as jitter between trials to facilitate later deconvolution of the fMRI signal (Dale, 1999). After all word pairs in the set (50 or 100) were presented, subjects were allowed to rest for a minute before the next Round began, wherein subjects saw the same set of word pairs again in a randomized sequence. In contrast to Experiment 1, subjects performed eight rounds of the task with no interruptions. Immediately following the fMRI session, subjects then engaged in a surprise memory test, identical to the interrupting memory test in Experiment 1. Word stimuli were identical to those in Experiment 1, as were the procedures for generating lists for each subject.

The tasks were presented using E-prime software (Psychology Software Tools, Pittsburgh, PA). For the scan session, the task was projected onto a screen at the head of the magnet bore using a BrainLogics MRI Digital Projection System. Subjects viewed the screen via a mirror attached to the radio frequency (RF) coil and indicated their response using a fiber optic response glove on their right hand connected to the presentation computer via a serial response box (BrainLogics, Psychology Software Tools, Pittsburgh, PA). Earplugs were provided to minimize discomfort due to scanner noise.

Behavioral analysis

Memory test: As in Experiment 1, surprise memory test data were analyzed by computing hit, miss, correct rejection (CR), and false alarm (FA) scores. These measures were assessed for differences by Item Value (positive or negative feedback in the discrimination learning task) and List-Length (50 items, 100 items) with mixed repeated-measures ANOVAs.

Concurrent discrimination task: Similar to Experiment 1, drift-diffusion models were fit separately for subjects who learned 50-pairs and those who learned 100-pairs. We again focused on the decision threshold (a) and drift-rate (v) parameters to explain variations in decision behavior associated with the list-length manipulation. For both List-lengths, we again expected that subjects would become more proficient decision-makers across Rounds, meaning that drift-rates should generally increase while thresholds should generally decrease across Rounds. Between List-lengths, however, we expected to replicate the pattern reported in Experiment 1, wherein subjects learning the shorter list would establish higher thresholds in early Rounds compared to the longer list. Given the extended learning period of this experiment (eight Rounds), we expected these thresholds to gradually converge as the discrimination responses become well-practiced and automatized.

Models were fit to subject RT and choice data with a similar procedure as described in Experiment 1. Mean drift-rate (v) and decision threshold (a) were free to vary across the eight levels of Round. This resulted in one value per Round for each of drift-rate and threshold, describing changes in behavior across learning in terms of these two parameters. The non-decision time (t_{er}) and starting point (z) parameters were estimated at the subject- and group-levels, while variance parameters (variances for drift, sv ; non-decision time, st ; and starting point, sz) were estimated only at the group level. Round-wise parameter estimates for drift-rate and threshold were intended for use as covariates in the imaging analyses.

Model parameters were estimated using three MCMC chains of 10,000 samples each, discarding the first 1,000 samples of each chain to ensure stabilization. After sampling, the Gelman-Rubin statistic indicated proper convergence for all parameters for both List-lengths (maximum deviance from 1.0 was 0.007 for the 50-pair group and 0.006 for the 100-pair group). To test whether the design of the primary model was indeed the best possible design, two alternative models were constructed for each List-length. One alternative model was designed with threshold (a) as the sole free parameter, holding drift-rate (v) constant, whereas another was designed with drift-rate (v) as free and threshold (a) as constant. Primary and alternative models were compared using the deviance information criterion (DIC), which measures the lack of fit of model estimates, penalizing for model complexity (e.g., degrees of freedom) (Spiegelhalter et al., 2002; Spiegelhalter et al., 2014). A lower DIC indicates a better fit, typically by a magnitude of 10 or more (Burnham and Anderson, 2004; Zhang and Rowe, 2014).

Imaging analysis

Image acquisition: Images were obtained on a Siemens Allegra 3-Tesla system. Anatomical images were acquired with a T1-weighted MP-RAGE sequence (repetition time [TR] = 1.54

s), echo time [TE] = 3.04 ms, flip angle [FA] = 8 degrees, inversion time [TI] = 800 ms, 192 sagittal slices, 1 mm³ isotropic voxels) and a T2-weighted spin-echo sequence (TR = 6.0 s, TE = 73 ms, FA = 150 degrees, 38 axial slices, 3.2 × 3.2 mm in-plane resolution, slices spaced 3.2 mm apart). Functional images sensitive to the BOLD contrast were acquired with a whole-brain echo-planar T2*-weighted sequence (TR = 1.5 s, TE = 25 ms, FA = 60 degrees, 29 axial slices, 3.125 × 3.125 mm in-plane resolution, 3.5 mm slice spacing). The first four images of each functional run were discarded to allow the magnetization state and RF signal to stabilize.

Functional image preprocessing: Image preprocessing and analysis was carried out using FIDL (Washington University, St. Louis). Imaging data were preprocessed to address noise and image artifacts, including within-TR slice-time acquisition adjustment, motion correction using a rigid-body translation and rotation algorithm (Snyder, 1996), within-run voxel intensity normalization to a mode of 1000 to facilitate cross-subject comparisons (Ojemann et al., 1997), and computation of a Talairach atlas space transformation matrix (Talairach and Tournoux, 1988). Single-subject analyses were conducted in data space. For group analysis, data were resampled to 2 mm isotropic voxels and transformed into Talairach atlas space.

Voxelwise imaging analysis: Subject-level data were analyzed using a voxelwise general linear model (GLM). In each subject's GLM, each time point was modeled as the sum of coded effects produced by model events and by error (Friston et al., 1994; Glover, 1999; Miezin et al., 2000; Ollinger et al., 2001). This approach makes no assumptions about the underlying shape of the blood-oxygen-level dependent (BOLD) response, but assumes that component signals sum linearly to produce the observed BOLD signal. Within each Round, a linear term captured signal drift across the run, and a constant modeled baseline signal. Each trial was modeled with a series of finite impulse response (FIR) basis functions wherein each time point was described with a delta function. The sequence of these functions corresponded to trial-level time series averages. These event-related effects were represented as an estimate of the percent of BOLD signal change relative to the baseline constant.

Voxelwise List-length ANOVA: Our first goal was to identify brain regions associated with the list-length manipulations. The primary purpose was to localize an MTL region that would putatively contribute to memory encoding and retrieval during and after concurrent discrimination learning. To do this, a GLM was created for each subject in which accuracy was modeled as a categorical event-related regressor, coding each decision (trial) as correct or as an error per round, modeled to 11 time points (16.5 s) across each trial. A 2-factor ANOVA was computed including a two-level between-subjects factor of List-length (50, 100 pairs) and a 11-level within-subjects factor of trial timepoint (one level for each modeled trial time point in the GLM). This ANOVA resulted in an image corresponding to the List-length × Time Point interaction. Voxels appearing in this image exhibit changes in event-related activity that statistically differs by List-length and modulates during a trial (i.e., a voxel whose activity is not flat over the time series of a trial). Peaks exceeding a z-transformed F-statistic of 5.0 were defined as center points for regions of interest, around

which 10 mm spheres were grown. This whole-brain interaction image was corrected for multiple comparisons (minimum z-transformed F-statistic of 3.0, $p = 0.05$ corrected, 45-voxel contiguity). Voxels within the spheres that failed to survive this correction (i.e., did not appear in the corrected image) were dropped from the regions of interest. Time series data were then extracted from these regions to examine the nature of the List-length effect (e.g., 50 condition > 100 condition or 100 condition > 50 condition).

As a secondary analysis, we also tested for neural differences due to List-length across learning. The goal of this analysis was to determine whether list-length influenced early and late stages of learning differently and what neural regions may account for that difference. We examined two epochs of learning corresponding to early (Rounds 1–4) and late learning (Round 5–8). A voxel-wise, mixed-effects repeated-measures ANOVA was then computed with factors of Epoch (early, late), List (50, 100), and Time (11 levels, one for each modeled time point in the single-subject GLMs). Because of low trial counts for error trials across subjects in the late learning epoch, this analysis tested correct trials only. Regions of interest were defined and corrected like described in the above procedure using the corrected interaction image for the Epoch x List x Time effect. Time series data were then extracted from the regions of interest to determine the direction of the effects underlying the interaction.

Voxelwise drift-diffusion model parameter ANOVA: Our second goal was to identify brain regions whose activity tracked with modulations in drift-diffusion parameters across Rounds in a manner that depended on List-length. A second GLM was created for each subject in which accuracy was modeled as a categorical event-related regressor, coding each decision (trial) as correct or incorrect modeled to 11 time points (16.5 s). Also included in each GLM were two covariates to describe changes in drift-diffusion model parameters across Rounds. These covariates captured correlation between changes in the parameter values and changes in the underlying BOLD signal. At the trial-level, the covariates were modeled as continuous measures across the trial with 11 FIR basis functions (11 time points, totaling 16.5 s) to describe how the correlation may change across a trial epoch. The covariate was entered as a single value per Round across the task. These covariates captured choice and RT behavior on the level of Round as described by the drift-diffusion model (i.e., all trials in Round 1 were assigned the Round 1 parameter value, all trials in Round 2 were assigned the Round 2 parameter value, etc.). Modulations in these parameters thus inherently reflected modulations across Rounds.

A 3-factor voxelwise ANOVA was then computed including a two-level within-subjects factor of Parameter (decision threshold, drift-rate), an 11-level within-subjects factor of Time Point (one for each trial-level modeled time point), and a two-level between-subjects factor of List-length (50 items, 100 items). The factor of Time Point was included to ensure that any significant effects would be related to trial-level fluctuations in activity (i.e., a main effect would correspond to voxels active, or not flat, during the task). The factor of Parameter measured the extent of correlation between one parameter and underlying BOLD signal versus the correlation of the other parameter and underlying signal (e.g., underlying activity correlating to a significantly greater degree for decision threshold over drift-rate). The factor of List-length, extrapolating from the results of Experiment 1, captured group-

level differences due to the relative encoding quality of declarative memories. This ANOVA ultimately resulted in an activation map of z-transformed F-statistics corresponding to the interaction of these three factors (i.e., trial-related activity that correlated with one parameter more so than the other in a manner that depended on List-length). This interaction separated activity that was statistically more attributable to modulations in one parameter versus the other in a manner that depended on List-length.

The interaction image was corrected for multiple-comparisons (minimum z-transformed F-statistic of 3.0, $p = 0.05$, 45-voxel contiguity). Contiguous voxel clusters that survived this correction were considered regions of interest. Time series data were then extracted from these regions of interest to examine the direction of the main effects of the List-length and Parameter factors. Regions were categorized according to the direction of these effects and which parameter the region was sensitive to. For example, a set of regions might be more sensitive to correlations with decision thresholds in the 50-pair list versus the 100-pair list.

Results, Experiment 2

Behavioral results

Memory performance after eight rounds of learning: Following an fMRI session of an eight-round concurrent discrimination task, subjects completed a surprise recognition and value memory test (Table 3, Fig. 4). Like in Experiment 1, we assessed corrected hit rates (Hit - FA) for recognition memory with a 2×2 repeated measures ANOVA with factors of Item Value (positive, negative) and List-length (50-pairs, 100-pairs). Selection was excluded as a factor because which items are selected varies across Rounds of the task, and the memory test occurred after all eight rounds instead of after Round 1. There was a main effect of Item Value ($F[1,31] = 12.32$, $p < 0.01$, $\eta_p^2 = 0.29$), indicating that on average across groups, recognition rates were higher for positive items than for negative items. There was no main effect of List-length on recognition memory ($F[1,31] = 1.98$, $p = 0.17$, $\eta_p^2 = 0.06$), indicating that on average, subjects were equally capable of recognizing items from the task regardless of how many items they had to learn. However, there was an interaction between Item Value and List-length ($F[1,31] = 54.31$, $p < 0.001$, $\eta_p^2 = 0.64$), indicating that recognition performance by outcome depended on list-length. To characterize this further, looking across list condition, the 50- and 100-pair groups were equally capable of recognizing positive old items ($t[31] = -1.02$, $p = 0.32$), but subjects who learned 50 items were better at recognizing negative old items versus the 100-pair group ($t[31] = 3.42$, $p < 0.01$). Looking within each group, subjects who learned the 100-pair list were better at recognizing positive items than negative items ($t[1,16] = 6.96$, $p < 0.001$), while subjects who learned the 50-pair list had the opposite tendency and were better at recognizing negative items over positive ones ($t[1, 15] = -3.36$, $p < 0.01$). Overall FA rates for recognition memory did not differ by List-length ($t[31] = -0.39$, $p = 0.70$).

When asked to explicitly identify more detailed information about the value associated with recognized items (i.e., whether the item was a positive item or a negative item in the scanned task), subjects in both groups performed equally (Fig. 4b). There were no main effects of Item Value ($F[1,31] = 1.76$, $p = 0.19$, $\eta_p^2 = 0.06$) or List-length ($F[1,31] = 0.53$, $p = 0.47$, $\eta_p^2 = 0.02$) and furthermore, no interaction between factors ($F[1,31] = 0.10$, $p = 0.76$, $\eta_p^2 =$

0.004). This suggests that while recognition memory performance varied by List-length and Item Value, memory for the value associated with recognized items was unaffected by these factors.

Concurrent discrimination performance and learning: Similar to Experiment 1, we again found that the learning curves were comparable regardless of List-length. In a 2×8 mixed-effects repeated-measures ANOVA with factors of Round and List-length, there was a main effect of Round on accuracy ($F[2.57, 79.81] = 205.12$, $p < 0.001$, $\eta_p^2 = 0.87$), indicating a general increase in accuracy across Rounds (Fig. 5a). List-length, however, had no effect on accuracy ($F[1, 31] = 1.12$, $p = 0.30$, $\eta_p^2 = 0.03$) and did not interact with Round ($F[2.57, 79.81] = 1.61$, $p = 0.20$, $\eta_p^2 = 0.05$). Thus, subjects showed similar improvements in the accuracy of their choices, regardless of how many items they had to learn.

Response times, however, were significantly affected by both Round and List-length (Fig. 5b). A significant main effect of Round on RT ($F[2.50, 77.46] = 28.91$, $p < 0.001$, $\eta_p^2 = 0.48$) indicated a general increase in speed across learning. There was a main effect of List-length ($F[1, 31] = 7.20$, $p = 0.01$, $\eta_p^2 = 0.19$), wherein subjects learning the shorter list (50-items) were on average faster than subjects learning the longer list (100-items). The increases in response speed across Rounds, however, did not depend on List-length, indicated by a null interaction ($F[2.50, 77.46] = 2.35$, $p = 0.09$, $\eta_p^2 = 0.07$). The relationship between these RT differences and the learning curves can be captured and described by the drift-diffusion model analysis.

Mirroring the approach of Experiment 1, we again used a drift-diffusion model analysis to quantify patterns of decision criteria and evidence quality across the two groups and across learning. Seven-parameter diffusion models were built for each of the two List-length groups with threshold (a) and drift-rate (v) parameters free to vary across eight levels of Round. To ensure that the above models were of adequate design, two alternative models, wherein one parameter (decision threshold or drift-rate) was fixed while the other was free to vary across Rounds, produced poorer fits for both the 50-pair (DIC = 10,325 and 9,324, respectively) and 100-pair (DIC = 24,876 and 23,132) lists relative to the models in which both parameters were free for 50-pairs (DIC = 8,927) and 100-pairs (DIC = 22,465). This suggests that the chosen model design (i.e., including both drift-rate and decision threshold as free parameters) was the best among these alternative models.

Average values for the free parameters from the chosen design are enumerated in Table 4. In a 2×8 repeated-measures ANOVA including factors of Round and List-length, drift-rate varied across Round ($F[2.04, 63.27] = 181.84$, $p < 0.001$, $\eta_p^2 = 0.85$), corresponding to a linear increase in the quality of evidence across learning. There was no main effect of List-length ($F[1, 31] = 0.29$, $p = 0.59$, $\eta_p^2 = 0.01$) or an interaction between Round and List-length for drift-rate ($F[2.04, 63.27] = 0.62$, $p = 0.54$, $\eta_p^2 = 0.02$). This suggests that evidence quality is comparable between the two groups and that it increases at similar rates. For decision thresholds, there was also a main effect of Round ($F[3.29, 101.99] = 13.55$, $p < 0.001$, $\eta_p^2 = 0.30$), indicating that thresholds, on average, varied across Rounds, but no main effect of List-length ($F[1, 31] = 1.46$, $p = 0.236$, $\eta_p^2 = 0.04$). However, there was an interaction of the two factors ($F[3.29, 101.99] = 4.24$, $p < 0.01$, $\eta_p^2 = 0.12$), indicating that

the pattern of changes in decision thresholds across Rounds depended on List-length. These patterns replicated the findings of Experiment 1 and are illustrated in Fig. 5.

Neural regions associated with memory differences—The first goal of our neuroimaging analysis was to identify regions associated with differences in memory due to the list-length manipulation. In a voxelwise ANOVA of List-length (50, 100) and trial Time Point (11 levels, one for each modeled time point of a trial), seven regions were localized corresponding to the interaction of these two factors, enumerated in Table 5. Activity in these regions statistically differed by List-length, and interestingly, all seven regions exhibited greater activity for the 50-pair condition versus the 100-pair condition. This set of regions comprised territory in the occipital lobe, inferior frontal gyrus, inferior parietal lobule and precuneus, medial frontal gyrus (pre-supplementary motor area), and the hippocampus (Fig. 6a). Of particular interest, the region in the hippocampus exhibits an activity pattern for the 50-pair condition wherein activity modulates across rounds, but for the 100-pair condition, activity remains at baseline throughout all rounds (Fig. 6b). This directly implicates the MTL in concurrent discrimination learning for the 50-pair condition and suggests that this set of regions, including the MTL, may underlie the behavioral memory differences between list-length groups in this experiment and in Experiment 1.

We performed a secondary analysis to explore regions whose activity reflects list-length changes that differ between early and late stages of learning. In particular, the goal of this analysis was to examine whether declarative memory differences relate to neural differences across learning. This analysis revealed regions in the medial frontal gyrus/pre-supplementary motor area (+06, +06, +56), bilateral body of the caudate nucleus (−19, +03, +26 and +16, +03, +28), and the left fusiform gyrus (−41, −78, −13). These regions are consistent with regions related to learning and decision-making identified in prior work using this task (Tremel et al., 2016). All of these regions exhibited an effect of list-length such that the 50-pair condition was associated with greater activation than the 100-pair condition. For the medial frontal gyrus and fusiform regions, this effect was greater during early learning than later learning. For the bilateral caudate body regions, this effect was greater during later learning.

Neural regions associated with decision-making differences—Our second goal was to identify regions associated with differences in decision-making across learning due to the list-length manipulation. Brain regions of interest were identified by testing for group-level List-length differences of the imaging regressors derived from the drift-diffusion models. In a voxelwise ANOVA of Parameter (decision threshold, drift-rate), List-length (50, 100), and Time (11 levels corresponding to each modeled time point), 15 regions were localized corresponding to the three-way interaction of these factors. This interaction identified regions that were differentially sensitive to one parameter over the other in a manner that depended on the list-length manipulation. Time series data were extracted from each region to determine which drift-diffusion parameter was the driver of the regional effect. In examining the main effects of Parameter, none of these regions tracked with drift-rate. Activity in all 15 regions correlated with modulations in decision thresholds, differentially by list-length. Thus, drift-rate effects were excluded from further analysis.

These regions were separated into two sets, one in which activity for the 50-pair group was predominant over that of the 100-pair group, and another set for the opposite pattern (i.e., direction of the main effect of list-length) (Fig. 7). Table 6 outlines the first set, containing nine regions whose activity exhibited stronger correlations with decision threshold modulations for the 100-pair group. This set included regions in the dorsomedial cerebellum, orbitofrontal cortex, left thalamus, the striatum (including the caudate head and anterior putamen), left cuneus, and the bilateral temporoparietal junction. The second set of regions is outlined in Table 7. This set contains six regions whose activity exhibited stronger correlations with changes in decision thresholds across learning for the 50-pair group versus the 100-pair group. It includes regions encompassing the right fusiform gyrus, right subcallosal area, and medial cortical regions in the precuneus and the anterior and posterior cingulate gyrus.

Discussion, Experiment 2

The behavioral results from Experiment 2 replicate the findings from Experiment 1. In both experiments, we observed that list-length affected recognition accuracy, with more accurate recognition when items are learned in a short list versus a long list. However, list-length did not affect value memory accuracy in either experiment. In Experiment 1, we found that selection had a greater effect on item memory in the 100-pair condition versus the 50-pair condition. From this, one would expect that memory after continuous learning would reflect a preference for items that were selected more often. Indeed, this is the case—recognition was better for positive items, which across learning, were items that subjects selected more often round after round (i.e., the goal of concurrent discrimination is to learn to select the positive item). Thus the early encoding differences from Experiment 1 may influence changes in decision threshold patterns in Round 2, which in turn, may influence which representations are drawn upon during decision-making. The ability to preferentially recognize positive items for the 100-pair group suggests that these subjects rely on representations that encode information about item-level associations to outcomes, rather than event-level representations about an experience.

In addition to the replicated and additional memory findings, the decision-making profiles of the 50- and 100-pair groups in this experiment parallel those observed in Experiment 1, despite the longer learning trajectory. In both experiments, evidence quality was comparable between groups, wherein drift-rate increased linearly across learning at similar rates for the 50- and 100-pair groups. Decision thresholds, however, were increased for the 50-pair group relative to the 100-pair group in the early rounds of the task when most of the learning is taking place (Rounds 2–4). The decision thresholds converge to similar levels in later rounds. This suggests that the linear evidence quality increases and more prudent decision process in Round 2 may translate into more confident or less effortful decision-making in later Rounds.

The behavioral replication is consistent with our predictions that the 50-pair group may rely on an MTL-based strategy, involving strong memory representations that would require more effortful evaluation in Round 2. As decisions are better learned, the required effort of retrieval decreases, which would be captured by the increase and subsequent decrease in

decision thresholds. Indeed, a voxelwise imaging analysis revealed a right hippocampal region in the MTL that was preferentially active for the 50-pair group versus the 100-pair group. Notably, this region's activity modulated with learning (i.e., rounds) for the short list condition but remained at baseline for the long list condition (Fig. 6b). This suggests that the list-length manipulation was successful at differentially engaging declarative memory during concurrent discrimination learning, and moreover that the MTL is relevant to discrimination learning performance for the 50-pair group.

Interestingly, this MTL region was unrelated to changes in decision thresholds across learning. Instead, another set of regions preferentially reflected the decision-making profile of the short list condition versus the long list condition. While these regions (Table 7) are outside of MTL territory, they do seem to be associated with episodic memory retrieval (Fletcher et al., 1995; Nielsen et al., 2005; Leech & Sharp, 2013) and with cognitive control processes (Allman et al., 2001; Leech & Sharp, 2013). More directly, the septal area has been linked to limbic and hippocampal memory systems, particular in reward contexts when outcomes are uncertain (Olds & Milner, 1954; Monosov & Hikosaka, 2012). Thus, while regions in the MTL proper may not directly relate to differences in threshold changes between list-lengths, regions associated with the retrieval and use of MTL-based representations do relate to these differences. This suggests that subjects learning the 50-pair list may indeed rely more heavily on mnemonic information encoded via the MTL versus subjects in the 100-pair group. Taken together, the MTL seems to be involved in mnemonic processes that may affect representations of information, but less directly involved in driving decision-making. It may be the case that the representations encoded and influenced by an MTL system contribute to decision-making in a concurrent discrimination task, meaning that the MTL proper indirectly contributes to decision-making.

In contrast, for the 100-pair group, we hypothesized that subjects might rely more on striatum-based representations (with overall weaker and less effective declarative memory due to the list-length increase), wherein stimulus-outcome associations would drive decision-making. While these representations become more accurate across learning, their retrieval would not require additional evaluation in Round 2 and later (i.e., the amount of information they contain is constant across learning, but their veracity increases). The voxelwise imaging analysis of the list-length manipulation, however, did not reveal any regions that are preferentially active for the 100-pair list. Nonetheless, the drift-diffusion model-informed imaging analysis identified a striatal network that tracked with decision-thresholds moreso for this long list group versus the short list group. Anatomically, this set of regions aligns with those associated with reinforcement learning, procedural memory, and striatum-based decision-making (Daw et al., 2011; Frank & Claus, 2006; Gläscher et al., 2010; Glimcher, 2011; Hare et al., 2008; Schultz, 1998; Schultz, 2013). This includes the orbitofrontal cortex, implicated in stimulus valuation (Schultz et al., 1998; Schultz et al., 2000; Frank & Claus, 2006) as well as the cuneus and temporoparietal junction, implicated in visual attention (Fink et al., 1996). The finding that these regions are related to changes in decision thresholds across learning suggests that the mnemonic information underlying decisions for the long-list group may be related to stimulus-outcome representations acquired via reinforcement learning. This is consistent with our prior work using this task, wherein estimates of reinforcement learning parameters intersected with decision-making

processes defined via drift-diffusion modeling, identifying regions in the putamen and caudate nucleus, like those in the present study (Tremel et al., 2016). Other studies of feedback-based learning have highlighted such a role for the striatum in guiding subsequent choice behavior based on outcomes experienced in the past (Poldrack et al., 2005; Schultz et al., 1998; Schultz et al., 2000; Seger & Cincotta, 2005; Seger et al., 2010; Tricomi & Fiez, 2008; Tricomi & Fiez, 2012; Williams & Eskandar, 2006).

One concern is that these threshold patterns could relate to differences in fatigue across subjects, wherein the 100-pair condition takes roughly double the time to perform compared to the 50-pair condition. There are several reasons why we believe that fatigue does not drastically impact these findings. First, the level of overall learning is identical between groups. If subjects in the 100-pair list were more fatigued, one would expect learning to be less robust. Second, decision thresholds suggest that fatigue is not an issue. In Round 2, early on in the task, both groups raise their thresholds relative to their identical Round 1 levels. Subjects in the 100-pair group maintain this level throughout, whereas it might be expected to drop as fatigue increases (e.g., reflecting less effort, control, motivation, or attention). Moreover, this pattern is replicated in both experiments using different subjects and different settings (scan versus behavioral session). Third, the major effects of interest that drive these findings emerge very early in learning, when subjects have engaged in the task for less than 15 minutes. Fatigue does, however, likely set in later in learning. There is some evidence of this in the wavering patterns of decision thresholds in Rounds 6, 7, and 8. Notably, however, the decision thresholds never drop back to Round 1 levels, suggesting that subjects remain engaged to a reasonable extent in both groups.

Notably, in a secondary analysis, we identified regions in the caudate and medial frontal gyrus that modulated across learning in a way that varied by list-length. These regions are consistent with regions identified in prior work, associated with learning and decision-making (Tremel et al., 2016). In the present context, this suggests that these regions may also relate to mnemonic processing in addition to learning and decision-making. We also identified a left fusiform gyrus region in the vicinity of the visual word form area (Cohen et al., 2000; Cohen et al., 2002), suggesting that changes in learning may directly affect visual representations of the words in the task, especially for the 50-pair condition. While this analysis may suggest that neural differences associated with changes in decision thresholds may not necessarily be comparable to changes due to learning, it is notable that this analysis is not comprehensive in that it examines two epochs of learning that average across major learning differences, especially in early rounds. By collapsing Rounds 1–4, we are effectively averaging across initial encoding in round one as well as the increases and decreases in decision threshold in the second, third, and fourth rounds. Additionally, this analysis focuses on correct trials only, which leaves out differences due to learning from errors. Errors may impact early learning much more than later learning since there are far fewer error trials as subjects over-learn the associations with practice. Thus, while the regions in this analysis are informative, the analysis is not exhaustive in its approach to localize putative learning-related regions.

General Discussion

In this study, we investigated the impact of manipulations of declarative memory on deterministic decision-making and learning. To manipulate declarative memory, we varied the number of item pairs that subjects attempted to learn in the context of a concurrent discrimination learning task. Across two experiments, we observed that manipulations of list-length produced a distinctive behavioral and neural profile of task performance that can be understood from a multiple-memory systems perspective.

List length effects on memory

Experiment 1 confirmed the prediction that declarative memory would be poorer for items experienced in the long-list as compared to short-list condition. These differences in overall recognition performance were not evenly distributed across the items. Rather, subjects in the long-list condition exhibited a larger selection effect, wherein items selected in Round 1 were remembered more accurately, especially when item selection produced a positive outcome. Subjects in the short-list condition, however, were able to remember selected and unselected items equally well, though overall they were most accurate at remembering the items with an experienced or inferred positive value (i.e., items that were experienced in the presence of negative feedback due to selection of the negative item, but inferred to be associated with positive feedback because of that outcome).

The presence of these early memory differences offer a potential explanation of the Experiment 2 results, wherein memory was tested after completing eight rounds of concurrent discrimination learning. At this point, whether a word was positive or negative had no influence on recognition or value memory for the 50-pair group. In contrast, subjects in the 100-pair group preferentially recognized items that were positive versus those that were negative. This is consistent with the idea that the 50-pair group encoded representations of the event (i.e., pair-level), wherein both positive and negative items are encoded to the same extent, versus item-level representations of the 100-pair group that depend in part on selection.

Interestingly, the poorer declarative memories of subjects in the long list versus the short list group were not accompanied by differences in learning trajectory during concurrent discrimination. That is, while the two groups formed qualitatively different declarative memories for the list items, they were able to draw upon the prior decision experiences equally well to support improvements in their selection of the positive items. This finding that recognition memory is particularly impacted by list-length while learning is more resistant is consistent with prior results in the concurrent discrimination literature. For instance, monkeys are capable of performing a concurrent discrimination task successfully despite MTL damage, but the ability to later recognize items is normally compromised (Ridley et al., 1989; Rehbein et al., 2005). In humans, amnesics exhibit similar patterns of behavior, wherein learning may be spared (albeit slower), but with no ability to later recognize items (Bayley et al., 2005). This has been contextualized as a product of intact habit learning via striatal reinforcement circuits. This interpretation is consistent with our results, wherein the subjects with the greatest deficits in recognition memory due to the list-length manipulation (100-pair group for both Experiment 1 and Experiment 2) seemed to be

more likely to encode information on an item-level rather than an event-level, as recognition memory for positive items was enhanced for items that they selected.

Bridging this gap between the monkey and human discrimination learning literatures is important given that there have been many inconsistencies left unexplained (e.g., Bayley et al., 2005 versus Corkin, 2002 or Malamut et al., 1984 versus Zola-Morgan et al., 1994). Altogether, this hints that decision-making and learning are separable processes from memory. While memory retrieval may be a key evidence provision mechanism for decision-making, there are other factors that determine what kinds of information are available and relevant for a decision.

List length effects on decision strategies

Using a drift-diffusion modeling approach, we investigated the decision processes underlying concurrent discrimination learning. When making a decision, information can come from a variety of sources, including multiple learning and memory systems (Dickerson & Delgado, 2015; Doll et al., 2015; Shohamy & Daw, 2015). Different types and sources of information can contribute in combination to a final decision output or response. Because of this, any two decision outputs may appear superficially similar, but may have been reached by different means. The present study offers support for this, wherein the learning curves of two groups of subjects appear similar on the surface, but the behavioral and neural profiles associated with task performance are distinct.

Behaviorally, in both Experiment 1 and Experiment 2, we observed list-length effects on the decision threshold adopted by subjects across learning. Those in the short-list condition elevated their decision thresholds in Round 2, which gradually decreased across learning. In contrast, the long-list group elevated their thresholds in Round 2 to a lower magnitude than the short list group, and this threshold maintained throughout learning. These results are consistent with the idea that the 50-pair condition should be more associated with a declarative memory strategy to support concurrent discrimination performance, while the 100-pair condition places greater weight on a procedural memory strategy, which would require a constant amount of effort. To support this, engagement of the MTL was associated with the 50-pair list but not with the 100-pair list. This idea is further supported by list-effect differences in the fit of the modeled decision threshold parameter to the hemodynamic signal acquired during concurrent discrimination performance, wherein medial cortical regions were associated with the 50-pair list and striatal regions with the 100-pair list.

With respect to the 100-pair group, we found that the threshold regressor identified regions including the dorsal striatum, orbitofrontal cortex, thalamus, cerebellum, and visual attention regions in the temporal and occipital lobes. Collectively, similar regions have been implicated in the encoding, storage, and processing of prediction signals about expected outcome values of particular stimuli (Cohen & Ranganath, 2007; Gläscher et al., 2010; Hare et al., 2008; Tremel et al., 2016). These prediction signals could be sufficient to drive decision-making in the task without major contributions from an MTL-based memory system, though at some cost to subsequent memory of the items. Learning via this approach is thus accomplished not by changing decision criteria (i.e., thresholds), but by updating representations of value and stimulus-outcome associations in areas like the orbitofrontal

cortex and the putamen, via prediction error signals in the dorsal striatum (Camille et al., 2011; FitzGerald et al., 2009; Hare et al., 2008; Padoa-Schioppa & Assad, 2006; Schultz et al., 1997; Schultz, 2013). This kind of approach would be especially sensitive to both selecting an item and to positive feedback in this task, because updates to these representations require one to experience an outcome. Consistent with evidence from human amnesics, a habit learning or reinforcement-centric approach that engages a striatum-based system seems to be a viable mechanism for discrimination (Bayley et al., 2005).

Representations arising from striatal activity may influence discrimination decisions directly and may provide additional decision evidence with or without MTL-based contributions (Bornstein & Daw, 2012; Bornstein & Daw, 2013; Doll et al., 2015; Tremel et al., 2016). Bornstein & Daw (2011) demonstrated that different subsections of the striatum contribute to decision-making in different ways. In particular the dorsomedial striatum acts as a model-based reinforcement learning agent, wherein planning of actions and their dynamic representations drive learning. This aligns with the finding in the present study that the action of selecting an item during discrimination learning can influence recognition of that item for the 100-pair list, whose decision-making profile better corresponded to activity of a striatal system. A striatal system may be more responsive to situations of uncertainty (Bornstein & Daw, 2012), which could potential arise as a result of poorer memory encoding during initial learning, as observed here in Experiment 1. This is further consistent with prior work, which identified differential roles in decision-making for the caudate and the putamen (Tremel et al., 2016). Altogether, it seems that the engagement of different memory systems may differentially rely on structural (e.g., list-length) and contextual factors (e.g., positive versus negative feedback, uncertainty, selection, etc.) of the task at hand.

In contrast to the apparent striatum-mediated decision-making of the 100-pair group, regions such as the posterior and anterior cingulate gyrus, septal nuclei, and the precuneus may be more closely linked to an MTL-based approach. Indeed, we found that the 50-pair list was associated with activity in a hippocampal region in the MTL, whereas the 100-pair list was not. However, the regions linked to decision threshold modulations lay outside the MTL proper. These regions have been implicated in control, recollection, and internal evaluation processes in other studies involving reinforcement and reward (Allman et al., 2001; Fletcher et al., 1995; Leech & Sharp, 2013; Maddock et al., 2001; Nielsen et al., 2005). It is plausible that in terms of decision-making, these regions might be involved in monitoring the internal state of memory quality. Thus, the MTL may not contribute directly to decision threshold changes in concurrent discrimination, but instead may influence later decision processes through other regions, such as those identified in Table 7. Generally, recollection of declarative memories of past experiences requires more effort and control than relying on stimulus-driven responses or less detailed memory like familiarity (Mandler, 1980; Jacoby, 1991; Eichenbaum, 2001). This additional effort ensures that individuals carefully consider their options and make the best decision based on available evidence. This effort could be captured by increases in decision thresholds in drift-diffusion models.

It is important to note that while the regions identified in the diffusion model imaging analysis are more related to threshold changes in one group versus the other, it is likely that both sets of regions are functionally contributing to concurrent discrimination performance

across list-lengths (albeit at different relative strengths). Thus, the findings in this study suggest that subjects leverage different types of mnemonic representations during concurrent discrimination learning depending on list-length. This is consistent with recent findings that different memory representations can be used to drive subsequent decision-making. For instance, in a secondary imaging analysis, we found evidence that visual representations in the fusiform gyrus may change across learning. We also found a hippocampal region that was more active for the 50-pair group compared to the 100-pair group. Mack & Preston (2016) demonstrated that reactivation of representations in occipito-temporal regions is associated with hippocampal-dependent memory, suggesting that subjects in the present task who learning the 50-pair list may be more prone to engaging in MTL-based memory that reactivates visual representations (particularly of words). Furthermore, activity in the MTL due to representational updating during early learning has been tied to activation in prefrontal regions, including bilateral medial frontal regions (Mack et al., 2016). This is consistent with the early differences we see in decision thresholds and that these early differences can be tied to activity in medial prefrontal regions for the 50-pair group. Therefore, while MTL-based mnemonic representations may provide useful evidence to both 50- and 100-pair groups, the relative impact of those representations during decision-making can be weighted by factors such as task structure (e.g., list-length) or episodic context of a trial (e.g., positive versus negative feedback) (Murty et al., 2016a; Murty et al., 2016b). This may partly explain the response bias of subjects in the 50-pair group who had a greater-than-chance propensity to response “negative” when explicitly identifying information about the value associated with an item. Context-dependent encoding (e.g., initially experiencing an item in the context of receiving negative feedback) is more associated with MTL-based processes and can influence decision-making if that context is recalled during a subsequent choice (Bornstein & Norman, 2017; Duncan & Shohamy, 2016).

While these results are consistent with many aspects of the literature, there are still some notably discrepancies with the concurrent discrimination neuropsychological literature. In the current study and our previous work (Tremel et al., 2016), the MTL did not map onto particular decision-making processes, whereas the striatum did. It is clear from the literature and from the results of the list-length analysis, that the MTL (specifically the hippocampus) is important for concurrent discrimination in some way. Both monkeys and humans with damage to this region exhibit performance impairments regularly (Buffalo et al., 1999; Corkin, 2002; Hood et al., 1999; Squire et al., 1988; Squire and Zola, 1996; Zola-Morgan et al., 1989; Zola-Morgan et al., 1994), albeit not always (Malamut et al., 1984; Phillips et al., 1988; Suzuki et al., 1993; Gaffan & Murray, 1992; Buffalo et al., 1998; Hood et al., 1999; Bayley et al., 2005; Chudasama et al., 2008). From this, the nature of how the MTL exactly contributes to concurrent discrimination learning remains an open question. It is possible that, in the case of human amnesics (Bayley et al., 2005), the slow trajectory of improvements via habit learning (presumably through a striatum-based system) is partly a result of engaging in a declarative memory strategy that is unsuccessful. The improvements may still arise via the striatal system, but hampered by the way subjects engage in the task. However, whether longer lists (e.g., 100) can recover a faster learning in amnesics is still unknown and is therefore difficult to speculate based on our data from the 100-item list.

Another possibility is that the MTL is necessary for shorter list-lengths to scaffold early learning that is then “proceduralized” with repeated practice. Without this early scaffold for a declarative memory strategy, a striatal system cannot drive learning to its fullest extent. This interpretation suggests that both systems are essential for adequate performance on these shorter lists, which is supported in the literature (Teng et al., 2000). Ultimately, however, the way in which the MTL or the striatum independently impact concurrent discrimination performance is still unresolved.

Conclusions

In this study, we investigated the relationship between decision-making processes and memory representations by examining changes in decision thresholds, the underlying neural systems, and memory performance during and after learning. Experiment 1 demonstrated that initial memories might be encoded differently depending on the amount of information that must be learned. A list-length manipulation revealed that memory profiles after an initial round of encoding differed, suggesting that subjects may leverage different types of memory to guide subsequent decision-making. These memory and decision effects depended on structural and contextual factors related to the task (e.g., list length, positive versus negative feedback, whether an item was selected, etc.). In Experiment 2, we examined a longer trajectory of learning and found that different neural systems were associated with these memory-related and decision-related effects. Specifically, we found that a striatal system was associated with decision threshold patterns when learning a longer list of items, whereas a medial cortical network was associated with those patterns when learning a shorter list. Furthermore, activity in the hippocampus was related to the list-length manipulation for the short-list but not for the long-list. Collectively, these experiments provide evidence that multiple types of mnemonic representations contribute to experienced-based decision-making. Contextual and structural factors can influence the types of evidence that are drawn upon during decision-making.

Acknowledgments

This work was supported by the National Science Foundation (Grant SBE-0354420 Supplement 0839229 to J.A.F.) and the National Institutes of Health (R21 MH106317).

References

- Allman JM, Hakeem A, Erwin JM, Nimchinsky E, Hof P. The anterior cingulate cortex. *Annals of the New York Academy of Sciences*. 2001; 935(1):107–117. [PubMed: 11411161]
- Balota DA, Yap MJ, Hutchison KA, Cortese MJ, Kessler B, Loftis B, ... Treiman R. The English lexicon project. *Behavior research methods*. 2007; 39(3):445–459. [PubMed: 17958156]
- Bayley PJ, Frascino JC, Squire LR. Robust habit learning in the absence of awareness and independent of the medial temporal lobe. *Nature*. 2005; 436(7050):550–553. [PubMed: 16049487]
- Bornstein AM, Daw ND. Multiplicity of control in the basal ganglia: computational roles of striatal subregions. *Current opinion in neurobiology*. 2011; 21(3):374–380. [PubMed: 21429734]
- Bornstein AM, Daw ND. Dissociating hippocampal and striatal contributions to sequential prediction learning. *European Journal of Neuroscience*. 2012; 35(7):1011–1023. [PubMed: 22487032]
- Bornstein AM, Daw ND. Cortical and hippocampal correlates of deliberation during model-based decisions for rewards in humans. *PLoS computational biology*. 2013; 9(12):e1003387. [PubMed: 24339770]

- Bornstein AM, Norman KA. Reinstated episodic context guides sampling-based decisions for reward. *Nature neuroscience*. 2017; 20(7):997. [PubMed: 28581478]
- Buffalo EA, Stefanacci L, Squire LR, Zola SM. A reexamination of the concurrent discrimination learning task: the importance of anterior inferotemporal cortex, area TE. *Behavioral neuroscience*. 1998; 112(1):3. [PubMed: 9517811]
- Buffalo EA, Ramus SJ, Clark RE, Teng E, Squire LR, Zola SM. Dissociation between the effects of damage to perirhinal cortex and area TE. *Learning & Memory*. 1999; 6(6):572–599. [PubMed: 10641763]
- Burnham KP, Anderson DR. Multimodel inference: understanding AIC and BIC in model selection. *Sociological methods & research*. 2004; 33(2):261–304.
- Camille N, Tsuchida A, Fellows LK. Double dissociation of stimulus-value and action-value learning in humans with orbitofrontal or anterior cingulate cortex damage. *Journal of Neuroscience*. 2011; 31(42):15048–15052. [PubMed: 22016538]
- Chudasama Y, Wright KS, Murray EA. Hippocampal lesions in rhesus monkeys disrupt emotional responses but not reinforcer devaluation effects. *Biological psychiatry*. 2008; 63(11):1084–1091. [PubMed: 18191111]
- Cohen L, Dehaene S, Naccache L, Lehéricy S, Dehaene-Lambertz G, Hénaff MA, Michel F. The visual word form area: spatial and temporal characterization of an initial stage of reading in normal subjects and posterior split-brain patients. *Brain*. 2000; 123(2):291–307. [PubMed: 10648437]
- Cohen L, Lehéricy S, Chochon F, Lemer C, Rivaud S, Dehaene S. Language specific tuning of visual cortex? Functional properties of the Visual Word Form Area. *Brain*. 2002; 125(5):1054–1069. [PubMed: 11960895]
- Cohen MX, Ranganath C. Reinforcement learning signals predict future decisions. *Journal of Neuroscience*. 2007; 27(2):371–378. [PubMed: 17215398]
- Coltheart M. The MRC psycholinguistic database. *The Quarterly Journal of Experimental Psychology*. 1981; 33(4):497–505.
- Corkin S. What's new with the amnesic patient HM? *Nature Reviews Neuroscience*. 2002; 3(2):153–160. [PubMed: 11836523]
- Craik FIM, Tulving E. Depth of processing and the retention of words in episodic memory. *Journal of Experimental Psychology: General*. 1975; 104(3):268–294.
- Dale AM. Optimal experimental design for event-related fMRI. *Human brain mapping*. 1999; 8(2–3):109–114. [PubMed: 10524601]
- Daw ND, Gershman SJ, Seymour B, Dayan P, Dolan RJ. Model-based influences on humans' choices and striatal prediction errors. *Neuron*. 2011; 69(6):1204–1215. [PubMed: 21435563]
- Dickerson KC, Delgado MR. Contributions of the hippocampus to feedback learning. *Cognitive, Affective, & Behavioral Neuroscience*. 2015; 15(4):861–877.
- Doll BB, Shohamy D, Daw ND. Multiple memory systems as substrates for multiple decision systems. *Neurobiology of learning and memory*. 2015; 117:4–13. [PubMed: 24846190]
- Duncan KD, Shohamy D. Memory states influence value-based decisions. *Journal of Experimental Psychology: General*. 2016; 145(11):1420. [PubMed: 27797556]
- Eichenbaum H. The hippocampus and declarative memory: cognitive mechanisms and neural codes. *Behavioural brain research*. 2001; 127(1–2):199–207. [PubMed: 11718892]
- Fernandez-Ruiz J, Wang J, Aigner TG, Mishkin M. Visual habit formation in monkeys with neurotoxic lesions of the ventrocaudal neostriatum. *Proceedings of the National Academy of Sciences*. 2001; 98(7):4196–4201.
- Fink GR, Markowitsch HJ, Reinkemeier M, Bruckbauer T, Kessler J, Heiss WD. Cerebral representation of one's own past: neural networks involved in autobiographical memory. *Journal of neuroscience*. 1996; 16(13):4275–4282. [PubMed: 8753888]
- FitzGerald TH, Seymour B, Dolan RJ. The role of human orbitofrontal cortex in value comparison for incommensurable objects. *Journal of Neuroscience*. 2009; 29(26):8388–8395. [PubMed: 19571129]
- Fletcher PC, Frith CD, Baker SC, Shallice T, Frackowiak RSJ, Dolan RJ. The mind's eye—precuneus activation in memory-related imagery. *Neuroimage*. 1995; 2(3):195–200. [PubMed: 9343602]

- Fox, J., Weisberg, S. An {R} Companion to Applied Regression. 2. Thousand Oaks CA: Sage; 2011.
URL: <http://socserv.socsci.mcmaster.ca/jfox/Books/Companion>
- Frank MJ, Claus ED. Anatomy of a decision: striato-orbitofrontal interactions in reinforcement learning, decision making, and reversal. *Psychological review*. 2006; 113(2):300. [PubMed: 16637763]
- Friston KJ, Tononi G, Reeke GN, Sporns O, Edelman GM. Value-dependent selection in the brain: simulation in a synthetic neural model. *Neuroscience*. 1994; 59(2):229–243. [PubMed: 8008189]
- Gaffan D, Murray EA. Monkeys (*Macaca fascicularis*) with rhinal cortex ablations succeed in object discrimination learning despite 24-hr intertrial intervals and fail at matching to sample despite double sample presentations. *Behavioral neuroscience*. 1992; 106(1):30. [PubMed: 1554436]
- Gläscher J, Daw N, Dayan P, O'Doherty JP. States versus rewards: dissociable neural prediction error signals underlying model-based and model-free reinforcement learning. *Neuron*. 2010; 66(4):585–595. [PubMed: 20510862]
- Glimcher PW. Understanding dopamine and reinforcement learning: the dopamine reward prediction error hypothesis. *Proceedings of the National Academy of Sciences*. 2011; 108(Supplement 3): 15647–15654.
- Glover GH. Deconvolution of impulse response in event-related BOLD fMRI. *Neuroimage*. 1999; 9(4):416–429. [PubMed: 10191170]
- Grasby P, Frith CD, Friston KJ, Simpson JFPC, Fletcher PC, Frackowiak RS, Dolan RJ. A graded task approach to the functional mapping of brain areas implicated in auditory—verbal memory. *Brain*. 1994; 117(6):1271–1282. [PubMed: 7820565]
- Hare TA, O'Doherty J, Camerer CF, Schultz W, Rangel A. Dissociating the role of the orbitofrontal cortex and the striatum in the computation of goal values and prediction errors. *Journal of neuroscience*. 2008; 28(22):5623–5630. [PubMed: 18509023]
- Hood KL, Postle BR, Corkin S. An evaluation of the concurrent discrimination task as a measure of habit learning: performance of amnesic subjects. *Neuropsychologia*. 1999; 37(12):1375–1386. [PubMed: 10606012]
- Jacoby LL. A process dissociation framework: Separating automatic from intentional uses of memory. *Journal of memory and language*. 1991; 30(5):513–541.
- Leech R, Sharp DJ. The role of the posterior cingulate cortex in cognition and disease. *Brain*. 2013; 137(1):12–32. [PubMed: 23869106]
- Mack ML, Preston AR. Decisions about the past are guided by reinstatement of specific memories in the hippocampus and perirhinal cortex. *NeuroImage*. 2016; 127:144–157. [PubMed: 26702775]
- Mack ML, Love BC, Preston AR. Dynamic updating of hippocampal object representations reflects new conceptual knowledge. *Proceedings of the National Academy of Sciences*. 2016; 113(46): 13203–13208.
- Maddock RJ, Garrett AS, Buonocore MH. Remembering familiar people: the posterior cingulate cortex and autobiographical memory retrieval. *Neuroscience*. 2001; 104(3):667–676. [PubMed: 11440800]
- Mahut H, Zola-Morgan STUART, Moss M. Hippocampal resections impair associative learning and recognition memory in the monkey. *Journal of Neuroscience*. 1982; 2(9):1214–1220. [PubMed: 7119874]
- Malamut BL, Saunders RC, Mishkin M. Monkeys with combined amygdalo-hippocampal lesions succeed in object discrimination learning despite 24-hour intertrial intervals. *Behavioral neuroscience*. 1984; 98(5):759. [PubMed: 6487412]
- Mandler G. Recognizing: The judgment of previous occurrence. *Psychological review*. 1980; 87(3): 252.
- Miezin FM, Maccotta L, Ollinger JM, Petersen SE, Buckner RL. Characterizing the hemodynamic response: effects of presentation rate, sampling procedure, and the possibility of ordering brain activity based on relative timing. *Neuroimage*. 2000; 11(6):735–759. [PubMed: 10860799]
- Mishkin M. A memory system in the monkey. *Philosophical transactions of the royal society of london B: Biological Sciences*. 1982; 298(1089):85–95.

- Monosov IE, Hikosaka O. Regionally distinct processing of rewards and punishments by the primate ventromedial prefrontal cortex. *Journal of Neuroscience*. 2012; 32(30):10318–10330. [PubMed: 22836265]
- Moody TD, Chang GY, Vanek ZF, Knowlton BJ. Concurrent discrimination learning in Parkinson's disease. *Behavioral neuroscience*. 2010; 124(1):1. [PubMed: 20141275]
- Morcom AM, Rugg MD. Getting ready to remember: the neural correlates of task set during recognition memory. *NeuroReport*. 2002; 13(1):149–152. [PubMed: 11924877]
- Murty VP, FeldmanHall O, Hunter LE, Phelps EA, Davachi L. Episodic memories predict adaptive value-based decision-making. *Journal of Experimental Psychology: General*. 2016; 145(5):548. [PubMed: 26999046]
- Murty VP, LaBar KS, Adcock RA. Distinct medial temporal networks encode surprise during motivation by reward versus punishment. *Neurobiology of learning and memory*. 2016; 134:55–64. [PubMed: 26854903]
- Nielsen FÅ, Balslev D, Hansen LK. Mining the posterior cingulate: segregation between memory and pain components. *Neuroimage*. 2005; 27(3):520–532. [PubMed: 15946864]
- Ojemann JG, Akbudak E, Snyder AZ, McKinsty RC, Raichle ME, Conturo TE. Anatomic localization and quantitative analysis of gradient refocused echo-planar fMRI susceptibility artifacts. *Neuroimage*. 1997; 6(3):156–167. [PubMed: 9344820]
- Olds J, Milner P. Positive reinforcement produced by electrical stimulation of septal area and other regions of rat brain. *Journal of comparative and physiological psychology*. 1954; 47(6):419. [PubMed: 13233369]
- Ollinger JM, Shulman GL, Corbetta M. Separating processes within a trial in event-related functional MRI: I. The method. *Neuroimage*. 2001; 13(1):210–217. [PubMed: 11133323]
- Padoa-Schioppa C, Assad JA. Neurons in orbitofrontal cortex encode economic value. *Nature*. 2006; 441(7090):223. [PubMed: 16633341]
- Peirce JW. PsychoPy—psychophysics software in Python. *Journal of neuroscience methods*. 2007; 162(1):8–13. [PubMed: 17254636]
- Peirce JW. Generating stimuli for neuroscience using PsychoPy. *Frontiers in neuroinformatics*. 2009; 2:10. [PubMed: 19198666]
- Phillips RR, Malamut BL, Bachevalier J, Mishkin M. Dissociation of the effects of inferior temporal and limbic lesions on object discrimination learning with 24-h intertrial intervals. *Behavioural Brain Research*. 1988; 27(2):99–107. [PubMed: 3358857]
- Poldrack RA, Sabb FW, Foerde K, Tom SM, Asarnow RF, Bookheimer SY, Knowlton BJ. The neural correlates of motor skill automaticity. *Journal of Neuroscience*. 2005; 25(22):5356–5364. [PubMed: 15930384]
- R Development Core Team. R: A language and environment for statistical computing. R Foundation for Statistical Computing; Vienna, Austria: 2016. URL: <http://www.R-project.org>
- Ratcliff R. A theory of memory retrieval. *Psychological review*. 1978; 85(2):59.
- Ratcliff R, McKoon G. The diffusion decision model: theory and data for two-choice decision tasks. *Neural computation*. 2008; 20(4):873–922. [PubMed: 18085991]
- Ratcliff R, Smith PL, Brown SD, McKoon G. Diffusion decision model: current issues and history. *Trends in cognitive sciences*. 2016; 20(4):260–281. [PubMed: 26952739]
- Rehbein L, Killiany R, Mahut H. Developmental study of the hippocampal formation in rhesus monkeys (*Macaca mulatta*): I. Early ablations spare discrimination learning but not recognition memory. *Behavioral neuroscience*. 2005; 119(3):635. [PubMed: 15998185]
- Ridley RM, Aitken DM, Baker HF. Learning about rules but not about reward is impaired following lesions of the cholinergic projection to the hippocampus. *Brain research*. 1989; 502(2):306–318. [PubMed: 2510907]
- Schultz W, Dayan P, Montague PR. A neural substrate of prediction and reward. *Science*. 1997; 275(5306):1593–1599. [PubMed: 9054347]
- Schultz W, Tremblay L, Hollerman JR. Reward prediction in primate basal ganglia and frontal cortex. *Neuropharmacology*. 1998; 37(4):421–429. [PubMed: 9704983]

- Schultz W, Tremblay L, Hollerman JR. Reward processing in primate orbitofrontal cortex and basal ganglia. *Cerebral cortex*. 2000; 10(3):272–283. [PubMed: 10731222]
- Schultz W. Updating dopamine reward signals. *Current opinion in neurobiology*. 2013; 23(2):229–238. [PubMed: 23267662]
- Seger CA, Cincotta CM. The roles of the caudate nucleus in human classification learning. *Journal of Neuroscience*. 2005; 25(11):2941–2951. [PubMed: 15772354]
- Seger CA, Peterson EJ, Cincotta CM, Lopez-Paniagua D, Anderson CW. Dissociating the contributions of independent corticostriatal systems to visual categorization learning through the use of reinforcement learning modeling and Granger causality modeling. *Neuroimage*. 2010; 50(2):644–656. [PubMed: 19969091]
- Shadlen MN, Shohamy D. Decision making and sequential sampling from memory. *Neuron*. 2016; 90(5):927–939. [PubMed: 27253447]
- Shohamy D, Daw ND. Integrating memories to guide decisions. *Current Opinion in Behavioral Sciences*. 2015; 5:85–90.
- Snyder, AZ. Quantification of brain function using PET. 1996. Difference image versus ratio image error function forms in PET-PET realignment; p. 131-137.
- Spiegelhalter DJ, Best NG, Carlin BP, Van Der Linde A. Bayesian measures of model complexity and fit. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*. 2002; 64(4):583–639.
- Spiegelhalter DJ, Best NG, Carlin BP, Linde A. The deviance information criterion: 12 years on. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*. 2014; 76(3):485–493.
- Squire LR, Zola-Morgan S, Chen KS. Human amnesia and animal models of amnesia: performance of amnesic patients on tests designed for the monkey. *Behavioral neuroscience*. 1988; 102(2):210. [PubMed: 3130073]
- Squire LR, Zola SM. Structure and function of declarative and nondeclarative memory systems. *Proceedings of the National Academy of Sciences*. 1996; 93(24):13515–13522.
- Suzuki WA, Zola-Morgan S, Squire LR, Amaral DG. Lesions of the perirhinal and parahippocampal cortices in the monkey produce long-lasting memory impairment in the visual and tactual modalities. *Journal of Neuroscience*. 1993; 13(6):2430–2451. [PubMed: 8501516]
- Talairach J, Tournoux P. Co-planar stereotaxic atlas of the human brain. 3-Dimensional proportional system: an approach to cerebral imaging. 1988
- Teng E, Stefanacci L, Squire LR, Zola SM. Contrasting effects on discrimination learning after hippocampal lesions and conjoint hippocampal–caudate lesions in monkeys. *Journal of Neuroscience*. 2000; 20(10):3853–3863. [PubMed: 10804225]
- Tremel JJ, Laurent PA, Wolk DA, Wheeler ME, Fiez JA. Neural signatures of experience-based improvements in deterministic decision-making. *Behavioural brain research*. 2016; 315:51–65. [PubMed: 27523644]
- Tricomi E, Fiez JA. Feedback signals in the caudate reflect goal achievement on a declarative memory task. *Neuroimage*. 2008; 41(3):1154–1167. [PubMed: 18445531]
- Tricomi E, Fiez JA. Information content and reward processing in the human striatum during performance of a declarative memory task. *Cognitive, Affective, & Behavioral Neuroscience*. 2012; 12(2):361–372.
- Turchi J, Devan B, Yin P, Sigrist E, Mishkin M. Pharmacological evidence that both cognitive memory and habit formation contribute to within-session learning of concurrent visual discriminations. *Neuropsychologia*. 2010; 48(8):2245–2250. [PubMed: 20144631]
- Wais PE, Wixted JT, Hopkins RO, Squire LR. The hippocampus supports both the recollection and the familiarity components of recognition memory. *Neuron*. 2006; 49(3):459–466. [PubMed: 16446148]
- Wiecki TV, Frank MJ. A computational model of inhibitory control in frontal cortex and basal ganglia. *Psychological review*. 2013; 120(2):329. [PubMed: 23586447]
- Williams ZM, Eskandar EN. Selective enhancement of associative learning by microstimulation of the anterior caudate. *Nature neuroscience*. 2006; 9(4):562. [PubMed: 16501567]

- Yonelinas AP, Jacoby LL. Dissociations of processes in recognition memory: effects of interference and of response speed. *Canadian Journal of Experimental Psychology/Revue canadienne de psychologie expérimentale*. 1994; 48(4):516. [PubMed: 7866392]
- Yonelinas AP. Components of episodic memory: the contribution of recollection and familiarity. *Philosophical Transactions of the Royal Society of London B: Biological Sciences*. 2001; 356(1413):1363–1374. [PubMed: 11571028]
- Zhang J, Rowe JB. Dissociable mechanisms of speed-accuracy tradeoff during visual perceptual learning are revealed by a hierarchical drift-diffusion model. *Frontiers in neuroscience*. 2014; 8
- Zola-Morgan S, Squire LR, Amaral DG, Suzuki WA. Lesions of perirhinal and parahippocampal cortex that spare the amygdala and hippocampal formation produce severe memory impairment. *Journal of Neuroscience*. 1989; 9(12):4355–4370. [PubMed: 2593004]
- Zola-Morgan S, Squire LR, Ramus SJ. Severity of memory impairment in monkeys as a function of locus and extent of damage within the medial temporal lobe memory system. *Hippocampus*. 1994; 4(4):483–495. [PubMed: 7874239]

Highlights

1. Hippocampal activity is sensitive to a list-length manipulation
2. Declarative memory profile is associated with higher decision thresholds
3. Medial cortical regions are associated with higher decisions thresholds
4. Procedural memory profile is associated with lower decision thresholds
5. Striatal network is associated with lower decision thresholds

Figure 1.

Experiment 1, surprise memory test after initial encoding. a) Corrected hit rate (hits minus false alarms) for recognition memory probe by learning outcome. Positive refers to items that associated with positive feedback (i.e., correct items), whereas negative refers to errors. b) Overall accuracy to the item value memory probe by learning outcome. c) Response bias to the item value memory probe. The proportion of total responses wherein the subject responded that an item was “positive” is plotted on the y-axis. A value of 0.5 (50%) indicates no bias, whereas above or below indicates a bias to respond “positive” or “negative,” respectively. Grey bars represent the 50-pair group, while white bars represent the 100-pair group. Error bars reflect standard error of the mean.

Figure 2.

Experiment 1, recognition memory as a function of whether the item was selected or not in Round 1. Mean corrected hit rate (hit minus false alarm rates) is plotted for positive and negative words, for the 50-pair (blue) and 100-pair (red) list-length groups, and according to whether the item was selected in Round 1 or not. The right panel illustrates the same data plotted instead as a difference score between the hit rate of selected item minus the hit rate of unselected items. Error bars reflect standard error of the mean.

Figure 3.

Experiment 1, concurrent discrimination learning task behavior and drift-diffusion model parameters. a) Mean group accuracy (proportion correct) across repetitions (Rounds) of the task item set. b) Mean response times (in seconds) across Rounds. c) mean decision threshold estimates across Rounds, estimated by the drift-diffusion model. d) mean drift-rate estimates across Rounds. N.B., learning was interrupted between rounds 1 and 2 by a surprise memory test. Blue represents the 50-pair group, while red represents the 100-pair group. Shaded area reflects standard error of the mean.

Figure 4.

Experiment 2, surprise memory test following extended learning task. a) Corrected hit rate (hits minus false alarms) for recognition memory probe by learning outcome. Positive refers to items that associated with positive feedback (i.e., correct items), whereas negative refers to errors. b) Overall accuracy to the feedback memory probe by learning outcome. Grey bars represent the 50-pair group, while white bars represent the 100-pair group. Error bars reflect standard error of the mean.

Figure 5.

Experiment 2, learning trajectory over eight rounds of a concurrent discrimination learning task and parameter estimates from a drift-diffusion model. a) Mean group accuracy (proportion correct) across repetitions (Rounds) of the task item set. b) Mean response times (in seconds) across Rounds. c) mean decision threshold estimates across Rounds, estimated from a drift-diffusion model. d) mean drift-rate estimates across Rounds. Blue represents the 50-pair group, while red represents the 100-pair group. Shaded area reflects standard error of the mean.

Figure 6.

Experiment 2, regions of interest differentially engaged by list-length. a) Activations indicate voxels with a list-length x time point (trial-level activity modulations) interaction. The blue circle highlights the right hippocampal region of interest. b) Time series of activity in the hippocampus for the 50-pair (blue) versus 100-pair (red) conditions. Each line corresponds to the evolution of average activity across a trial for correct responses and for each Round. Darker shades represent earlier rounds, while lighter shades represent later rounds.

Figure 7.

Experiment 2, regions of interest associated with decision threshold changes, depending on list-length. Activations indicate voxels that reflect changes in decision threshold parameters across learning differentially for the 50-pair list (blue to green) and the 100-pair list (red to yellow).

Table 1

Experiment 1, recognition memory and item value memory results after initial encoding experience. Corrected hit rate (Corr. Hit) is the raw hit proportion minus the false alarm rate, adjusting for response bias. Standard error of the mean is reported in parentheses.

List	Stim	Hit	FA	Corr. Hit	Value Mem
50	All	0.56 (0.03)	0.20 (0.02)	0.37 (0.03)	0.62 (0.02)
	Pos	0.60 (0.04)	0.20 (0.02)	0.40 (0.03)	0.65 (0.02)
	Neg	0.53 (0.03)	0.20 (0.02)	0.33 (0.03)	0.58 (0.03)
100	All	0.52 (0.03)	0.24 (0.04)	0.27 (0.03)	0.56 (0.03)
	Pos	0.56 (0.03)	0.24 (0.04)	0.31 (0.04)	0.59 (0.03)
	Neg	0.48 (0.03)	0.24 (0.04)	0.24 (0.03)	0.52 (0.04)

Table 2

Experiment 1, values of free parameters across Rounds for the 50-pair and 100-pair list groups. R1-R3, value for each Round; a , decision threshold; v , drift-rate. Standard error of the mean is reported in parentheses.

Param	R1	R2	R3
a , 50	0.94 (0.02)	1.49 (0.02)	1.54 (0.02)
a , 100	1.01 (0.02)	1.41 (0.02)	1.42 (0.02)
v , 50	-0.06 (0.02)	0.21 (0.04)	0.70 (0.04)
v , 100	-0.02 (0.03)	0.24 (0.03)	0.62 (0.06)

Experiment 2, post-learning recognition memory and item value memory results. Corrected hit rate (Corr. Hit) is the raw hit proportion minus the false alarm rate, adjusting for response bias. Standard error of the mean is reported in parentheses.

Table 3

List	Stim	Hit	FA	Corr. Hit	Value Mem
50	All	0.93 (0.02)	0.07 (0.02)	0.86 (0.03)	0.86 (0.05)
	Pos	0.90 (0.02)	0.07 (0.02)	0.83 (0.03)	0.90 (0.04)
	Neg	0.96 (0.02)	0.07 (0.02)	0.89 (0.03)	0.81 (0.08)
100	All	0.87 (0.03)	0.08 (0.03)	0.79 (0.03)	0.89 (0.03)
	Pos	0.94 (0.02)	0.08 (0.03)	0.86 (0.03)	0.92 (0.02)
	Neg	0.80 (0.04)	0.08 (0.03)	0.72 (0.03)	0.86 (0.06)

Table 4

Experiment 2, values of free parameters across Rounds for the 50-pair and 100-pair list groups. R1-R8, value for each Round; a , decision threshold; v , drift-rate. Standard error of the mean is reported in parentheses.

Param	R1	R2	R3	R4	R5	R6	R7	R8
a , 50	1.35 (0.02)	1.92 (0.02)	1.81 (0.02)	1.74 (0.02)	1.68 (0.02)	1.66 (0.02)	1.65 (0.02)	1.53 (0.02)
a , 100	1.26 (0.02)	1.59 (0.02)	1.65 (0.02)	1.66 (0.02)	1.66 (0.02)	1.66 (0.02)	1.57 (0.02)	1.65 (0.02)
v , 50	-0.14 (0.06)	0.37 (0.05)	0.87 (0.06)	1.52 (0.06)	1.87 (0.06)	2.28 (0.06)	2.65 (0.06)	2.77 (0.07)
v , 100	-0.02 (0.06)	0.28 (0.05)	0.74 (0.06)	1.40 (0.06)	1.83 (0.06)	2.21 (0.06)	2.48 (0.06)	2.83 (0.07)

Experiment 2, regions of interest related to the declarative memory manipulation of List-length. These regions were defined by a voxelwise interaction between List-length (50, 100 items) and trial-level Time Point (11 levels). *x,y,z*, Talairach atlas coordinates of regional peak; BA, approximate Brodmann area; vx, region volume in voxels; R, right; L, left; Inf, inferior; G, gyrus.

Table 5

ROI	anatomical location	x	y	z	BA	vx
1	R Inf Occipital G	27	-85	-8	18	125
2	Medial Frontal G	-3	28	46	8	102
3	R Lingual G	21	-68	6	18	138
4	L Inf Parietal Lobule	-56	-34	39	40	109
5	L Inf Frontal G	-49	38	8	46	38
6	R Precuneus	9	-71	49	7	39
7	R Hippocampus	24	-44	-5	--	45

Table 6

Experiment 2, regions of interest whose activity correlates with modulations in decision thresholds to a greater extent for the 100-pair list than the 50-pair list. *x,y,z*, Talairach atlas coordinates of regional center of mass; BA, approximate Brodmann area; vx, region volume in voxels; R, right; L, left; Cblm, cerebellum; C, cortex; Ant, anterior; J, junction.

ROI	anatomical location	x	y	z	BA	vx
1	L Dorsomedial Cblm	-1	-49	-24	--	54
2	L Orbitofrontal C	-1	51	-10	10	86
3	L Ant Putamen	-13	13	-2	--	46
4	L Cuneus	-25	-89	4	18	60
5	L Thalamus	-11	-23	14	--	45
6	L Caudate Head	-17	15	12	--	64
7	R Temporoparietal J	45	-69	28	39	47
8	R mTG/sTG/PTO	31	-53	30	39	45
9	L Angular G/PTO	-27	-69	30	39	59

Table 7

Experiment 2, regions of interest whose activity correlates with modulations in decision thresholds to a greater extent for the 50-pair list than the 100-pair list. *x,y,z*, Talairach atlas coordinates of regional center of mass; BA, approximate Brodmann area; vx, region volume in voxels; R, right; L, left; G, gyrus.

ROI	anatomical location	<i>x</i>	<i>y</i>	<i>z</i>	BA	vx
1	R Fusiform G	37	-79	-14	19	48
2	R Subcallosal G	17	19	-10	25	54
3	R Cingulate	1	45	10	32	60
4	L Cingulate	-1	-47	36	31	103
5	R Precuneus	7	-55	36	31	123
6	L Precuneus	-7	-71	38	7	48