More Than the Sum of Its Parts: A Role for the Hippocampus in Configural Reinforcement Learning

Highlights

- Reinforcement learning models can disentangle configural from elemental learning

- People engage in configural learning, even when elemental learning is more efficient

- BOLD activity in the hippocampus tracks configurally learned values during choice

- Configural learning is related to interactions between the hippocampus and striatum

Authors

Katherine Duncan, Bradley B. Doll, Nathaniel D. Daw, Daphna Shohamy

Correspondence

duncan@psych.utoronto.ca

In Brief

Duncan et al. investigate how people learn to predict outcomes using cue configurations. They show that configural learning is characterized by unique computational, behavioral, and neural signatures, including hippocampal activity, interactions between the hippocampus and striatum, and enhanced flexible knowledge.

CellPress

Article

More Than the Sum of Its Parts: A Role for the Hippocampus in Configural Reinforcement Learning

Katherine Duncan,[1,5,*] Bradley B. Doll,[3] Nathaniel D. Daw,[2] and Daphna Shohamy[3,4]
[1]Department of Psychology, University of Toronto, Toronto, ON, Canada
[2]Neuroscience Institute and Department of Psychology, Princeton University, Princeton, NJ, USA
[3]Department of Psychology, Columbia University, New York, NY, USA
[4]Zuckerman Mind, Brain, Behavior Institute and Kavli Center for Neuroscience, Columbia University, New York, NY, USA
[5]Lead Contact
*Correspondence: duncan@psych.utoronto.ca
https://doi.org/10.1016/j.neuron.2018.03.042

SUMMARY

People often perceive configurations rather than the elements they comprise, a bias that may emerge because configurations often predict outcomes. But how does the brain learn to associate configurations with outcomes and how does this learning differ from learning about individual elements? We combined behavior, reinforcement learning models, and functional imaging to understand how people learn to associate configurations of cues with outcomes. We found that configural learning depended on the relative predictive strength of elements versus configurations and was related to both the strength of BOLD activity and patterns of BOLD activity in the hippocampus. Configural learning was further related to functional connectivity between the hippocampus and nucleus accumbens. Moreover, configural learning was associated with flexible knowledge about associations and differential eye movements during choice. Together, this suggests that configural learning is associated with a distinct computational, cognitive, and neural profile that is well suited to support flexible and adaptive behavior.

INTRODUCTION

The brain simultaneously represents the world at multiple levels of resolution, from low-level features (such as lines) to their high-level configurations (such as a painting), yet people tend to consciously perceive configurations rather than the features they comprise (Kimchi, 1994; Navon, 1977). This tendency is typically explained in terms of perceptual constraints, such as eccentricity and size (Lamb and Robertson, 1990; Navon and Norman, 1983). Yet there is another possible explanation for the adaptive origin of this phenomenon: a tendency to represent configurations could emerge because configurations are often better predictors of outcomes than the cues of which they are

made. Currently, very little is known about how the brain uses configurations to predict outcomes and how this form of learning adaptively changes between circumstances and across individuals.

We investigated which neural processes are responsible for learning to predict outcomes based on configurations of cues, how this learning changes based on the predictive strength of the configuration, and which individuals are more inclined to learn about configurations. Our central hypothesis was that learning about the predictive strength of configurations involves interactions between stimulus-stimulus associative mechanisms in the hippocampus and outcome learning processes in the ventral striatum. To test this, we combined behavioral measurements of learning (accuracy, reaction time, and eye tracking) with computational models of reinforcement learning and fMRI.

Reinforcement learning provides a formal framework for understanding trial-and-error learning. This approach has most often been applied to learning about the values of options, each represented by a discrete stimulus, such as a single shape. However, the stimuli in our natural environments rarely consist of a single element, but instead comprise configurations of elements. We leveraged reinforcement learning models that have been widely used to characterize learning about a single cue (e.g., Dayan and Daw, 2008; Glimcher, 2011; O'Doherty et al., 2003; Schultz, 1998; Sutton and Barto, 1998), and we extended them to understand how people learn to predict outcomes based on configurations of cues.

There are two qualitatively different ways through which the values of cue configurations can be learned. One could compute a configuration's value by combining the values associated with each cue (Rescorla and Wagner, 1972). This strategy can be thought of as *elemental*, because it does not treat the configuration as having its own predictive strength but instead grounds the value of the configuration in the value of the elements. Alternatively, one could treat each unique configuration as a separate unit with its own learned value. This strategy is consistent with the observation that people and animals tend to perceive configurations rather than cues (Kimchi, 1994; Navon, 1977) and that combinations of cues are treated differently from the sum of their elements (Pearce, 1987, 1994; Rudy and Sutherland, 1995). Despite empirical evidence that humans can learn about both
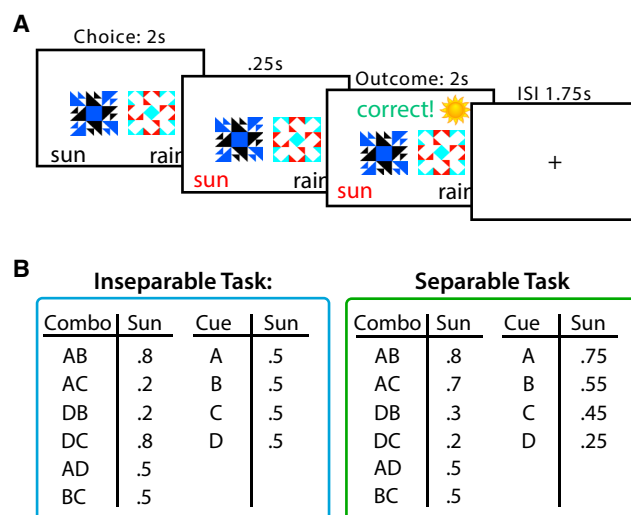
CellPress

A

B

Inseparable Task:					Separable Task			
Combo	Sun	Cue	Sun		Combo	Sun	Cue	Sun
AB	.8	A	.5		AB	.8	A	.75
AC	.2	B	.5		AC	.7	B	.55
DB	.2	C	.5		DB	.3	C	.45
DC	.8	D	.5		DC	.2	D	.25
AD	.5				AD	.5		
BC	.5				BC	.5		

Figure 1. Task Design and Performance
(A) Trial schematic. Both tasks had the same trial structure: participants first saw a pair of cues that they could use to predict weather outcomes. After making a prediction, they were presented with an outcome that allowed them to learn contingencies through trial and error.
(B) Task structures. In the inseparable task, individual items were not predictive of the weather, instead only combinations could be used to predict the outcomes. In the separable task, individual cues and their combinations both were predictive of weather outcomes.

configurations and elements (Melchers et al., 2008) as well as computational frameworks that emphasize the use of structure in learning (e.g., Gershman et al., 2010), little is known about how the brain supports reinforcement learning when the cues are configural. This is a critical question because adopting a configural strategy may require additional neural machinery to differentiate overlapping configurations.

We hypothesized that the hippocampus is well positioned to build and disambiguate configurations in the service of reinforcement learning. The hippocampus is known for its role in episodic and relational memory, both of which depend on the formation of associative links among sets of stimuli (Cohen and Eichenbaum, 1993; Davachi, 2006; Eichenbaum et al., 2007; Eichenbaum and Cohen, 2014). The role of the hippocampus in associative binding is consistent with its position at the apex of the visual processing hierarchy (Amaral and Lavenex, 2006; Van Essen et al., 1992). Moreover, the need to separate overlapping conditions for reinforcement learning parallels a similar requirement for interference reduction in hippocampal-dependent memory, namely, pattern separation between memories with overlapping elements (Marr, 1971; Norman and O'Reilly, 2003; Treves and Rolls, 1992). Indeed, in lesion studies of conditioning in rodents, the hippocampus has been implicated in contextual conditions that require differentiating overlapping combinations of cues (for a review, see Rudy and Sutherland, 1995).

We further hypothesized that the use of configurations during reinforcement learning would require bringing these hippocampal configural representations in contact with striatal reinforcement learning mechanisms. We reasoned that this relationship

should be reflected in functional connectivity between the hippocampus and the nucleus accumbens. The nucleus accumbens is known to play an important role in updating value in reinforcement learning (O'Doherty et al., 2004), and it also has strong anatomical connections with the anterior/ventral hippocampus (Groenewegen et al., 1987; Kelley and Domesick, 1982). This pathway provides a route through which configural associations in the hippocampus could be integrated with outcome signals in the nucleus accumbens (Ito et al., 2008; Pennartz et al., 2011). In particular, functional interactions between these structures could allow outcomes to be assigned to configural associations in the hippocampus.

To test these hypotheses, we sought to understand the neural and computational mechanisms by which people learn to predict outcomes based on configurations of cues, how changes in the predictive strength of cues modulate learning, and how people vary in their tendency to perceive and use configurations versus single cues during learning. We designed a study in which we manipulated the extent to which outcomes could be predicted by configurations versus cues, using two variants of a probabilistic classification task. As shown in Figure 1, participants predicted a weather outcome (rain or sun) based on visual cues. In the inseparable task, the outcome could be predicted *only* based on the configuration and not by the separable contributions of its elements. In the separable task, outcomes could be predicted by *either* the elements or their configuration, so that people could use either strategy. Earlier studies with probabilistic classification tasks had shown that, without any constraints, people vary in the extent to which they use a single cue or a combination of cues to predict outcomes (Gluck et al., 2002; Meeter et al., 2006; Poldrack et al., 2001; Shohamy et al., 2004). Here we sought to manipulate the structure of the task and leverage trial-by-trial learning models to obtain a tighter measurement of elemental versus configural learning and to establish a direct link between different computational strategies and their neural mechanisms.

Participants performed both tasks while being scanned with fMRI. We developed reinforcement learning models to quantify the extent to which choices reflected configural versus elemental learning within subjects, across subjects, and between tasks. We also characterized behavior by separately quantifying eye gaze, choices, reaction time (RT), and explicit learning of cue and configural probabilities, each of which tracked the model estimates. We found that people tended to engage in configural learning even when it was less efficient than elemental learning and that the extent to which they did so depended on the relative strength with which configurations versus elements predicted the outcome. Moreover, as we hypothesized, the tendency to predict outcomes based on configurations, both within and across individuals, was related to hippocampal responses and patterns of activity during choice and to functional connectivity between the hippocampus and the nucleus accumbens. Conversely, the tendency to predict outcomes based on elements was related to lateral occipital response during choice, consistent with cue representations becoming more configural across the visual processing hierarchy.
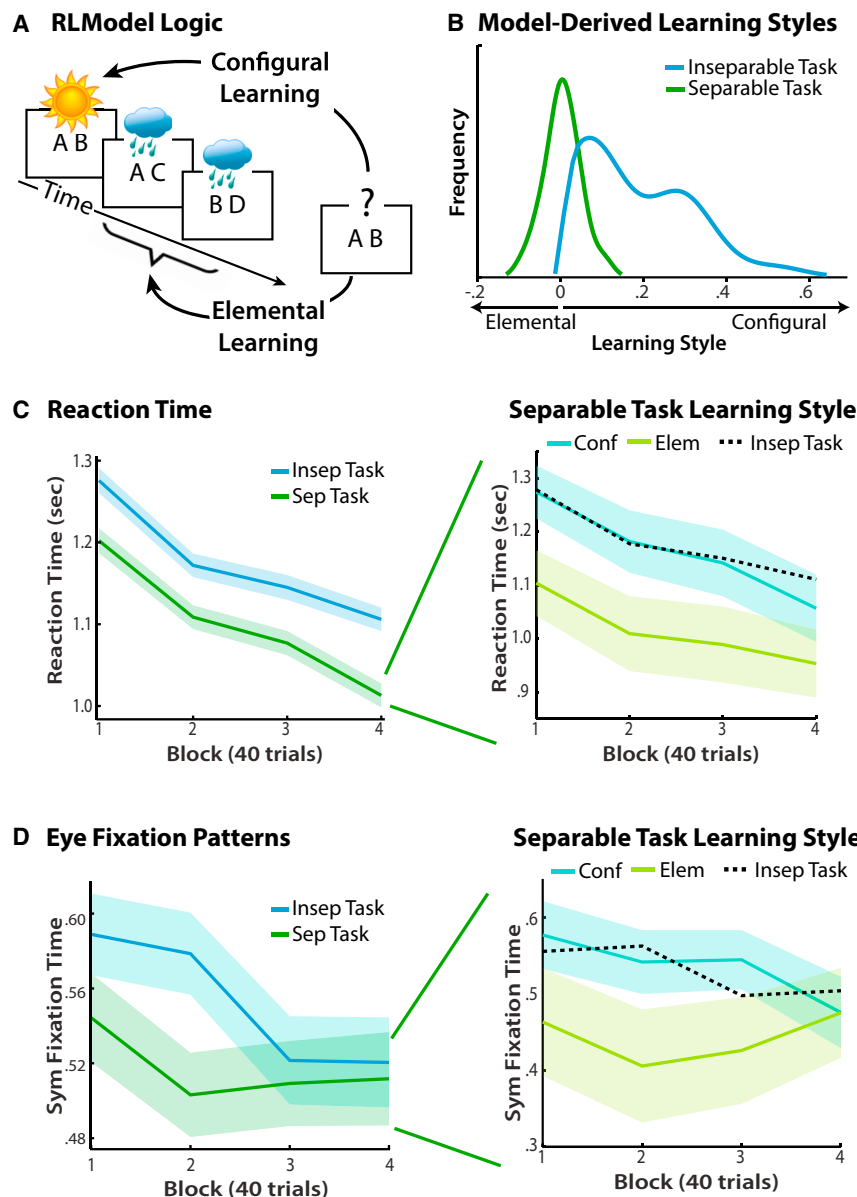
A   RLModel Logic

B   Model-Derived Learning Styles

C   Reaction Time

Separable Task Learning Style

D   Eye Fixation Patterns

Separable Task Learning Style

Figure 2. Learning Style Decomposition

(A) Schematic of reinforcement learning (RL) model logic. We used an RL model to quantify the extent to which each choice was predicted based on recent experiences with the presented cue combination (configural learning) or experiences with the individual cues (elemental learning).

(B) Distribution of RL model-derived learning styles. Learning styles were estimated by comparing the variance in choices explained by a model that only includes experiences with cue combinations versus a model that only includes experiences with individual cues. The probability distribution functions for each task are plotted.

(C) Reaction times are related to task and RL model estimates. On the left, mean RT in each task is plotted for all participants. On the right, RT is plotted separately for participants classified as configural or elemental learners in the separable task. For comparison, the dashed line plots mean RT for all participants in the inseparable task and error bars reflect 95% confidence intervals.

(D) Eye fixation patterns are related to task and RL model estimates. Symmetric fixation time indexes the proportional difference in time spent viewing each cue (0 = viewing only one cue; 1 = equal viewing) and is plotted for each task on the left. On the right, symmetric viewing time is plotted separately for participants classified as configural or elemental learners in the separable task. For comparison, the dashed line plots symmetric fixation time for all participants in the inseparable task and error bars reflect 95% confidence intervals.

RESULTS

Reinforcement Learning Models Quantify Variability in Configural versus Elemental Learning across Tasks and Individuals

Figure 1 shows the task design. On each of 240 trials, participants (n = 26) saw two cues and were asked to predict a weather outcome (sun versus rain). After making a choice, participants received feedback. Each participant performed two versions of the task. The versions differed only in how strongly the configurations and the cues predicted each outcome. In the inseparable task, only the configuration of the cues, not any individual cue, was predictive of the outcome. In the separable task, the individual cues were good predictors of the outcomes, as were the configurations.

To quantify the extent to which learning in both tasks involved elemental versus configural strategies on a trial-by-trial basis, we developed a variant of reinforcement learning models (Sutton and Barto, 1998). In general, reinforcement learning models are used to formally describe how a learner uses trial-by-trial reinforcement to update his/her responses. The basic logic is that each choice is based on a prediction about the outcome, a prediction that is updated for future choices if the actual outcome differs from the expected outcome. In the present task, the prediction on a trial might be formed based either on previous experience with the configuration seen on that trial or on previous experience with each element of the configuration separately. Because the same elements occur in different configurations, these two strategies make divergent predictions about trial-by-trial choices. Figure 2A shows a schematic of the approach of the variant of the model we used here; full modeling details are described in the STAR Methods and Table 1 compares our full model to partial variants. Briefly, the model leveraged participants' reinforcement history and choice patterns to separately estimate the choice values associated with each configuration and each element, providing an estimate of configural versus elemental values for each trial (within participants; Configural

CellPress

Table 1. Comparison of Reinforcement Learning Models

Model	Inseparable Task AIC	Separable Task AIC
Null	9,131.5	9,033.1
2 Betas	7,676.4	8,068.3
1 Alpha	7,320	7,691.8
Update unchosen	7,290.1	7,883.2
Full (10 betas, 2 alphas, only update chosen)	7,234.4	7,690.7

Choice Index) and of the degree to which each participant engaged in configural versus elemental learning (between participants; Learning Style Score).

We found that participants showed significant learning in both tasks but differed in their learning strategies. In the inseparable task, selection of the optimal weather outcome increased across trials (final proportion correct = 0.76, SE = 0.04; Beta = 0.007, SE = 0.001, z = 5.87, p < 0.000001). Consistent with the task design, the model estimated that 25 of 26 participants in the inseparable task showed a greater reliance on configural learning (Figure 2B). Participants also showed significant learning in the separable task (final proportion correct = 0.80, SE = 0.04; Beta = 0.005, SE = 0.001, z = 5.74, p < 0.000001), with roughly half of the participants in the separable task relying more on configural learning while the other half relied more on elemental learning, demonstrating that people learn about configurations even when they do not have to (Figure 2B; task difference: t(27) = 6.02, p < 0.00001). Participants maintained the same learning styles throughout each task, as measured by the correlation across the first and second halves of each task (inseparable r(27) = 0.66; separable r(27) = 0.46). Importantly, this Learning Style Score was not correlated with overall performance (final trial probability correct) in the separable task (r = 0.16, p = 0.4), indicating that both strategies were effective for learning this task.

Configural versus Elemental Learning Is Reflected in Choice Speed and Visual Processing of Choice Options

We validated the model estimates of learning style on separate behavioral measures, using RT and eye gaze (Figures 2C and 2D). For each measure, we first compared across the inseparable and separable tasks, and then we explored whether the model-derived learning strategy explained individual differences within the separable task.

RT differed between the inseparable and the separable tasks: participants were slower to make correct choices in the inseparable task (inseparable = 1,174 ms, SE = 39; separable = 1,100 ms, SE = 42; Beta = −74.5, SE = 22.7, F(1,25) = 10.7, p = 0.003; interaction with block: p > 0.57; Figure 2C). We asked whether RT was related to individual differences in learning styles within the separable task. We found that RT in the separable task was positively correlated with the degree to which participants relied on configural learning (Beta = 1,855, SE = 896, F(1,24) = 4.3, p = 0.05; Figure 2C). Moreover, while elemental learners' RTs in the separable task differed from the full group's RTs in the inseparable task (1,013 versus 1,174 ms; Beta = −120,

SE = 42, F(1,10) = 8.1, p = 0.02), the RTs of configural learners in the separable task were more comparable (1,163 versus 1,174 ms; Beta = −41, SE = 22, F(1,14) = 3.5, p = 0.08). Together, these findings indicate that making choices based on configurations, rather than individual cues, takes longer.

We next examined whether eye movements were related to task condition and to model estimates of configural and elemental learning. We measured the proportion of time that participants spent looking at each of the two cues while making their choice. We reasoned that, because configural learning relies on both cues, configural learners should distribute their looking time more symmetrically across the cues than elemental learners, who could base their decisions on a single informative cue. Consistent with this logic, symmetric viewing time was significantly greater in the inseparable compared to the separable task (inseparable = 1,174 ms, SE = 39; separable = 1,100 ms, SE = 42; Beta = 0.03, SE = 0.01, F(1,16) = 8.3, p = 0.01; interaction with block: p > 0.57), and configural learners in the separable task were numerically, but not significantly, more symmetric in their cue viewing (Beta = −0.81, SE = 0.46, F(1,15) = 3.1, p = 0.10; Figure 2D). As with the RT difference, elemental learners in the separable task differed in their fixation behavior from the full group in the inseparable task (Beta = 0.04, SE = 0.02, F(1,7) = 5.4, p = 0.05), but configural learners did not (Beta = 0.02, SE = 0.01, F(1,8) = 2.9, p = 0.12). Moreover, configural learning in the separable task was also related to an increased number of fixations (Beta = 10.8, SE = 3.3, F(1,15) = 11.5, p = 0.004), further linking configural learning to the increased integration of information across the cues. Lastly, both eye movement metrics (symmetric viewing time and fixation count) were correlated with RT in each task (r > 0.5, p < 0.04), suggesting that these behavioral metrics track with configural versus elemental learning style.

These individual differences in learning style observed in the separable task raise the possibility that perhaps people have a predisposition to engage in one learning style, regardless of the task structure. However, if learning style was only an individual trait, we would expect that it should be correlated across tasks, which it was not (r = 0.07, p = 0.73). Thus, the between-subject effects do not appear to reflect differences in how people learn, in general, but rather they may reflect differences in how participants interact with the structure of the cue-outcome associations built into each task.

In summary, a detailed analysis of participants' behavior shows converging evidence from choices, RT, eye tracking, and reinforcement learning models for how and when people predict outcomes based on configurations of cues. This behavior is captured in two measurements that can be used to probe neural activity: (1) the relative use of configural and elemental learning on a trial-by-trial basis, which we refer to as the Configural Choice Index, and (2) individual differences in configural versus elemental learning, which we refer to as the Learning Style Score.

Choices Based on Configurations Are Related to BOLD Activity in the Hippocampus

Having validated our modeling approach, we next used these model-based estimates to test the hypothesis that configural
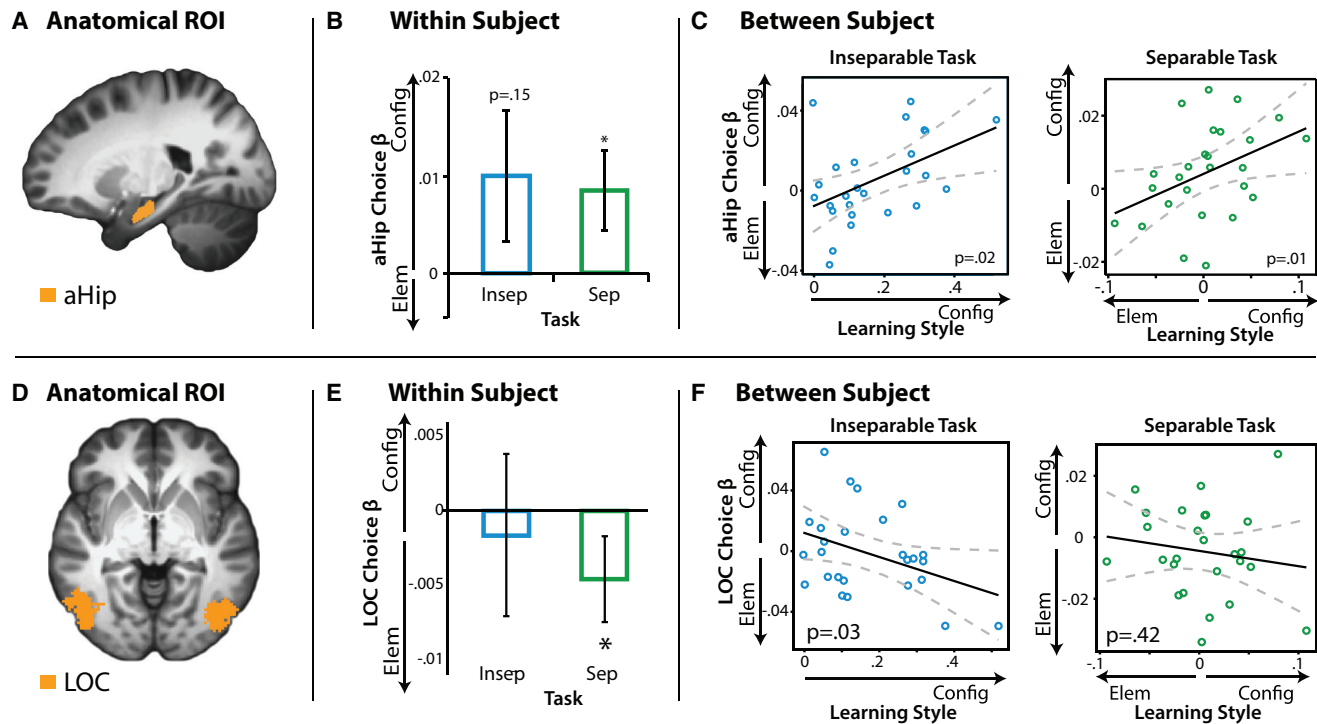
Figure 3. Configural Choices Are Related to Hippocampal BOLD Responses Both within and across Subjects

(A) Example anterior hippocampal (aHip) ROI. Anatomical ROIs were drawn for each participant in their native space.

(B) The aHip BOLD responses track configural choice value. Trial-specific estimates of aHip BOLD responses during each choice were used to predict whether that choice was based on configurally or elementally learned values (Configural Choice Index). The aHip was more active when participants made choices that reflected configurally as compared to elementally learned values in the separable task and marginally more active in the inseparable task. Error bars indicate the SE of the β estimate.

(C) Individual differences in configural learning are related to BOLD aHip responses. Correlation between individual differences in configural versus elemental learning and degree to which aHip tracks configural choice values are shown.

(D) Lateral occipital cortex (LOC) ROI. The ROI was generated with an anatomically constrained meta-analysis for object processing.

(E) LOC BOLD responses track elemental choice value but only in the separable task. Trial-specific estimates of LOC BOLD responses during each choice were used to predict whether that choice was based on configurally or elementally learned values (Configural Choice Index); negative betas indicate greater activity during elemental choices. Error bars indicate the SE of the β estimate.

(F) Individual differences in elemental learning are related to LOC responses but only in the inseparable task. Lines depict group-level linear effects ± 2 SE.

learning is related to BOLD responses in the hippocampus. We focused on an anatomically defined *a priori* region of interest, the anterior hippocampus (aHip; Figure 3A), which is known to support stimulus-stimulus encoding and which has strong anatomical connections to the ventral striatum (Cohen and Eichenbaum, 1993; Groenewegen et al., 1987). We began by investigating aHip choice activity, predicting that it would be greatest when participants made choices using configural as compared to elemental learning (Configural Choice Index).

Consistent with this prediction, aHip choice activity was correlated with the use of configurations, both within and between subjects as well as across tasks. In the inseparable task, which encouraged configural learning, we found that those participants who engaged in more configural learning showed a stronger correlation between aHip choice response and the Configural Choice Index (Beta = 0.02, SE = 0.007, F(1,26) = 7.2, p = 0.01; Figure 3C). This pattern was replicated in the separable task (Beta = 0.01, SE = 0.004, F(1,24) = 7.0, p = 0.01; Figure 3C). Moreover, within subjects, aHip BOLD responses reliably

tracked the Configural Choice Index in the separable task (Beta = 0.008, SE = 0.005, F(1,27) = 4.2, p = 0.05; Figure 3B). A similar positive relationship was seen in the inseparable task, though it did not reach statistical significance (Beta = 0.01, SE = 0.008, F(1,27) = 2.2, p = 0.15; Figure 3B). This relationship between BOLD activity and the Configural Choice Index was not found in the striatum in either task (p values > 0.23; Figure S2). Thus, we find complementary and independent evidence from both tasks that the hippocampus plays a role in configural learning.

The inseparable task depends on configural learning and, therefore, supports a role for the hippocampus in predicting outcomes based on configurations. However, this task was contrived to require configural learning, therefore raising the question of whether the hippocampus supports configural learning when it is not required. This question is addressed by the findings from the separable task, where configural learning was not required and where learning strategy was not correlated with learning success, suggesting that the relationship between

aHip choice activity and configural learning is driven by configural learning per se.

To more specifically investigate the neural mechanisms underlying elemental choices, we performed exploratory analyses in the perirhinal cortex (PRC) and lateral occipital cortex (LOC)—two regions that are known to contribute to object processing (Buckley and Gaffan, 1998; Grill-Spector et al., 2001; Kourtzi and Kanwisher, 2000; Murray and Richmond, 2001). We found that BOLD activity in the PRC was not reliably related to elemental choices, either within or across subjects (p values > 0.1). However, in the LOC we found that BOLD activity was related to elemental learning: in the separable task, LOC activity was greater within subjects when participants made choices using elemental learning (Beta = −0.01, SE = 0.005, F(1,30) = 4.87, p = 0.04; Figure 3D). Similarly, in the inseparable task, the tendency to engage in elemental learning was related to choice-related activity in the LOC across subjects (Beta = −0.02, SE = 0.008, F(1,27) = 5.26, p = 0.03; Figure 3F), despite the fact that, in this condition, engaging in elemental learning hurt performance. Together, these results suggest that earlier portions of the ventral processing stream complement the role of the hippocampus in configural learning through their involvement in elemental learning.

Configural Learning Is Related to Functional Connectivity between the Hippocampus and Nucleus Accumbens

The finding that the hippocampus contributes to feedback-driven learning about configurations raises questions about the mechanism underlying these effects. In particular, as in standard reinforcement learning tasks, each trial of our tasks had a choice phase, during which the participant saw the shapes and made a response, followed by an outcome phase, during which the participant learned whether the choice was correct or incorrect. Studies of reinforcement learning typically analyze these two phases separately, and the outcome phase is commonly associated with responses in the nucleus accumbens (O'Doherty et al., 2004). Thus, a question of relevance here is how the responses in the hippocampus to configurations, at the time of choice, interact with responses at the time of outcome, putatively in the nucleus accumbens? To answer this question and to assess whether interactions between the hippocampus and the nucleus accumbens are associated with configural learning, we used two approaches: (1) a staggered beta series analysis to measure the coordination between choice and outcome responses and outcome and later choice responses, and (2) a background connectivity analysis to measure sustained functional connectivity throughout the course of learning.

The staggered beta series analysis first assessed whether trial-by-trial fluctuations in aHip BOLD responses during choice predicted outcome responses in the nucleus accumbens, above and beyond choice responses in the nucleus accumbens itself (Figure 4A). We first asked whether differences in the predictive structure of the inseparable and separable tasks modulated the strength of choice-outcome coupling. We found that aHip choice responses were more correlated with nucleus accumbens outcome responses during the inseparable task than in the separable task (inseparable = 0.18, SE = 0.03; separable =

0.09, SE = 0.04; Beta = 0.21, SE = 0.09, F(1,29) = 4.6, p = 0.04; Figure 4B). Importantly, the reverse did not hold: nucleus accumbens choice responses were not more correlated with aHip feedback response in the inseparable task (p = 0.31; see Table S1 for a full set of control analyses), consistent with directional interactions between the hippocampus and nucleus accumbens. Interestingly, choice-outcome coupling between the dorsal/lateral striatum and nucleus accumbens was greater in the separable task as compared to the inseparable task (Figure S3), suggesting that elemental learning may be more related to within-striatum connectivity.

We next asked whether individual differences in choice-outcome coupling between the aHip and nucleus accumbens were related to configural learning within tasks, again controlling for nucleus accumbens choice responses. We found that participants who relied more on configural learning in the inseparable task showed greater correlations between aHip choice responses and nucleus accumbens outcome response (Beta = 0.13, SE = 0.06, F(1,24) = 4.09, p = 0.05), and a similar relationship was seen in the separable task (Beta = 0.16, SE = 0.08, F(1,23) = 4.44, p = 0.05; Figure 4C). The directionality of these interactions, however, was only partially supported by control analyses (Tables S2 and S3). Specifically, while nucleus accumbens choice→aHip outcome correlations were not significantly correlated with configural learning in the inseparable task (p = 0.12), there was a trend toward this relationship in the separable task (p = 0.06). Thus, while these between-participant analyses support the hypothesis that configural learning is related to functional connectivity between the aHip and nucleus accumbens, there is less evidence that this relationship is mediated by aHip responses driving nucleus accumbens outcome responsiveness. The weak temporal asymmetry observed in these analyses may either reflect bleeding between choice and outcome period estimates in our model or an imperfect specificity in the timing of the coupling across regions.

Thus far we examined the relationship between choice and outcome activations within a trial; participants, however, must also use feedback to update their representations for future choices. Therefore, we next explored how outcome-related activity influences subsequent choice-related activity. We defined the next relevant choice as occurring on the next trial on which participants were presented with the same configuration of options (Figure 4D) to relate outcome→subsequent choice coupling to configural learning. In the separable task, we found that nucleus accumbens outcome activity was more correlated with subsequent aHip choice activity in participants who were more configural in their learning (Beta = 0.05, SE = 0.02, F(1,24) = 6.18, p = 0.01; Figure 4C). As with the within-trial analyses, this relationship was observed while controlling for aHip outcome's correlation with aHip subsequent choice, and, further, it was asymmetric (i.e., aHip outcome was not related to nucleus accumbens subsequent choice in the same manner, p = 0.96). This relationship was also found to be in the same direction in the inseparable task, but it was not statistically significant (Beta = 0.03, SE = 0.02, F(1,24) = 1.76, p = 0.22; Figure 4C; aHip outcome→nucleus accumbens subsequent choice p = 0.54). Neither was there a significant interaction with task

A Choice -> Outcome Coupling

B Within Subject

C Between Subject

D Outcome -> Subsquent Choice Coupling

E Within Subject
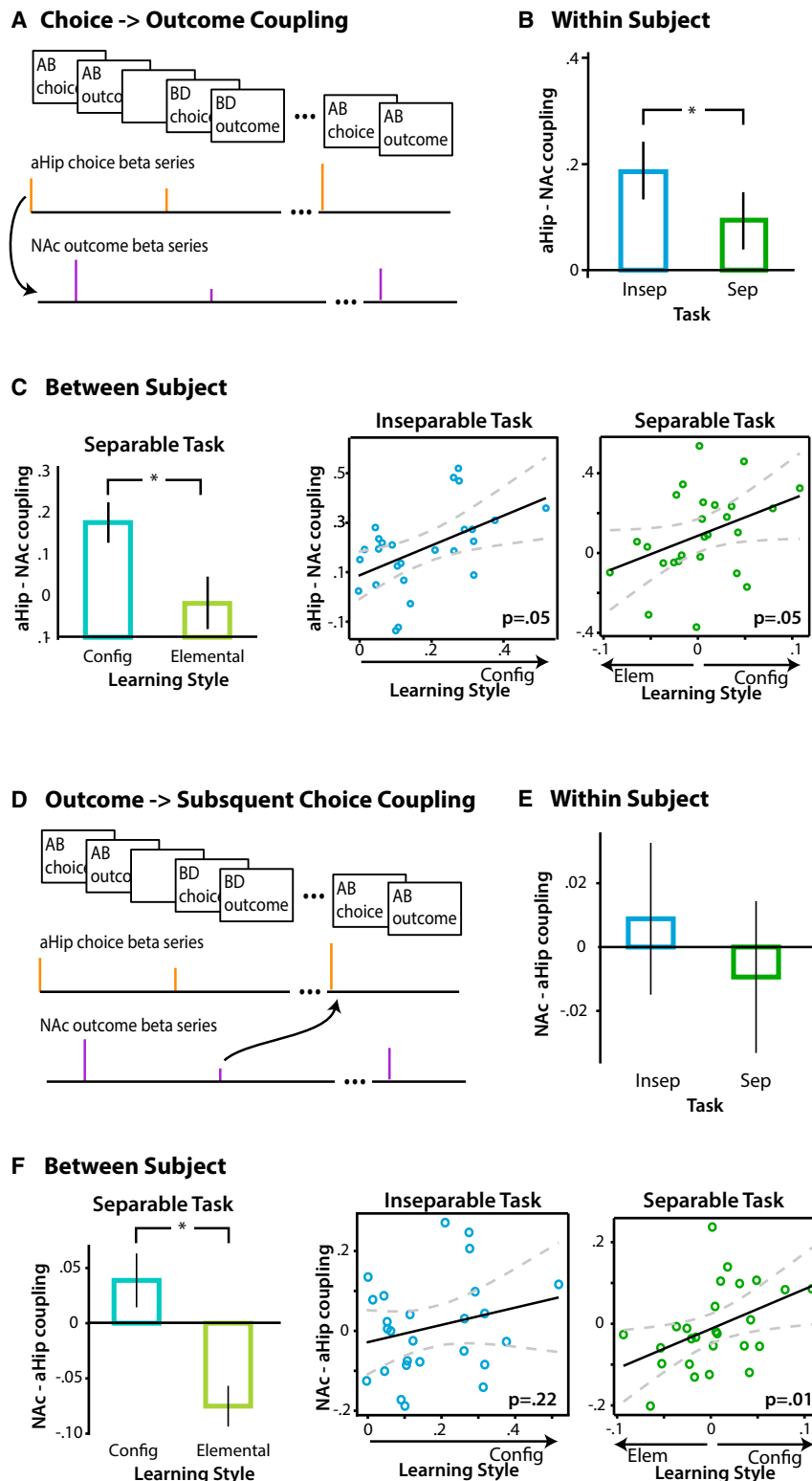

F Between Subject

Figure 4. Configural Learning Is Related to Hippocampal-Nucleus Accumbens Outcome Functional Connectivity

(A) Schematic illustration of choice-outcome connectivity analysis. Within-trial choice and outcome BOLD activations were used to assess how choice-related activations in aHip drive subsequent nucleus accumbens (NAc) outcome activations on a trial-by-trial basis.

(B) Task differences in aHip-NAc coupling. The aHip choice responses were more correlated with NAc outcome responses in the inseparable task.

(C) aHip-NAc coupling and individual learning differences. The left graph plots average aHip-NAc coupling for participants classified as configural or elemental learners in the separable task. The middle graph plots the positive relationship between aHip-NAc choice-outcome coupling configural learning during the inseparable task.

(D) Schematic illustration of outcome-subsequent choice connectivity analysis. Across-trial choice and outcome BOLD activations were used to assess how NAc outcome-related activations drove aHip choice-related activation when participants next saw the relevant configuration.

(E) No outcome-subsequent choice differences across tasks. The aHip choice responses were more correlated with NAc outcome responses in the inseparable task.

(F) NAc-aHip coupling and individual learning differences. The left graph plots NAc-aHip coupling for participants classified as configural or elemental learners in the separable task. Correlations between NAc-aHip outcome-subsequent choice coupling and configural learning during the inseparable task (middle) and the separable task (right) are shown. Lines depict group-level linear effects ± 2 SE.
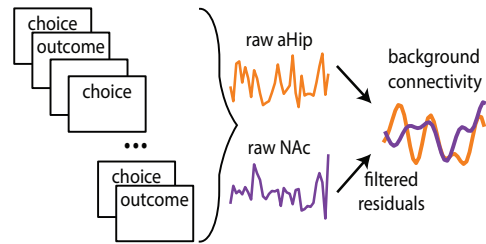
behavioral impact of outcome→subsequent choice coupling is less robust than choice-outcome coupling within trials.

We additionally measured background connectivity to investigate whether functional interactions operating in the background of trial-evoked responses were related to learning. We first filtered out trial-evoked responses using a combination of general linear models (GLMs) and temporal frequency filters, and then we predicted nucleus accumbens BOLD fluctuations with the aHip (Figure 5A). In both tasks, participants who engaged in more configural learning had stronger functional connectivity between the aHip and nucleus accumbens throughout learning (inseparable: Beta = 0.0037, SE = 0.0016, $F_{(1,24)}$ = 5.0, p = 0.04;
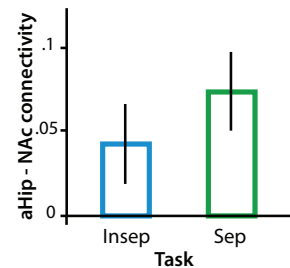
(p = 0.54). Thus, while there is some evidence that outcome-related activity in the nucleus accumbens is related to hippocampal activity the next time that choice reappears, the

separable: Beta = 0.0035, SE = 0.0016, $F_{(1,24)}$ = 5.2, p = 0.03; Figure 5C). By contrast, dorsal and lateral regions of the striatum did not show this relationship with the nucleus accumbens
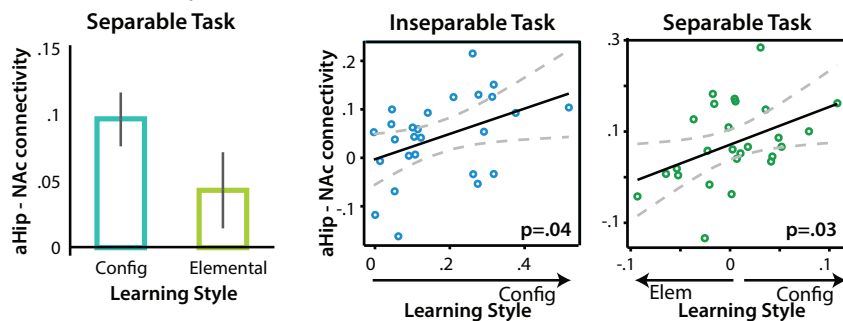
Figure 5. Configural Learning Is Related to Hippocampal-Nucleus Accumbens Background Connectivity

(A) Schematic illustration of background connectivity analysis. Background BOLD responses were isolated by filtering raw signal to remove task-evoked activations (using a GLM) and bandpass filtered to 0.08–0.009 Hz. Connectivity was estimated by predicting NAc background signal with background signal from the aHip.

(B) Task differences in background connectivity. Background connectivity between the aHip and NAc was not influenced by the task demands.

(C) aHip-NAc background connectivity and individual learning differences. The left graph plots average aHip-NAc background connectivity for participants classified as configural or elemental learners in the separable task. The middle graph plots the positive relationship between aHip-NAc background connectivity and configural learning in the inseparable task. The right graph plots the same relationship during the separable task. Lines depict group-level linear effects ± 2 SE.
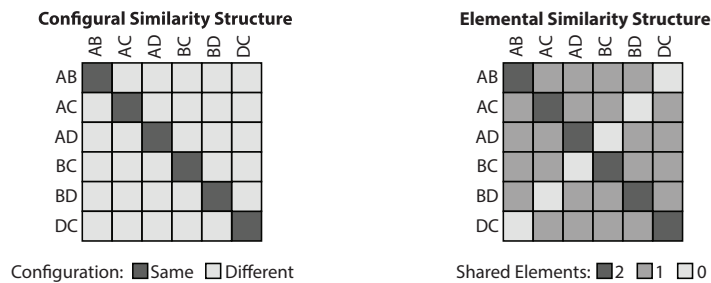
(Figure S3). Furthermore, when both background connectivity and choice-outcome coupling were included in a single GLM to predict each participant's learning style, aHip-nucleus accumbens background connectivity continued to be reliably associated with configural learning (inseparable: Beta = 0.63, SE = 0.29, t(21) = 2.2, p = 0.04; separable: Beta = 0.19, SE = 0.08, t(21) = 2.3, p = 0.04). This suggests that ongoing shifts in the strength of functional connectivity between these regions set the stage for different ways of learning. The strength of aHip-nucleus accumbens background connectivity, however, did not significantly differ between tasks (separable = 0.07, SE = 0.02; inseparable = 0.04, SE = 0.02; p > 0.15; Figure 5C), suggesting that these low-frequency fluctuations may be less influenced by task demands but instead bias the nature of learning within tasks.

Patterns of Choice-Related Activity in the aHip Reflect Configural Representations Better Than Elemental Representations

We used a representational similarity analysis to assess whether, in addition to being more active while participants used configural learning to make choices, the patterns of activity elicited in the aHip reflected the configuration of cues rather than individual elements. Specifically, we generated two representational similarity matrix (RSM) models. The configural model predicts that activity patterns evoked on trials containing the same configurations would be positively correlated but that all other trial pairs (including those that share one element) would be less correlated. Conversely, the elemental model predicts that the strength of correlation would be a linear function of the number of shared elements, i.e., 0 shared < 1 shared < 2 shared (Figure 6). For the inseparable task, which required configural learning, aHip activ-

ity patterns were reliably correlated with the configural model (mean = 0.05, SE = 0.01; t(25) = 3.39, p = 0.002), and they were significantly more correlated with the configural than the elemental model (mean = 0.02, SE = 0.01; t(25) = 2.20, p = 0.04) (Figure 6C). Moreover, the configural model was a significantly better predictor of aHip activation patterns in the inseparable task than in the separable task (t(25) = 2.80, p = 0.01) (Figure 6C). Findings from the separable task were not well fit by either model (configural = −0.02, SE = 0.02; elemental = −0.01, SE = 0.02; p values > 0.22; model comparison: t(25) = 0.60, p = 0.55; task*model interaction: F(1,25) = 2.94; p = 0.1), and neither model predicted patterns of activity in striatal regions of interest (ROIs; p values > 0.2). The Learning Style Score, however, was not reliably correlated with the degree to which aHip patterns correlated with the configural model across participants (configural model correlation – elemental model correlation; inseparable task r = 0.18, p = 0.38; separable task r = −0.04, p = 0.83). Taken together, these findings suggest that, when required by the task at hand, the aHip represents the configurations of cues to guide actions, consistent with its well-documented role in relational memory.
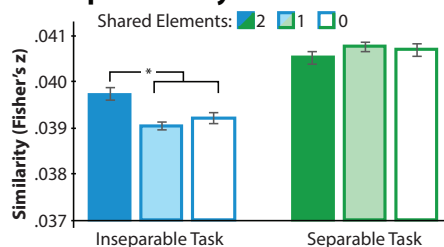
Configural Learning Results in More Flexible Knowledge Than Elemental Learning

The analyses so far show that people can learn to predict outcomes based on either elements or their configuration, that these approaches differ behaviorally and neurally, but that they do not relate to overall levels of accuracy during learning. Given that configural learning requires more resources, effort, and time than elemental learning, this raises questions about the possible benefits of configural learning. One possibility is that engaging hippocampal systems during learning may result in more flexible and explicit knowledge (e.g., Foerde et al., 2006; Reber et al., 1996; Eichenbaum and Cohen, 2001). We tested this idea using

A Similarity Analysis Logic

Configural Similarity Structure

Elemental Similarity Structure

Configuration: ■ Same □ Different

Shared Elements: ■ 2 ▨ 1 □ 0

B aHip Similarity

Shared Elements: ■ 2 ▨ 1 □ 0

C aHip RSAT
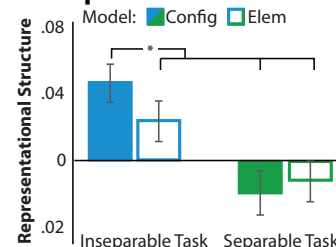
Model: ■ Config □ Elem

Figure 6. Patterns of Hippocampal Choice-Activity Reflect Configural Rather Than Elemental Content

(A) Schematics of models used to conduct representational similarity analyses (RSAs). Configural model (left) predicts high similarity across trials with the same configuration but is blind to the elements. Elemental model (right) predicts that similarity will be a linear function of the number of overlapping elements across trials.

(B) Hippocampal similarity structure organized by number of shared elements. Similarity is quantified as Pearson's correlation coefficients (Fisher z transformed for statistical comparisons).

(C) RSA model comparison. aHip similarity structure is most correlated with the configural model (filled bars) during the inseparable task.

Error bars reflect SEM.

participants' post-task ratings to determine whether individual differences in learning style were related to participants' explicit knowledge of cue-weather associations. Immediately following each task, participants rated the likelihood of rain and sun based on each cue and cue configuration (Figure S1A). This type of test has been used to assess hippocampal learning (Foerde et al., 2006; Reber et al., 1996), so we reasoned that rating accuracy should be correlated with learning success to the extent that hippocampal processes support learning performance. We found a strong relationship between accuracy of ratings and learning accuracy in the inseparable task (r(25) = −0.76, p < 0.00001; Figure S1B). Further, we found that this correlation depended on how participants learned in the separable task, with greater configural learning associated with stronger correlations between learning and rating accuracy (Beta = 2.56, F(1,22) = 4.0, p = 0.06; Figure S1C). This analysis reveals a potential implication of these learning differences: configural learning is more tightly linked with explicit awareness of outcome contingencies.

DISCUSSION

We found that learning to make decisions based on configurations of cues is not a simple extension of cue-response learning; configural learning is related to qualitative differences in the engagement of, and interaction between, distinct learning and memory systems. Specifically, basing choices on cue configurations rather than individual cues was associated with greater BOLD activity in the hippocampus and greater functional connectivity between the hippocampus and nucleus accumbens, both within and between subjects. Configural learning was also more closely associated with subsequent explicit knowledge of experienced contingencies, suggesting that engaging hippocampal circuits during learning influences the flexibility with which people can

then use learned contingencies. All of these results are consistent with a model in which the ability to learn about configurations holistically, and distinct from their elements, depends on representational pattern separation-like processes supported by the hippocampus and their interactions with the ventral striatum.

These findings call for greater consideration of how information is integrated across multiple cues, from perception to learning to decision-making. Although conscious perception favors the configural level of representation, these high-level representations must be constructed from disparate features (e.g., color, shape, texture, and sound) to form coherent wholes. Variability in how cue combination (Ernst and Bülthoff, 2004) parses the environment may determine what we perceive as discrete items, dictating the units that are encoded by learning and memory systems and associated with outcomes. In turn, learning statistical regularities in the co-occurrence of features can guide how features are integrated into perceptual units (Fiser and Aslin, 2002; Saffran et al., 1996) and neural representations (Schapiro et al., 2012, 2013). Statistical regularities in stimulus co-occurrence have also been used to explain animals' expectations of reinforcement in conditioning experiments (Gershman et al., 2010), including the extent of their reliance on stimulus configurations versus elements (Courville et al., 2005). Interestingly, the statistical regularity of feature combinations has also been shown to influence whether elements or configurations of elements form the basis of learning (Turk-Browne et al., 2008). Rather than manipulating the co-occurrence of cues or their perceptual features, we manipulated their predictive power. In the separable task, where individual cues were predictive, participants were more likely to integrate cues in the decision, combining separate elementally learned values. By contrast, in the inseparable task, participants were more likely to integrate the cues in memory via hippocampal processing and attach a value to that configural representation. Thus, our findings add to the statistical learning literature by demonstrating that learning influences the integration of information at multiple stages of processing, from perception to decisions.

CellPress

Our findings also have implications for a long-standing challenge for reinforcement learning models, namely, credit assignment (Roelfsema et al., 2010; Roelfsema and van Ooyen, 2005; Sutton, 1984). The basic tenets of reinforcement learning are straightforward: positive outcomes should increase the value of the states that generated them. What is much less straightforward, however, is determining which states actually generated the outcome (e.g., Gershman et al., 2010; Niv et al., 2015). Much like in everyday experience, multiple cues were simultaneously presented in our tasks, creating ambiguity as to which permutation of the possible cues was most informative. We find that the brain has multiple mechanisms for assigning credit in these situations, with interactions between the hippocampus and nucleus accumbens implicated in assigning credit to the cue configurations. Interestingly, temporal ambiguity is also a major source of credit assignment problems, with delays between cues and outcomes limiting learning (Kobayashi and Schultz, 2008; Maddox et al., 2003; Sutton, 1984). The hippocampus may also play role in bridging these gaps (Foerde et al., 2013; Foerde and Shohamy, 2011), suggesting that incorporating the unique learning attributes of the hippocampus into reinforcement learning and decision-making theories could generate new progress in this field (see also Gershman and Daw, 2017; Gluth et al., 2015; Peters and Büchel, 2010; and Gershman et al., 2010).

The involvement of the hippocampus in reinforcement learning has implications for ongoing debates concerning the organization of multiple memory systems in the brain. Reinforcement learning has traditionally been attributed to habitual striatal mechanisms (Knowlton et al., 1996), while it is often argued that the hippocampus is uninvolved in such learning or even competes against it (Foerde et al., 2006; Knowlton et al., 1994; Poldrack et al., 2001; Wimmer et al., 2014). Many of these conclusions have rested on data from feedback-based probabilistic classification tasks, such as the weather prediction task, similar to the task used here. This task has been used to demonstrate comparatively intact probabilistic learning in medial temporal lobe amnesia and negative relationships between hippocampal and striatal BOLD responses (Knowlton et al., 1994; Poldrack et al., 2001). Critically, elemental learning is much more efficient than configural learning in the classic weather prediction task, because each of the four abstract cues is presented in isolation and in a large number of combinations (Gluck et al., 2002; Shohamy et al., 2004). Indeed, this observation has been quantified in prior work where individual differences in the use of different strategies were reported in a task that left the strategy options unconstrained (Gluck et al., 2002; Meeter et al., 2006; Shohamy et al., 2004).

By contrast, the underlying statistical relationships between cues and outcomes in our tasks fostered or even required configural learning. These changes resulted in a shift toward configural learning, revealing a new role for hippocampal processes in feedback-based learning. This finding suggests that many of the foundational results in the multiple memory systems literature may only hold for elemental learning, rather than reflecting essential dissociations between hippocampal and striatal involvement in feedback-based learning more generally. In that sense, our findings contribute to an emerging framework in which the striatum and the hippocampus cooperate—rather than compete—to support learning, memory, and decision-making (Dickerson et al., 2011; Murty et al., 2015; Wimmer and Shohamy, 2012; Davidow et al., 2016).

Our findings shed light both on the conditions that elicit configural learning and on individual differences in this tendency. On one hand, we found that many participants tend to engage in configural learning even in the separable task, when they were not required to and were specifically instructed against it. At first blush, this suggests that certain individuals may have an *a priori* tendency to process configurations. At the same time, however, we found that strategies did not correlate across the two tasks, indicating that configural learning in the separable task is not merely a trait that the participants bring with them. Rather, the learning style differed as a function of task structure. The individual differences in learning style are nevertheless notable, especially because all participants received explicit instructions prior to each task, but clearly some participants did not follow the elemental instructions provided in the separable task. Perhaps participants deviated from instructions because the strong relationships between configurations and weather in the task were somehow more salient than the instructed element-weather relationships. If this is the case, participants would not have intentionally adopted a configural strategy, but rather it would have been elicited through experience with the task (see also Doll et al., 2009).

In providing support for cooperative memory systems, our findings also argue for a more interactive way of thinking about configural learning. The configural theory of hippocampal function was first proposed by Sutherland and Rudy (1989) to account for the effects of isolated hippocampal lesions on instrumental and Pavlovian conditioning (Alvarado and Rudy, 1995; Dusek and Eichenbaum, 1998; Kumaran et al., 2009; but see Gallagher and Holland, 1992 and Saunders and Weiskrantz, 1989). Configural and relational models have since become arguably the most pervasive framework for understanding hippocampal contributions to episodic memory (e.g., Davachi, 2006; Eichenbaum and Cohen, 2014; and Eichenbaum et al., 2007) and other forms of learning (Chun and Phelps, 1999; Goldfarb et al., 2016). Our findings are consistent with the proposition that the hippocampus supports configural representations, but they also suggest that these representations may not be sufficient for learning. Rather, configural reinforcement learning may be the product of interactions between the hippocampus and striatal regions, which are poised to process outcomes and initiate motor output. Indeed, disconnection between the hippocampus and the nucleus accumbens has been shown to impair context-dependent learning (Ito et al., 2008). Our work adds to this lesion approach by using computational models to quantitatively characterize the online dynamics of these regions and tie them to the dynamics of learning.

More broadly, because context can be conceptualized as a configuration of items, understanding how and why people engage in configural learning is instrumental for identifying the circumstances under which decisions depend on context. Learning that the value of some actions depends on context is essential for healthy behavior. As one example, when engaging in configural learning related to everyday behaviors, whether

CellPress

consuming chocolate, alcohol, or drugs, rewards are associated with a constellation of cues (e.g., alcohol, bar, and friends) such that each cue on its own would not trigger a behavior. The combination of cues, however, would also be required to unlearn response associations. Thus, to reshape behavior, one must go beyond characterizing habits to understand the learning process that generated them. Our work provides an approach to disentangle the contributions of configural and elemental learning in an experimental setting, thereby opening the door to developing new manipulations and interventions.

STAR★METHODS

Detailed methods are provided in the online version of this paper and include the following:

- KEY RESOURCES TABLE
- CONTACT FOR REAGENT OR RESOURCE SHARING
- EXPERIMENTAL MODEL AND SUBJECT DETAILS
- METHOD DETAILS
  - Behavioral Procedures
  - MRI Methods
  - Eye Tracking Procedures
- QUANTIFICATION AND STATISTICAL ANALYSIS
  - Learning Success Model
  - Reinforcement Learning Model
  - FMRI Statistical Analysis

SUPPLEMENTAL INFORMATION

Supplemental Information includes four figures and three tables and can be found with this article online at https://doi.org/10.1016/j.neuron.2018.03.042.

AUTHOR CONTRIBUTIONS

All authors designed the experiment and analyses. K.D. and B.B.D. performed the experiment. K.D. analyzed the data. K.D., D.S., and N.D.D. wrote the paper.

DECLARATION OF INTERESTS

The authors declare no competing interests.

REFERENCES

Al-Aidroos, N., Said, C.P., and Turk-Browne, N.B. (2012). Top-down attention switches coupling between low-level and high-level areas of human visual cortex. Proc. Natl. Acad. Sci. USA 109, 14675–14680.

Alvarado, M.C., and Rudy, J.W. (1995). Rats with damage to the hippocampal-formation are impaired on the transverse-patterning problem but not on elemental discriminations. Behav. Neurosci. 109, 204–211.

Amaral, D.G., and Lavenex, P. (2006). Hippocampal neuroanatomy. In The Hippocampus Book, P. Andersen, R. Morris, D. Amaral, T. Bliss, and J. O'Keefe, eds. (Oxford University Press).

Bates, D., Mächler, M., Bolker, B., and Walker, S. (2015). Fitting linear mixed-effects models using lme4. J. Stat. Softw. 67, 1–48.

Behzadi, Y., Restom, K., Liau, J., and Liu, T.T. (2007). A component based noise correction method (CompCor) for BOLD and perfusion based fMRI. Neuroimage 37, 90–101.

Brainard, D.H. (1997). The Psychophysics Toolbox. Spat. Vis. 10, 433–436.

Buckley, M.J., and Gaffan, D. (1998). Perirhinal cortex ablation impairs visual object identification. J. Neurosci. 18, 2268–2275.

Chun, M.M., and Phelps, E.A. (1999). Memory deficits for implicit contextual information in amnesic subjects with hippocampal damage. Nat. Neurosci. 2, 844–847.

Cohen, N.J., and Eichenbaum, H. (1993). Memory, Amnesia, and the Hippocampal System (Cambridge, MA: The MIT Press).

Cordes, D., Haughton, V.M., Arfanakis, K., Carew, J.D., Turski, P.A., Moritz, C.H., Quigley, M.A., and Meyerand, M.E. (2001). Frequencies contributing to functional connectivity in the cerebral cortex in "resting-state" data. AJNR Am. J. Neuroradiol. 22, 1326–1333.

Courville, A.C., Daw, N.D., and Touretzky, D.S. (2005). Similarity and discrimination in classical conditioning: a latent variable account. In Advances in Neural Information Processing Systems 17: Proceedings of the 2004 Conference, NIPS 2004 (Neural Information Processing Systems Foundation).

Cox, R.W. (1996). AFNI: software for analysis and visualization of functional magnetic resonance neuroimages. Comput. Biomed. Res. 29, 162–173.

Davachi, L. (2006). Item, context and relational episodic encoding in humans. Curr. Opin. Neurobiol. 16, 693–700.

Davidow, J.Y., Foerde, K., Galván, A., and Shohamy, D. (2016). An upside to reward sensitivity: the hippocampus supports enhanced reinforcement learning in adolescence. Neuron 92, 93–99.

Daw, N.D., O'Doherty, J.P., Dayan, P., Seymour, B., and Dolan, R.J. (2006). Cortical substrates for exploratory decisions in humans. Nature 441, 876–879.

Dayan, P., and Daw, N.D. (2008). Decision theory, reinforcement learning, and the brain. Cogn. Affect. Behav. Neurosci. 8, 429–453.

Dickerson, K.C., Li, J., and Delgado, M.R. (2011). Parallel contributions of distinct human memory systems during probabilistic learning. Neuroimage 55, 266–276.

Doll, B.B., Jacobs, W.J., Sanfey, A.G., and Frank, M.J. (2009). Instructional control of reinforcement learning: a behavioral and neurocomputational investigation. Brain Res. 1299, 74–94.

Duncan, K., Tompary, A., and Davachi, L. (2014). Associative encoding and retrieval are predicted by functional connectivity in distinct hippocampal area CA1 pathways. J. Neurosci. 34, 11188–11198.

Dusek, J.A., and Eichenbaum, H. (1998). The hippocampus and transverse patterning guided by olfactory cues. Behav. Neurosci. 112, 762–771.

Eichenbaum, H., and Cohen, N.J. (2001). From Conditioning to Conscious Recollection: Memory Systems of the Brain (Oxford University Press).

Eichenbaum, H., and Cohen, N.J. (2014). Can we reconcile the declarative memory and spatial navigation views on hippocampal function? Neuron 83, 764–770.

Eichenbaum, H., Yonelinas, A.P., and Ranganath, C. (2007). The medial temporal lobe and recognition memory. Annu. Rev. Neurosci. 30, 123–152.

Ernst, M.O., and Bülthoff, H.H. (2004). Merging the senses into a robust percept. Trends Cogn. Sci. 8, 162–169.

Fiser, J., and Aslin, R.N. (2002). Statistical learning of new visual feature combinations by infants. Proc. Natl. Acad. Sci. USA 99, 15822–15826.

Foerde, K., and Shohamy, D. (2011). Feedback timing modulates brain systems for learning in humans. J. Neurosci. 31, 13157–13167.

Foerde, K., Knowlton, B.J., and Poldrack, R.A. (2006). Modulation of competing memory systems by distraction. Proc. Natl. Acad. Sci. USA *103*, 11778–11783.

Foerde, K., Race, E., Verfaellie, M., and Shohamy, D. (2013). A role for the medial temporal lobe in feedback-driven learning: evidence from amnesia. J. Neurosci. *33*, 5698–5704.

Gallagher, M., and Holland, P.C. (1992). Preserved configural learning and spatial learning impairment in rats with hippocampal damage. Hippocampus *2*, 81–88.

Gershman, S.J., and Daw, N.D. (2017). Reinforcement Learning and Episodic Memory in Humans and Animals: An Integrative Framework. Annu. Rev. Psychol. *68*, 101–128.

Gershman, S.J., Blei, D.M., and Niv, Y. (2010). Context, learning, and extinction. Psychol. Rev. *117*, 197–209.

Glimcher, P.W. (2011). Understanding dopamine and reinforcement learning: the dopamine reward prediction error hypothesis. Proc. Natl. Acad. Sci. USA *108* (*Suppl 3*), 15647–15654.

Gluck, M.A., Shohamy, D., and Myers, C. (2002). How do people solve the "weather prediction" task?: individual variability in strategies for probabilistic category learning. Learn. Mem. *9*, 408–418.

Gluth, S., Sommer, T., Rieskamp, J., and Büchel, C. (2015). Effective Connectivity between Hippocampus and Ventromedial Prefrontal Cortex Controls Preferential Choices from Memory. Neuron *86*, 1078–1090.

Goldfarb, E.V., Chun, M.M., and Phelps, E.A. (2016). Memory-Guided Attention: Independent Contributions of the Hippocampus and Striatum. Neuron *89*, 317–324.

Grill-Spector, K., Kourtzi, Z., and Kanwisher, N. (2001). The lateral occipital complex and its role in object recognition. Vision Res. *41*, 1409–1422.

Groenewegen, H.J., Vermeulen-Van der Zee, E., te Kortschot, A., and Witter, M.P. (1987). Organization of the projections from the subiculum to the ventral striatum in the rat. A study using anterograde transport of Phaseolus vulgaris leucoagglutinin. Neuroscience *23*, 103–120.

Insausti, R., Juottonen, K., Soininen, H., Insausti, A.M., Partanen, K., Vainio, P., Laakso, M.P., and Pitkänen, A. (1998). MR volumetric analysis of the human entorhinal, perirhinal, and temporopolar cortices. AJNR Am. J. Neuroradiol. *19*, 659–671.

Ito, R., Robbins, T.W., Pennartz, C.M., and Everitt, B.J. (2008). Functional interaction between the hippocampus and nucleus accumbens shell is necessary for the acquisition of appetitive spatial context conditioning. J. Neurosci. *28*, 6950–6959.

Kelley, A.E., and Domesick, V.B. (1982). The distribution of the projection from the hippocampal formation to the nucleus accumbens in the rat: an antero-grade- and retrograde-horseradish peroxidase study. Neuroscience *7*, 2321–2335.

Kimchi, R. (1994). The role of wholistic/configural properties versus global properties in visual form perception. Perception *23*, 489–504.

Kleiner, M., Brainard, D., Pelli, D., Ingling, A., Murray, R., and Broussard, C. (2007). What's new in psychtoolbox-3. Perception *36*, 1–16.

Knowlton, B.J., Squire, L.R., and Gluck, M.A. (1994). Probabilistic classification learning in amnesia. Learn. Mem. *1*, 106–120.

Knowlton, B.J., Mangels, J.A., and Squire, L.R. (1996). A neostriatal habit learning system in humans. Science *273*, 1399–1402.

Kobayashi, S., and Schultz, W. (2008). Influence of reward delays on responses of dopamine neurons. J. Neurosci. *28*, 7837–7846.

Kourtzi, Z., and Kanwisher, N. (2000). Cortical regions involved in perceiving object shape. J. Neurosci. *20*, 3310–3318.

Kriegeskorte, N., Mur, M., and Bandettini, P. (2008). Representational similarity analysis - connecting the branches of systems neuroscience. Front. Syst. Neurosci. *2*, 4.

Kumaran, D., Summerfield, J.J., Hassabis, D., and Maguire, E.A. (2009). Tracking the emergence of conceptual knowledge during human decision making. Neuron *63*, 889–901.

Lamb, M.R., and Robertson, L.C. (1990). The effect of visual angle on global and local reaction times depends on the set of visual angles presented. Percept. Psychophys. *47*, 489–496.

Maddox, W.T., Ashby, F.G., and Bohil, C.J. (2003). Delayed feedback effects on rule-based and information-integration category learning. J. Exp. Psychol. Learn. Mem. Cogn. *29*, 650–662.

Marr, D. (1971). Simple memory: a theory for archicortex. Philos. Trans. R. Soc. Lond. B Biol. Sci. *262*, 23–81.

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

Melchers, K.G., Shanks, D.R., and Lachnit, H. (2008). Stimulus coding in human associative learning: flexible representations of parts and wholes. Behav. Processes *77*, 413–427.

Mumford, J.A., Davis, T., and Poldrack, R.A. (2014). The impact of study design on pattern estimation for single-trial multivariate pattern analysis. Neuroimage *103*, 130–138.

Murray, E.A., and Richmond, B.J. (2001). Role of perirhinal cortex in object perception, memory, and associations. Curr. Opin. Neurobiol. *11*, 188–193.

Murty, V.P., DuBrow, S., and Davachi, L. (2015). The simple act of choosing influences declarative memory. J. Neurosci. *35*, 6255–6264.

Navon, D. (1977). Forest before trees: the precedence of global features in visual perception. Cogn. Psychol. *9*, 353–383.

Navon, D., and Norman, J. (1983). Does global precedence really depend on visual angle? J. Exp. Psychol. Hum. Percept. Perform. *9*, 955–965.

Niv, Y., Daniel, R., Geana, A., Gershman, S.J., Leong, Y.C., Radulescu, A., and Wilson, R.C. (2015). Reinforcement learning in multidimensional environments relies on attention mechanisms. J. Neurosci. *35*, 8145–8157.

Norman, K.A., and O'Reilly, R.C. (2003). Modeling hippocampal and neocortical contributions to recognition memory: a complementary-learning-systems approach. Psychol. Rev. *110*, 611–646.

Norman-Haignere, S.V., McCarthy, G., Chun, M.M., and Turk-Browne, N.B. (2012). Category-selective background connectivity in ventral visual cortex. Cereb. Cortex *22*, 391–402.

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

O'Doherty, J., Dayan, P., Schultz, J., Deichmann, R., Friston, K., and Dolan, R.J. (2004). Dissociable roles of ventral and dorsal striatum in instrumental conditioning. Science *304*, 452–454.

Pearce, J.M. (1987). A model for stimulus generalization in Pavlovian conditioning. Psychol. Rev. *94*, 61–73.

Pearce, J.M. (1994). Similarity and discrimination: a selective review and a connectionist model. Psychol. Rev. *101*, 587–607.

Pelli, D.G. (1997). The VideoToolbox software for visual psychophysics: transforming numbers into movies. Spat. Vis. *10*, 437–442.

Pennartz, C.M.A., Ito, R., Verschure, P.F.M.J., Battaglia, F.P., and Robbins, T.W. (2011). The hippocampal-striatal axis in learning, prediction and goal-directed behavior. Trends Neurosci. *34*, 548–559.

Peters, J., and Büchel, C. (2010). Episodic future thinking reduces reward delay discounting through an enhancement of prefrontal-mediotemporal interactions. Neuron *66*, 138–148.

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

Power, J.D., Mitra, A., Laumann, T.O., Snyder, A.Z., Schlaggar, B.L., and Petersen, S.E. (2014). Methods to detect, characterize, and remove motion artifact in resting state fMRI. Neuroimage *84*, 320–341.

R Development Core Team. (2015). R: A language and environment for statistical computing (R Foundation for Statistical Computing).

Reber, P.J., Knowlton, B.J., and Squire, L.R. (1996). Dissociable properties of memory systems: differences in the flexibility of declarative and nondeclarative knowledge. Behav. Neurosci. *110*, 861–871.

Rescorla, R.A., and Wagner, A.R. (1972). A theory of Pavlovian conditioning: variations in the effectiveness of reinforcement and nonreinforcement. In Classical conditioning II: current research and theory, A.H. Black and W.F. Prokasy, eds. (Appleton-Century-Crofts), pp. 64–99.

Roelfsema, P.R., and van Ooyen, A. (2005). Attention-gated reinforcement learning of internal representations for classification. Neural Comput. *17*, 2176–2214.

Roelfsema, P.R., van Ooyen, A., and Watanabe, T. (2010). Perceptual learning rules based on reinforcers and attention. Trends Cogn. Sci. *14*, 64–71.

Rudy, J.W., and Sutherland, R.J. (1995). Configural association theory and the hippocampal formation: an appraisal and reconfiguration. Hippocampus *5*, 375–389.

Saffran, J.R., Aslin, R.N., and Newport, E.L. (1996). Statistical learning by 8-month-old infants. Science *274*, 1926–1928.

Saunders, R.C., and Weiskrantz, L. (1989). The effects of fornix transection and combined fornix transection, mammillary body lesions and hippocampal ablations on object-pair association memory in the rhesus monkey. Behav. Brain Res. *35*, 85–94.

Schapiro, A.C., Kustner, L.V., and Turk-Browne, N.B. (2012). Shaping of object representations in the human medial temporal lobe based on temporal regularities. Curr. Biol. *22*, 1622–1627.

Schapiro, A.C., Rogers, T.T., Cordova, N.I., Turk-Browne, N.B., and Botvinick, M.M. (2013). Neural representations of events arise from temporal community structure. Nat. Neurosci. *16*, 486–492.

Schönberg, T., Daw, N.D., Joel, D., and O'Doherty, J.P. (2007). Reinforcement learning signals in the human striatum distinguish learners from nonlearners during reward-based decision making. J. Neurosci. *27*, 12860–12867.

Schultz, W. (1998). Predictive reward signal of dopamine neurons. J. Neurophysiol. *80*, 1–27.

Shohamy, D., Myers, C.E., Onlaor, S., and Gluck, M.A. (2004). Role of the basal ganglia in category learning: how do patients with Parkinson's disease learn? Behav. Neurosci. *118*, 676–686.

Smith, S.M., Jenkinson, M., Woolrich, M.W., Beckmann, C.F., Behrens, T.E.J., Johansen-Berg, H., Bannister, P.R., De Luca, M., Drobnjak, I., Flitney, D.E., et al. (2004). Advances in functional and structural MR image analysis and implementation as FSL. Neuroimage *23* (*Suppl 1*), S208–S219.

Sutherland, R.J., and Rudy, J.W. (1989). Configural association theory: the role of the hippocampal formation in learning, memory, and amnesia. Psychobiology *17*, 129–144.

Sutton, R.S. (1984). Temporal credit assignment in reinforcement learning. PhD thesis (University of Massachusetts Amherst).

Sutton, R.S., and Barto, A.G. (1998). Reinforcement learning: an introduction (Cambridge, MA: The MIT Press).

Tambini, A., and Davachi, L. (2013). Persistence of hippocampal multivoxel patterns into postencoding rest is related to memory. Proc. Natl. Acad. Sci. USA *110*, 19591–19596.

Tompary, A., Duncan, K., and Davachi, L. (2015). Consolidation of Associative and Item Memory Is Related to Post-Encoding Functional Connectivity between the Ventral Tegmental Area and Different Medial Temporal Lobe Subregions during an Unrelated Task. J. Neurosci. *35*, 7326–7331.

Treves, A., and Rolls, E.T. (1992). Computational constraints suggest the need for two distinct input systems to the hippocampal CA3 network. Hippocampus *2*, 189–199.

Turk-Browne, N.B., Isola, P.J., Scholl, B.J., and Treat, T.A. (2008). Multidimensional visual statistical learning. J. Exp. Psychol. Learn. Mem. Cogn. *34*, 399–407.

Tustison, N.J., Cook, P.A., Klein, A., Song, G., Das, S.R., Duda, J.T., Kandel, B.M., van Strien, N., Stone, J.R., Gee, J.C., and Avants, B.B. (2014). Large-scale evaluation of ANTs and FreeSurfer cortical thickness measurements. Neuroimage *99*, 166–179.

van de Pol, M., and Wright, J. (2009). A simple method for distinguishing within- versus between-subject effects using mixed models. Anim. Behav. *77*, 753–758.

Van Essen, D.C., Anderson, C.H., and Felleman, D.J. (1992). Information processing in the primate visual system: an integrated systems perspective. Science *255*, 419–423.

Wimmer, G.E., and Shohamy, D. (2012). Preference by association: how memory mechanisms in the hippocampus bias decisions. Science *338*, 270–273.

Wimmer, G.E., Braun, E.K., Daw, N.D., and Shohamy, D. (2014). Episodic memory encoding interferes with reward learning and decreases striatal prediction errors. J. Neurosci. *34*, 14901–14912.

Yarkoni, T., Poldrack, R.A., Nichols, T.E., Van Essen, D.C., and Wager, T.D. (2011). Large-scale automated synthesis of human functional neuroimaging data. Nat. Methods *8*, 665–670.

CellPress

STAR★METHODS

KEY RESOURCES TABLE

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Software and Algorithms		
MATLAB R2015A	MathWorks	https://www.mathworks.com
Psychtoolbox-3	Brainard 1997; Pelli 1997; Kleiner et al., 2007	http://psychtoolbox.org/
FSL 5.0.1	FMRIB; Smith et al., 2004	https://fsl.fmrib.ox.ac.uk/fsl/fslwiki
AFNI 2011.12.21.0938	Cox, 1996;	https://afni.nimh.nih.gov
ANTs 2.1	stnava; Tustison et al., 2014	http://stnava.github.io/ANTs/
Neurosynth	Yarkoni et al., 2011	http://neurosynth.org/
R version 3.3.0	R Development Core Team, 2015	https://www.r-project.org
LME4	Bates et al., 2015	http://cran.r-project.org/web/packages/lme4/

CONTACT FOR REAGENT OR RESOURCE SHARING

Further information and requests for resources should be directed to and will be fulfilled by Katherine Duncan (duncan@psych.utoronto.ca).

EXPERIMENTAL MODEL AND SUBJECT DETAILS

We recruited thirty right-handed participants (twenty women, ten men) from the New York University and Columbia University communities with normal or corrected to normal vision. Two participants were removed from analyses due to early termination of their scanning sessions and two additional participants were removed because they exhibited both excessive motion (2 sd above group mean) and missed responses (> 10% & > 2 sd above group mean). The remaining twenty-six participants had a mean age of 23.7 with a range of 18-29. Informed consent was obtained at the beginning of the session and all experimental procedures were approved by the Institutional Review Board.

METHOD DETAILS

Behavioral Procedures
Participants performed the Inseparable and Separable Tasks in counterbalanced order while undergoing fMRI scanning. Both tasks were modifications of the "weather prediction" task (Knowlton et al., 1994, 1996; Poldrack et al., 2001; Shohamy et al., 2004). Participants learned to predict sunny or rainy outcomes based on abstract cues through trial-and-error. Each task included four unique cues which were combined into six pairs of cues. The two tasks differed, however, in how the cues could be used to predict the weather.

In the Inseparable Task, which was designed to selectively tap configural learning, each of six combinations of cues was probabilistically associated with weather outcomes, but the cues were paired such that each cue was not predictive of weather outcomes. In the Separable Task, each of the four cues had an independent relationship with weather outcomes, allowing participants to use elemental learning. The structure of the Separable Task was also conducive to configural learning, more so than the structure of the standard Weather Prediction task, because: (1) cues were always presented in pairs, never alone; and (2) only six combinations were used, rather than 16. Immediately before performing each task, participants were told about the task structure for the sake of transparency about the design. In the Inseparable Task, they were told that individual images were not predictive of the weather and that a good strategy would be to learn which combinations are strong predictors; in the Separable Task, they were told that each image independently predicted the weather outcomes and that a good strategy would be to learn which images are strong predictors.

Each task was broken into four scanned runs that were 60 trials long. Visual stimuli were presented using Psychtoolbox (Brainard, 1997; Kleiner et al., 2007; Pelli, 1997) and were projected onto a screen that was viewed through a mirror attached to the participant's head coil. Trials began with the presentation of a pair of cues for 2 s, during which participants were asked to predict whether it would be sunny (index finger key) or rainy (middle finger key). Their selections were indicated on the display and, after a 0.25 s delay, feedback appeared at the top of the display for 2 s. A 1.75 s fixation-cross separated trials. After completing each task, participants were presented with the 6 combinations and 4 cues from the preceding task one at a time. They were asked to rate the likelihood of weather outcomes based on these cues, on a continuous scale from "always sunny" to "always rainy."

MRI Methods

A 3T Siemens Allegra MRI system with a whole-head coil was used for all scans. Functional data was collected using an echo-planar pulse (EPI) sequence (TR = 1500ms, TE = 17ms, FOV = 240 × 192mm, 34 interleaved slices, 3 × 3 x 3mm voxel size, flip angle = 76 degrees) with oblique coronal slices aligned perpendicular to the hippocampal long axis. The field of view was reduced in the phase encode direction and the TE was minimized to reduce the total read out time and, thus, minimize distortions and artifacts. All functional sequences were 6.3 minutes long (252 volumes). A T1-weighted high-resolution MPRAGE (magnetization-prepared rapid-acquisition gradient echo) sequence (1x1x1 voxel size, 176 sagital slices) was also acquired to obtain full brain coverage. Finally, a field map sequence was collected and used to correct for distortions in EPI images using in-house scripts that apply the correction in k-space.

Preprocessing of functional data was conducted with FSL (FMRIB Software Library; online at https://fsl.fmrib.ox.ac.uk/fsl/fslwiki; Smith et al., 2004), AFNI (Analysis of Functional NeuroImages; online at https://afni.nimh.nih.gov; Cox, 1996), and ANTS (Advanced Normalization Tools; online at http://stnava.github.io/ANTs/; Tustison et al., 2014). The first 8 volumes (12 s) were discarded to allow for signal normalization and differences in slice acquisition timing were corrected for using sinc interpolation (FSL slicetimer). Next, registration-targets for the session were generated by aligning and averaging the first volume of each EPI scan (before T1 normalization). Each EPI volume and the T1-weighted anatomical scans were aligned to the registration-target using rigid transformations optimized using AFNIs 3dvolreg and ANTs, respectively. A priori anatomical subcortical ROIs were automatically segmented using FSL's FIRST segmentation tool, followed by visual inspection and correction. The hippocampus was additionally segmented into anterior and posterior portions by partitioning it at the posterior extent of the uncus. To perform additional exploratory analyses, we (1) hand segmented a perirhinal cortex ROI using established procedures (Insausti et al., 1998) and (2) generated a lateral occipital complex ROI using the Neurosynth tool (Yarkoni et al., 2011) to produce an anatomically constrained meta-analysis identifying voxels that are sensitive to objects.

Eye Tracking Procedures

Eye position was monitored during the scanning session with an infrared videographic camera equipped with a telephoto lens (Eyelink 1000, SR research Ltd., Kanata, Ontario, Canada) at a minimum of 250 Hz. Nine-point calibrations were performed and validated at the beginning of the session. After resampling all data to 250 Hz, eye blinks and periods of lost signal (+/− 5 samples) were removed from data. Full trials were removed if less than 70% of the samples could be labeled as fixations and participants were included in the analysis if at least 50% of the trials in both tasks reached this cutoff (17/26 participants). Symmetric looking time was calculated during the response period (before response) as 1 minus (the absolute difference in time spent looking at each cue divided by the sum of time spent looking at either cue).

QUANTIFICATION AND STATISTICAL ANALYSIS

Learning Success Model

Logistic regressions predicting optimal choice (0 versus 1) with trial number were run separately for each participant. Final learning was estimated using the resulting regression equations as the likelihood of making an optimal choice on the final trial. Group performance was assessed using a mixed-effects GLM (lme4 package: Bates et al., 2015) in the R programming language, optimized with restricted maximum likelihood. The model predicted optimal choice (0 versus 1) with trial number and task and used a logistic linking function. It additionally included random effects terms for all coefficients across participants.

Reinforcement Learning Model

We used a reinforcement learning (RL) model (Sutton and Barto, 1998) to derive tailored estimates of the degree to which participants' choices were driven by elemental or configural learning in each task. This model leveraged participants' reinforcement history and choice patterns to separately estimate the choice values associated with each configuration and each element. For example, consider the following sequence of outcomes: AB- > Sun; AC- > Rain; BD- > Rain (Figure 2A). If then presented with 'AB', a participant making choices using recent experiences with the configuration 'AB' would be more likely to select 'Sun', whereas a participant making choices using recent experiences with the elements 'A' and 'B' would be more likely to select 'Rain'.

The model assumes participants learn the value, Q, of choosing Sun or Rain given a particular state, S. States are determined by how the 'stimulus' is identified, namely by either the configuration of cues presented on the trial, Sc, or by each element, Se_1, Se_2. The value of the weather predicted on the current trial, W_p, is updated based on the difference between the expected value of the prediction and the feedback received (correct = 1; incorrect = 0). This difference is called the prediction error, δ, and is calculated separately for each presented state:

$$\delta_{ct} = r_t - Qt - 1(W_p \,|Sc)$$

$$\delta_{e1t} = r_t - Qt - 1(W_p \,|Se1)$$

CellPress

$$\delta_{e2t} = r_t - Qt - 1(W_p \mid Se2).$$

On each trial, the observed states (one configuration and two elements) are updated in light of their prediction error. (States not presented remain unchanged.) The degree to which these prediction errors update the value of the predicted weather depends on the learning rate parameter, α:

$$Qt(W_p|Sc) = Qt - 1(W_p|Sc) + \alpha_c \delta_{ct-1}$$

$$Qt(W_p|Se_1) = Qt - 1(W_p|Se_1) + \alpha_e \delta_{e1t-1}$$

$$Qt(W_p|Se_2) = Qt - 1(W_p|Se_2) + \alpha_e \delta_{e2t-1}.$$

As α approaches 1, the value of the predicted weather will approach the most recently received feedback. As α approaches 0, the value of the predicted weather will be minimally updated by feedback. We fit separate α parameters to participants' behavior for configural and element states (2 total), restricting their range to be between 0 and 1. The value of the non-predicted weather was not updated.

The state-contingent weather prediction values were combined to compute the probability, P, of predicting each outcome using a softmax (logistic) choice rule:

$$D_{ct} = (Q(Sun|Sc) - Q(Rain|Sc))$$

$$D_{e1t} = (Q(Sun|Se_1) - Q(Rain|Se_1))$$

$$D_{e2t} = (Q(Sun|Se_2) - Q(Rain|Se_2))$$

$$P(Pred_{sun,t}) = \frac{1}{1 + \exp(-(\beta c * D_{ct} + \beta e_1 * D_{e1t} + \beta e_2 * D_{e2t}))}.$$

Here, the inverse temperature parameters, β, control how closely the differences in weather prediction values, D, govern choices. Importantly, separate β parameters were fit to participants' behavior for each configural and element state (10 total). This allowed us to capture participants' tendency to weight specific elements or configurations more in their decisions. This made the model flexible enough to accommodate participants who only learned the optimal weather predictions for some configurations or elements.

We estimated the 12 free parameters (α_c, α_e, β_{AB}, β_{AC}, β_{AD}, β_{BC}, β_{BD}, β_{CD}, β_A, β_B, β_C, β_D) for each participant by minimizing the sum of the negative log likelihoods of choices given the estimated probability, P, of each choice using constrained nonlinear optimization (fmincon, MATLAB). To avoid local optima, we repeated the search five times with each parameter starting at a random point. Model comparison (Table 1) verifies the utility of including multiple alpha and beta terms in the model. We also conducted likelihood ratio tests on each participant (Figure S4) to confirm that these model decisions improved fit at the individual level. Notably, including separate beta parameters for each configuration/element, significantly improved model fit in nearly all participants compared to a model that assumed participants weighted all configuration and all elements equally (1 configural and 1 elemental beta).

We used this model to estimate the *Configural Choice Index*, which reflects the relative use of configurally versus elementally learned values on a trial-by-trial basis by comparing the configural and elemental Q values. Specifically, we measured the difference between the predicted and non-predicted weather Q values, separately for configural states, D_{cpt}, and elemental states, D_{ept}.

$$D_{cpt} = (Q(Pred|Sc) - Q(NonPred|Sc))$$

$$D_{ept} = w_1(Q(Pred|Se_1) - Q(NonPred|Se_1)) + w_2(Q(Pred|Se_2) - Q(NonPred|Se_2))$$

$$w_1 = \frac{\beta e_1}{\beta e_1 + \beta e_2} \quad w_2 = \frac{\beta e_2}{\beta e_1 + \beta e_2}.$$

The two elemental state Q values were averaged by weighting each according to the corresponding β parameter that was fit to the individual participant. This takes into account individual differences in the extent to which each cue was weighted in the decision

making process (e.g., making choices using only the most predictive cues). The choice-type score was then calculated by subtracting D_{ept} from D_{cpt}. We used the average α_c and α_e across participants to generate this trial-by-trial Q-value. This approach reduces over-fitting and the noisiness in individual participants estimates (Daw et al., 2006; Schönberg et al., 2007) and ensured that individual differences in learning were not driven by differences in the model specifications.

We also used this model to obtain a single estimate of the degree to which each participant engaged in elemental or configural learning in each task (*Learning Style Score*) by comparing the proportion of variance (pseudo r^2: LLE_{model}/LLE_{chance}) explained by reduced models that removed either all configural states or all elemental states. The variance explained by the elemental only model was subtracted from the that explained by the configural only model to produce a continuous estimate of each participant's relative use of configural versus elemental learning. This score was used to identify how individual differences in learning style were related to other behavioral and neural variables.

In the Inseparable Task, the full model provided a better fit for 25/26 participants over a model that only had elemental values (likelihood ratio test, p < 0.05), reflecting the importance of configurations in the tasks' design. By contrast, adding elemental choice values uniquely explained choice variance in 8/26 participants as compared to a configural only model (likelihood ratio test, p < 0.05), indicating that some participants used elemental learning despite the task being designed to encourage configural learning. Combining these tendencies into the *Learning Style Score* revealed that all but one participant in the Inseparable Task showed greater reliance on configural learning (Figure 2B).

A mixture of learning styles was found in the Separable Task. Including elemental choice values significantly improved the model fit in 18/26 subjects, while including configural choice values significantly improved model fit in 15/26 participants (likelihood ratio test, p < 0.05). Distilling these differences into the *Learning Style Score* revealed that roughly half the participants relied more on configural learning while the other half relied more on elemental learning.

FMRI Statistical Analysis
Configural Choice Analysis
To assess how BOLD activations in different learning and memory systems related to the types of choices that people made on a trial-by-trial basis, we first estimated BOLD responses to individual trials. We generated these 'beta-series' with a voxel-wise GLM run on concatenated runs in native space. The GLM contained separate regressors for the onset of each choice period (HRF convolved impulse response) to obtain single trial estimates. It also contained regressors for correct and incorrect feedback (HRF convolved impulse response at onset of outcome) and nuisance parameters (motion and global signal, see below). The resulting trial-specific beta estimates were then averaged across voxels within anatomically defined ROIs (aHip, NAc, D/LS, PRC, LOC) and entered into a mixed effects GLM to predict whether the corresponding choices were more consistent with configural or elemental learning (*Configural Choice Index* described above). To assess whether the relationship between BOLD activation and choice was modulated by individual differences in learning, we additionally ran a model that included individual differences in Learning Style Score and their interaction with each beta-series. In all models, each fixed effect, apart Learning Style Score, was also included as a random effect, grouped by participant. Predictors included in all linear mixed-effects models were mean centered within subject to disentangle within and between subject effects (van de Pol and Wright, 2009). Significance tests for all linear mixed-effects models were performed using the Kenward-Roger correction for degrees of freedom.
Staggered Beta-Series Analyses
We also assessed how these choice activations were related to outcome activations on a trial-by-trial basis. We used another voxel-wise GLM to obtain trial-specific estimates of outcome activations. The model contained separate regressors for the onset of each outcome phase (HRF convolved impulse response) along with regressors for the onset of the choice period (HRF convolved impulse response), with trials binned according to optimal choices, non-optimal choices, and no optimal choice possible along with nuisance parameters (motion and global signal, see below). The trial-specific outcome betas were then averaged across voxels within ROIs (aHip, NAc, D/LS). It should be noted that the time separating choice and outcome trial phases was not jittered to mitigate previously identified relationships between delayed feedback and hippocampal involvement in reinforcement learning (Foerde et al., 2013; Foerde and Shohamy, 2011). The cost of this decision, however, is that choice and outcome can only be disentangled via the short intervening delay and the inconsistent relationship between choice types and outcome types (e.g., we included two configurations with unpredictable weather outcomes).

We first used these outcome betas along with choice phase betas to assess how activation within one region at the time of choice was related to activations in other regions at the time of outcomes (choice→outcome connectivity). Specifically, we predicted outcome activations in the NAc with choice-phase activations in the aHip and D/LS using a mixed GLM. Adopting the logic of Granger-Causality methods, we also included NAc choice responses to control for temporal autocorrelation within regions. A model including betas from both tasks was used to estimate the effect of task by including the task and the interaction between task and each beta-series (aHip, D/LS, and NAc choice responses). Separate models were also run using beta-series from each task to estimate the effect of configural versus elemental learning on inter-region coupling within tasks and how it interacted with individual differences Learning Style Score. All fixed effects, apart from Learning Style Score, were also included as random effects, grouped by participant.

CellPress

To better understand how outcome impacts learning, we next assessed how activation within one region at the time of outcome was related to subsequent activations in other regions on the next trial in which participants saw the configuration in question (outcome→subsequent choice connectivity). Specifically, we predicted choice activations in the aHip with the last relevant outcome-response in the NAc and D/LS using a mixed GLM. As with the choice→outcome analysis, here we also included aHip outcome responses to control for the temporal autocorrelation within region. A model including betas from both tasks was used to estimate the effect of task by including the task and the interaction between task and each beta-series. Separate models were also run using betas from each task to estimate the effect of configural versus elemental learning on inter-region coupling within tasks and how it interacted with individual differences Learning Style Score. All fixed effects, apart from Learning Style Score, were also included as random effects, grouped by participant.

To further investigate the directionality of results obtained from the above models, we ran a series of control analyses (Tables S1–S3). Specifically, we reversed the regions (e.g., predicting aHip outcome with NAc choice) or measured functional connectivity within the same phase of a trial (choice→choice or outcome→outcome) in control models.

Background Connectivity Analysis

Background functional connectivity between ROIs was measured by correlating low-frequency fluctuations in BOLD responses across regions. Importantly, this 'background connectivity' procedure takes steps to remove the contributions of trial-evoked responses so that this resulting measure reflects shifts in functional connectivity that are related to extended states of cognitive processing. This approach has previously been used to study sustained attention (Al-Aidroos et al., 2012; Norman-Haignere et al., 2012) and episodic memory processes (Duncan et al., 2014; Tompary et al., 2015) and is ideal for studying the relationship between functional connectivity and learning, as the process of learning by definition occurs over extended periods of time.

We first used a two-step procedure, similar to that reported in Duncan et al. (2014) and Tompary et al. (2015), to reduce the influence of signals that may artificially inflate estimates of functional connectivity. First, a voxel-wise GLM was used to regress out trial-evoked responses and nuisance parameters for each run. Separate regressors were included for choice (2 s boxcar Regressors: optimal response; non-optimal response; no optimal response possible) and outcome (1.5 s boxcar Regressors: correct feedback; incorrect feedback) phases. Regressors were generated by convolving 'active time points' with the canonical HRF, its temporal derivative, and it dispersion derivative, as supplied by AFNI. This basis set flexibly captures variability in the amplitude, initiation and duration of BOLD responses in each voxel. Nuisance parameters were included by adapting the guidelines set in Power et al. (2014). These include the Voltara expansion of the six motion parameters estimated in motion correction (motion, $motion^2$, $motion_{t-1}$, $motion_{t-1}^2$) and their derivatives, along with global signal. In contrast to resting state scanning, the global signal in task scans can contain important functional information. To account for this difference, we used a similar method to that described in Behzadi et al. (2007) to (1) restrict contributing voxels to those that were not modulated by the task (omnibus F-stat < 1.13, p > 0.2) and to (2) use the first 5 principle components obtained from a temporal principal components analysis rather than the mean across all voxels. The latter step allowed us to capture the heterogeneous consequences that motion and physiological factors have across voxels. Second, we then band-pass filtered the residuals of this model, leaving only signal between 0.009 and 0.08 Hz. This band of frequencies is the highest contributor to inter-region correlations during resting state analyses (Cordes et al., 2001). Critically, this band is also outside the task frequency (.167 Hz) and, thus, also filters out responses that were consistently elicited by choice or outcome phases.

Thus, we filtered out trial-related BOLD responses using a flexible basis set to model the HRF (canonical HRF + temporal derivative + dispersion derivative) followed by a frequency filtering procedure which removes any signal which fluctuates at the task frequency. This second stage could be particularly useful for removing trial-related BOLD responses which do not conform to canonical HRF models. It should be noted, however, that alternative basis sets (e.g., sine waves or tent functions) do not make assumption about HRF shape and so could efficiently accomplish both goals in a single step. We chose this two-step procedure instead to capitalize on both canonical HRF basis sets' capacity to capture the most plausible BOLD responses and frequency-filtering's capacity to more flexibly remove remaining signals which systematically occur at different phases of a trial. Although the outcomes of these two approaches should be similar, future work directly comparing them would benefit the field.

For each anatomical ROI, we extracted mean timeseries from the filtered residuals. These were z-scored and concatenated across runs. We then used mixed GLMs to estimate the functional connectivity in each task. Models predicted NAc timeseries with aHip and D/LS timeseries. A model including time points from both tasks was used to estimate the effect of the task by including the task and the interaction between task and each timeseries. Separate models were also run using time points from each task to estimate the effect of configural versus elemental learning on functional connectivity. These models included individual differences in Learning Style Score and their interaction with each timeseries. All fixed effects, apart Learning Style Score, were also included as random effects, grouped by participant.

Representational Similarity Analysis (RSA)

We used RSA to assess whether hippocampal choice activity patterns better reflected configurations of cues or the elements which comprise them (Kriegeskorte et al., 2008). Trial-specific choice activations (t-stats) were estimated in each voxel within anatomically defined ROIs (see above), but inclusion in the analysis was restricted to voxels which showed a reliable positive or negative activation

across trials (p < 0.05; Tambini and Davachi, 2013). We then computed similarity by correlating activation vectors across pairs of trials for each ROI (restricted to separate runs; Mumford et al., 2014). We averaged Fisher z transformed Pearson correlation coefficients across all trial pairs which compared the same pair of configurations (e.g., all trials comparing AB patterns to BC patterns) to construct a similarity matrix. We then correlated these matrices with different models (Figure 6A). The *Configural Model* predicted that patterns would be more similar for trial pairs which shared the same configuration as compared to those which did not, but was blind to the elements. By contrast, the *Elemental Model*, predicted that similarity should linearly scale according the number of shared elements (0, 1, or 2).