Article

Neuron

Neurocomputational Dynamics of Sequence Learning

Highlights

- Human subjects learn to detect patterns in sequences of images while undergoing fMRI
- Behavior is well explained by a Bayesian pattern-learning model
- Distinct networks track uncertainty about the pattern and the predicted image

Authors

Arkady Konovalov, Ian Krajbich

Correspondence

krajbich.1@osu.edu

In Brief

Konovalov and Krajbich present a computational model for how the human brain detects deterministic patterns in its environment. Using fMRI, they identify distinct brain networks that track uncertainty about the specific temporal prediction and the underlying pattern.

Konovalov & Krajbich, 2018, Neuron *98*, 1282–1293
June 27, 2018
https://doi.org/10.1016/j.neuron.2018.05.013

CellPress

CellPress

Neuron
Article

Neurocomputational Dynamics of Sequence Learning

Arkady Konovalov[1,2] and Ian Krajbich[2,3,4,*]
[1]Laboratory for Social and Neural Systems Research, Department of Economics, University of Zurich, 8006 Zurich, Switzerland
[2]Department of Economics, The Ohio State University, 1945 North High Street, 410 Arps Hall, Columbus, OH 43210, USA
[3]Department of Psychology, The Ohio State University, 1827 Neil Avenue, 200E Lazenby Hall, Columbus, OH 43210, USA
[4]Lead Contact
*Correspondence: krajbich.1@osu.edu
https://doi.org/10.1016/j.neuron.2018.05.013

SUMMARY

The brain is often able to learn complex structures of the environment using a very limited amount of evidence, which is crucial for model-based planning and sequential prediction. However, little is known about the neurocomputational mechanisms of deterministic sequential prediction, as prior work has primarily focused on stochastic transition structures. Here we find that human subjects' beliefs about a sequence of states, captured by reaction times, are well explained by a Bayesian pattern-learning model that tracks beliefs about both the current state and the underlying structure of the environment, taking into account prior beliefs about possible patterns in the sequence. Using functional magnetic resonance imaging, we find distinct neural signatures of uncertainty computations on both levels. These results support the hypothesis that structure learning in the brain employs Bayesian inference.

INTRODUCTION

One of the big questions of model-based learning (Daw et al., 2011) is how people learn and update the structure of the decision environment (i.e., a model, a mental map, or a transition rule). Learning contingencies between one's actions and "states of the world" can be beneficial in achieving specific goals or earning higher rewards (Kool et al., 2016). Important examples of such contingencies are sequential events, which one often needs to memorize and retrieve daily (e.g., a sequence of turns on a road or steps in a recipe).

While there has been much research into the neural representations of sequences and probabilistic reinforcement learning, there has been a lack of research on the neurocomputational mechanisms of deterministic sequence learning (i.e., learning of patterns in sequences of states), where the decision-maker can infer the underlying structure and use this knowledge to make better predictions. In particular, it is unclear how the brain distinguishes deterministic patterns from random sequences. This problem is fundamentally different from probabilistic learning, where one knows that two states are linked; the issue is just estimating the strength of that link (i.e., the probability). While many studies in behavioral psychology have explored sequence learning using the serial reaction time (SRT) task, comparing deterministic and probabilistic sequences (Martini et al., 2013; Shanks et al., 2003; Vandenberghe et al., 2006; Wilkinson and Shanks, 2004), or implicit and explicit learning (Jiménez et al., 2006), or studying the impact of attention on sequence learning (Jiménez and Mendez, 1999; Shanks et al., 2005), these studies lack a unified model-based approach that could explain sequence learning as a results of an optimal (Bayesian) inference process (Sebastiani et al., 2000). Importantly, while many studies investigate sequences learned over many training blocks, we are interested in how subjects detect the presence of a predictable sequence.

Evidence from many cognitive and neurobiological studies suggests that the brain employs a number of mechanisms that provide neural representations of sequences, including but not limited to learning of transition rules, chunking, and more complex tools such as learning of nested tree structures for understanding language (Dehaene et al., 2015). Previous studies have identified a large number of brain regions potentially involved in sequence learning: the hippocampus (FitzGerald et al., 2017; Gheysen et al., 2010; Kumaran and Maguire, 2006b, 2006a, 2007; Schendan et al., 2003; Terada et al., 2017; Van Opstal et al., 2008; Yang and Li, 2012) (which has been repeatedly shown to play the key role in encoding and retrieving sequential structures), inferior frontal gyrus (Doeller et al., 2006; van der Graaf et al., 2006; Huettel et al., 2002) (IFG), insula (van der Graaf et al., 2006; Huettel et al., 2002; Kalm et al., 2013; Kumaran and Maguire, 2006a; Yang and Li, 2012), caudate (Huettel et al., 2002; Kalm et al., 2013; Kumaran and Maguire, 2006a; Schendan et al., 2003; Schwarb and Schumacher, 2009), putamen (Huettel et al., 2002; Lungu et al., 2014; Schwarb and Schumacher, 2009; Seidler et al., 2005; Werheid et al., 2003; Wymbs et al., 2012), anterior and posterior cingulate cortex (van der Graaf et al., 2006; Seidler et al., 2005; Werheid et al., 2003), ventral striatum (Doeller et al., 2006), intraparietal sulcus (Kumaran and Maguire, 2006a; Wymbs et al., 2012) (IPS), and precuneus (Lungu et al., 2014; Yang and Li, 2012); however, the exact nature and dynamics of computations and function of each specific region remain unclear.

At the same time, a parallel literature on reinforcement learning has used computational models to explain how the brain learns and updates values in reward-based environments (O'Doherty et al., 2007). In this paradigm, a decision-maker can sometimes employ a "model-based" learning strategy that learns and uses the structure of the decision environment to plan reward-maximizing actions (Beierholm et al., 2011; Daw et al., 2011; Doll et al., 2012, 2015; Gläscher et al., 2010; Huys et al., 2012;

Lee et al., 2014). Previous research in this area has primarily focused on stochastic transition models (Bornstein and Daw, 2012, 2013; Daw et al., 2011; Gläscher et al., 2010; Lee et al., 2014) (typical for Markov decision processes). In this type of model, the decision-maker updates their beliefs about coming states using a transition matrix that computes conditional probabilities of states given the previously observed state. On the other hand, a number of recent studies have suggested that the brain can also employ hidden structure learning using a space of causal structures and a set of prior beliefs, which are updated in a Bayesian fashion once some evidence toward one or the other structure is collected (Gershman and Niv, 2010, 2015; Gershman et al., 2015). The neural underpinnings of structure learning remain an open question, although recent evidence suggests that the hippocampus (Aggleton et al., 2007) and orbitofrontal cortex (Wilson et al., 2014) (OFC) play central roles, with the hippocampus engaging in structure encoding and the OFC serving as a "repository" for structure representations. The exact role of these and other regions involved in structure learning (such as regions encoding state prediction errors, SPEs), however, is not yet fully understood.

Here we combine these two approaches with an application to sequence learning. We are interested in whether stochastic transition models or deterministic structure models provide a better explanation for sequence learning, and how the best model might be implemented in the brain. To investigate these questions, we use a simple modified SRT task where subjects observe sequences of slowly revealed images and are incentivized to identify each image as quickly as possible. If the subjects detect patterns, they can respond before the images are fully revealed. This design allows us to measure learning using a continuous reaction time (RT) measure, without the need to collect additional measures (e.g., explicit confidence ratings). Across sequences we vary whether there is a pattern, and if there is, the length of the pattern. Varying the length of the patterns is important, since it helps to ensure that drops in RT and changes in brain activity are due to learning and not simply due to image number in the sequence.

With this design, we find that subjects detect patterns more quickly than can be explained by the standard stochastic transition model. Instead, learning (represented by RTs) is best explained by a Bayesian pattern-learning model that assumes prior beliefs about types of structures, or patterns, that the decision-maker can encounter. In this model, subjects have beliefs about the identity of the next state, as well as the structure of the environment (what kind of pattern, if any). From these beliefs, we construct state- and structure-level measures of uncertainty (entropy) and study which regions of the brain track each of them.

In summary, we provide a detailed model-based account of sequence learning in the human brain, identifying two distinct networks that separately encode uncertainty about the state and uncertainty about the underlying structure.

RESULTS

Behavioral Results

We carried out an fMRI experiment, where adult human subjects performed a modified SRT task (Figure 1). As in a standard SRT task, prior to the experiment, subjects learned fixed one-to-one correspondence between three fingers and three images (a face, a hand, and a landscape). During the experiment, we showed subjects sequences of these images and asked them to respond to each image with the correct button press.

We added two important modifications to the standard SRT design (STAR Methods). First, we incentivized both accuracy and RT with monetary rewards, with higher reward for accuracy. We aimed (successfully) to achieve maximal accuracy: subjects pressed the correct button in 97% of trials, while still responding to the RT incentive. Moreover, subjects correctly guessed the final image in 96% of sequences with patterns. This means that it is unlikely that subjects were guessing, which benefits our belief updating analyses, though it does preclude any analyses relating accuracy to neural activity or comparing correct and error trials. Second, we presented not the static images themselves but animations that smoothly transformed from a scramble of all images to a single image within 3 s (Figure 1). We did this so that subjects' RTs would change substantially once they discovered a pattern (in a typical SRT experiment, the difference between full uncertainty and no uncertainty does not exceed 30 ms [Bornstein and Daw, 2012], while in our experiment it exceeded 1 s).

We organized the images into 50 short sequences of 12 images each and informed subjects that some of the sequences they could encounter in the experiment would have repeating patterns. We created 30 sequences with repeated patterns of length 2, 3, 4, and 6, and 20 sequences that had no apparent patterns (conditional probability of each image given previous history was close to 0.33); each subject experienced the same set of sequences. We did not use any cues to indicate the sequence type and presented all sequences in a pseudorandom order (see STAR Methods).

As we expected, subjects responded faster to images in sequences with repeated patterns (1.05 s versus 1.4 s, $t\,[24] = -17$, $p < 0.001$), with RTs decreasing rapidly toward the end of the sequence (Figure 2A), and shorter patterns exhibiting earlier decreases in RT. We sought to identify a belief-updating model that would explain both the difference between the two conditions and the rate of the RT decrease in sequences with patterns.

Computational Modeling

A common way to model beliefs about states of the world in a learning task is with a Markov transition matrix, which reflects conditional beliefs about future states based on previous history (Bornstein and Daw, 2012, 2013; Konovalov and Krajbich, 2016; Otto et al., 2013) (typically, of the first order, where the conditional distribution depends only on one previously observed period). In a sequence with a repeating pattern, the posterior state distribution can depend on a longer history of states, so we used k-th order transition matrices, which are updated using a Dirichlet posterior (see STAR Methods). We use this model primarily as a benchmark, as it is considered a good approximation of the belief-updating process (Otto et al., 2013). We found the third-order matrix to be the best fitting model, meaning that the probability of seeing a particular image at a given point in time is based on the identities of the past three images. These

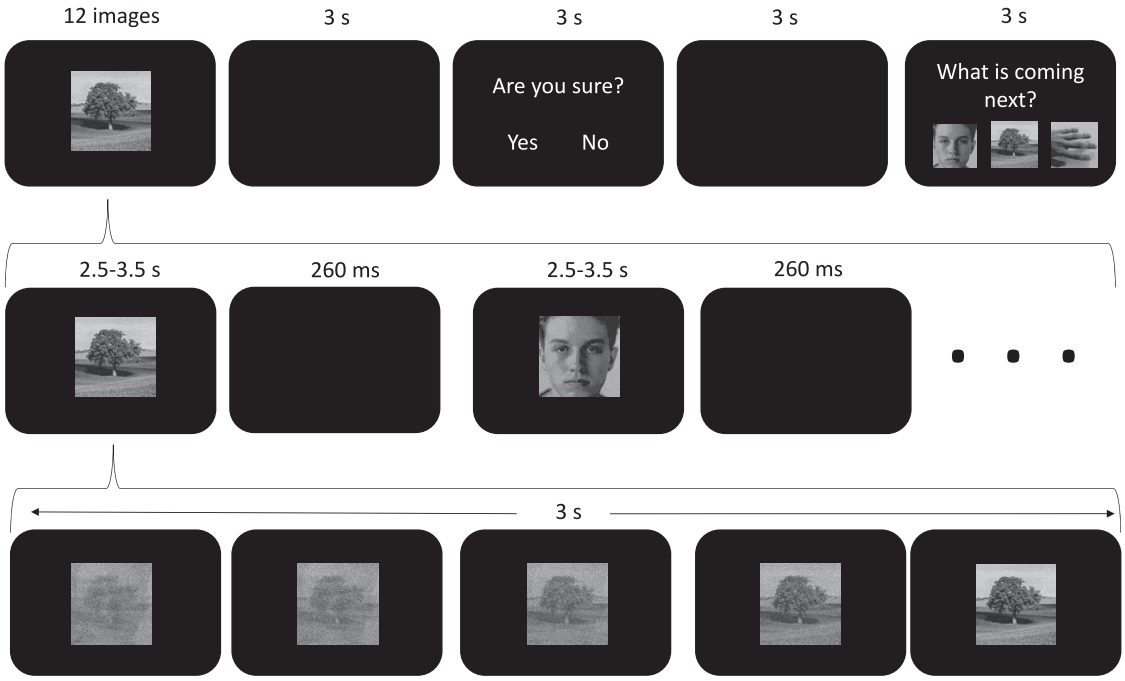

Figure 1. Serial Reaction Time Task

(Top row) Subjects observed 50 sequences (30 with repeating patterns of length 2–6 and 20 pseudorandomized) with 12 images each, in the fMRI scanner. At the end of each sequence, subjects indicated whether they knew the next image in the sequence and made a guess by pressing the corresponding button. (Middle row) Images (a hand, a face, and a landscape) were presented one at a time, as an animation that lasted for 3 s, plus or minus random jitter (in the case of positive jitter the image remained static after 3 s), regardless of response, and were separated by 260 ms of black screen. (Bottom row) Each animation started as a scramble of all three images with added Gaussian noise, with one image gradually emerging over time and the noise rate decreasing. If a subject pressed a correct key, a white frame appeared around the image until the end of the animation.

probabilities are used to calculate the Shannon entropy of beliefs $p(\cdot)$ about states a_i given observed history h,

$$E_a = -\sum_i p(a_i|h)\log p(a_i|h),$$

and surprise,

$$S_a = -\log\left(p(a_i|h)\right),$$

which we use as predictors of RTs. However, this model does not support some qualitative characteristics of the data: for instance, it predicts rapid one-trial drops in RTs when the observer learns the pattern (Figure S1), while the data show more gradual decline over two to three trials (we also explored a version of the model with more gradual belief updating, but it demonstrated slow learning and provided a worse fit to the data; see STAR Methods). To confirm that the gradual decline in RTs was not just a result of averaging across subjects and pattern types, we also examined the individual subject data. We found no evidence for rapid one-shot learning (see Figure S2).

We improved predictions of the transition matrix model by implementing a Bayesian pattern-learning model that assumes that subjects update both beliefs about the states of the world (images) and possible patterns (i.e., structures) (STAR Methods).

The model assumes that all sequences can be classified into seven categories, or "structures": patterns of length 1 to 6 (for instance, length 2, "hand-face-hand-face…," or length 3, "hand-face-landscape-hand-face-landscape…"), and "no pattern" (i.e., randomly generated sequences with no discernible pattern). To decrease the degrees of freedom, we assumed that subjects had common prior beliefs about the distribution of these structures before entering the experiment. As a robustness check, we also estimated a model in which these priors correspond to the real frequencies of the seven possible structures in the experiment, and the results were similar. For each possible structure, the subject estimates its likelihood given the observed history. For instance, after a history of "hand-face-landscape-face," the subject would assign probability 0 to patterns of length 1–3 but other non-zero probabilities to longer patterns and no pattern. The model assumes that subjects then use these posterior structure beliefs to calculate the probability that the next image is a face, hand, or landscape.

In other words, subjects calculate beliefs about each specific state coming next in the sequence as a compound probability that takes into account both the probability of a specific structure given the observed history and the probability of each state given that structure. Here the beliefs about the underlying structure represent "structure learning," while updating beliefs about the states of the world using these structural beliefs represents inference based on the previously learned structure.

As with the transition matrix model, we calculated state-level entropy for these beliefs, as well as the Bayesian surprise

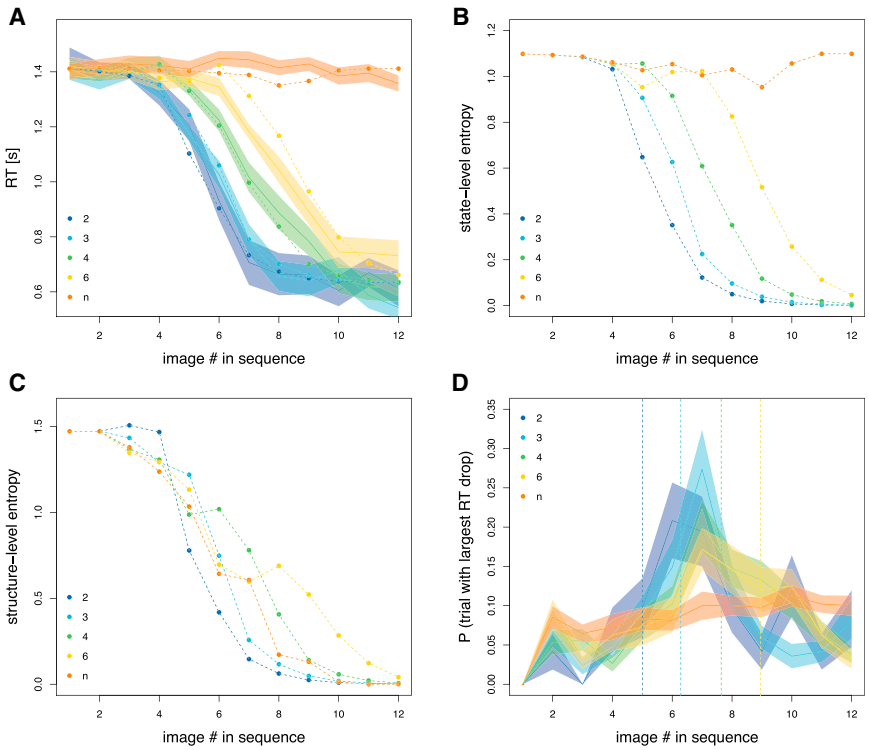

Figure 2. Behavioral Model Fits

Solid lines show data, and dotted lines show model simulations based on the best-fitting parameters. Colored areas denote SE clustered at the subject level. Data are split by the pattern length (with "n" denoting no pattern).

(A) Reaction times in seconds as a function of image number in a sequence.

(B) Predicted dynamics of state-level entropy.

(C) Predicted dynamics of structure-level entropy.

(D) Probability that the current trial has the largest drop in RT, compared to the previous trial, within the sequence. The vertical dotted lines are the single-peak predictions of the model. See also Figure S1.

produced by the model, to predict subjects' RTs. The pattern model explained the data better than the third-order Markov model, both qualitatively (Figure 2A) and quantitatively (Table S1), taking into account the increased number of free parameters (6 for 7 prior beliefs). Importantly, it predicted a fast but gradual decline in RTs when subjects learned the specific patterns (Figure 2A). In addition to the state-level entropy (Figure 2B), we also calculated structure-level entropy (uncertainty about the underlying pattern, Figure 2C) and used it as a separate variable of interest for fMRI analyses (it did not provide additional explanatory power for the RT analysis). Note that increases in the structure entropy in patterns of length 4 and 6 are caused by a flip in beliefs from greatly favoring no pattern to slightly favoring a pattern; this leads to an increase in entropy.

Neural Correlates of Pattern Learning

Given that pattern learning is characterized by rapid declines in RT, we first took a model-free approach and investigated neural activity associated with decreases in RT. We chose two measures that might reflect different aspects of the learning process.

First, for each sequence, for every subject, we identified the one trial (image within the sequence) with the largest decrease in RT compared to the previous trial. Typically, these RT drops happened in the middle of the sequence, with shorter patterns exhibiting earlier drops (Figure 2D), reflecting the most informative point in a sequence. We included all the data in the analysis and controlled for the condition (pattern length), as well as RT on each trial (STAR Methods, GLM1).

Second, we ran an analysis including the change in RT on each trial (RT on the current trial minus RT on the previous trial), which reflects the entire learning process. Again we controlled for the condition and RT on each trial (STAR Methods, GLM2).

We included the resulting contrast images in a second-level analysis using non-parametric methods; all reported results include FWE-corrected voxel clusters at $p < 0.05$ and cluster-forming threshold $p = 0.0001$ (SnPM default, also see the discussion by Woo et al., 2014).

Contrasting pattern versus no-pattern sequences, we identified three clusters that were more active in the pattern case: the left angular gyrus (peak $x = -56$, $y = -56$, $z = 38$, all coordinates reported in the MNI space), the precuneus (peak $x = -14$, $y = -64$, $z = 0$), and the right superior frontal gyrus (SFG, peak $x = 16$, $y = 10$, $z = 56$). More importantly, we found that the largest drops in RT were associated with increased activity in the right intraparietal sulcus (IPS, peak $x = 38$, $y = -50$, $z = 42$) and bilaterally in the lateral prefrontal cortex (lPFC, including regions of the middle frontal gyrus and middle orbital gyrus, right peak $x = 32$, $y = 42$, $z = 12$, left peak $x = -38$, $y = 50$, $z = 24$) (Figures 3A and 3B; Table S2). Additionally, two clusters in the left precuneus showed decreased activity during these trials (peaks $x = -16$, $y = -50$, $z = 8$ and $x = -4$, $y = -62$, $z = 24$).

Previous studies have associated this frontoparietal network with a number of different paradigms, including learning (Lee et al., 2014), attention (Corbetta and Shulman, 2002), and memory (Linden et al., 2003) tasks; based on evidence from the Neurosynth database (Yarkoni et al., 2011) (http://www.neurosynth.org), the map of activity reflecting drop in RTs is primarily correlated with maps related to working memory (the top two keywords associated with the map using the image decoding tool were "working" and "working memory").

We also ran a GLM that additionally included an interaction between the condition (pattern versus no pattern) and RT drop, since we would expect RT drops to correspond to learning in the pattern condition, but not in the no-pattern condition. However, this interaction did not yield any significant results, suggesting that RT drops in the no-pattern condition may have reflected subjects thinking (incorrectly) that they had learned something. Indeed, drops in RTs (and entropy) can also happen in sequences with no patterns, so it is important to apply a model-based analysis that can account for such situations.

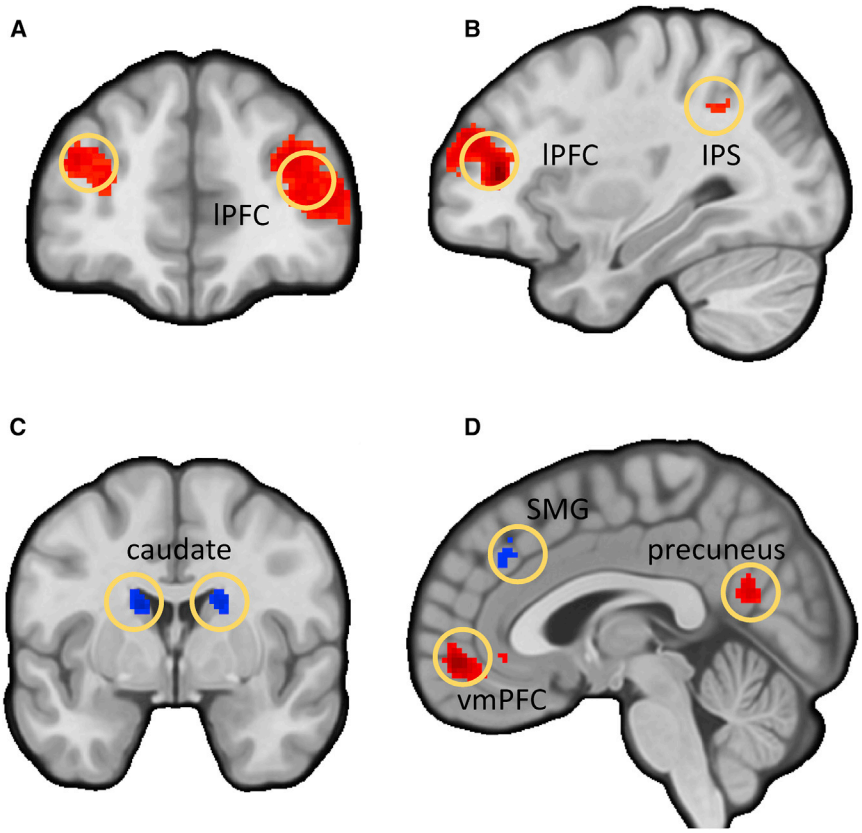

Figure 3. Model-Free Analysis of RT Changes

(A and B) BOLD contrast of single trials within each sequence that have the largest RT decrease relative to the previous trial, compared with all other trials in the sequence. Coronal (A, y = 48) and sagittal (B, x = 32) slices depict positive activations in the lPFC and right IPS.

(C and D) Neural maps showing significant clusters of activity modulated by trial-by-trial change in RT. Coronal (C, y = 0) and sagittal (D, x = 2) slices depict positive correlations in the vmPFC and precuneus, and negative correlations in the bilateral caudate and SMG/ACC. All images shown with threshold at $p < 0.05$, FWE-corrected ($t = 4.42$). See also Tables S2 and S3.

We also identified a number of clusters showing both positive and negative correlations with changes in RT (Table S3). Notably, clusters in the ventromedial prefrontal cortex (vmPFC, peak x = 2, y = 54, z = −6), precuneus (x = −8, y = −60, z = 16), left fusiform (x = −28, y = −34, z = −16), and right hippocampus (x = 26, y = −14, z = −22) showed positive relationship with changes in RT (Figures 3C and 3D), implying that activity in these areas was typically decreasing toward the middle of a sequence and the increasing toward the end. Clusters in the ventrolateral PFC (vlPFC) (x = 20, y = 56, z = −6), superior medial gyrus (SMG)/anterior cingulate cortex (ACC) (x = 4, y = 38, z = 34), bilateral caudate (x = 18, y = 0, z = 22 and x = −14, y = 0, z = 20), and right inferior frontal gyrus (IFG, x = 34, y = 24, z = −4) showed a negative correlation, with activity typically increasing in the beginning of the sequence and decreasing toward the end.

While illustrative, these analyses do not provide new insight into the specific roles of these regions in the learning process. As noted in the introduction, many of these regions have been implicated in sequence learning. As our behavioral fits demonstrate, RTs could be influenced by different underlying processes, which are not necessarily directly reflected in behavior (such as structure-level inference). To further investigate the role of these regions in the sequence learning process, we employed model-based analysis of the BOLD data.

Neural Signatures of State-Level Uncertainty

Here we sought to identify neural signatures of the time series of entropy predictions from the best-fitting (Bayesian pattern-learning) model. We calculated entropy of beliefs on both levels, as well as surprise about the current image, and included them as parametric modulators at each stimulus (image animation) onset in the single-subject analyses, controlling for RT and pattern length (2–6) (GLM3, STAR Methods). The voxel-wise parameter estimates for these regressors indicate how BOLD activity in specific regions covaries with image prediction uncertainty, with positive correlation reflecting *higher* activity when the subject is uncertain about the current observed outcome.

We found that higher-state-level entropy situations are correlated with higher activity in the same frontoparietal network we observed in our model-free analyses. This includes bilateral lPFC (left peak x = −30, y = 56, z = 8, right peak x = 30, y = 54, z = 2) and the IPS (left peak x = −38, y = −50, z = 40, right: peak x = 42, y = −56, z = 48) (Figures 4A and 4B; Table S4). These regions have been previously shown to reflect SPE in reinforcement learning tasks (IPS; Gläscher et al., 2010) and reliability of model-based/model-free learning systems (vlPFC; Lee et al., 2014), also referred to as inferior lateral PFC, extending to dorsolateral PFC), which is consistent with higher belief uncertainty. To illustrate differences in activity in the IPS and vlPFC between our conditions, we extracted the raw BOLD signal using masks defined as 10 mm spheres around the peak activation in Gläscher et al. (2010) and Lee et al. (2014) and plotted levels of activity for patterns of different lengths (Figures 4C and 4D). We found that the IPS and (to a lesser extent) vlPFC activity monotonically reflected the pattern length, increasing as the complexity of the pattern increased and levelling off in the later part of the sequence.

Conversely, a large number of regions showed negative correlation with state-level entropy (i.e., more active with more certainty). Notably, we observed significant effects in the bilateral hippocampus and parahippocampal gyrus (left peak, x = −20, y = −18, z = −18; right peak, x = 24, y = −18, z = −18), bilateral anterior insula (left, x = −48, y = 4, z = 8; right, x = 42, y = 4,

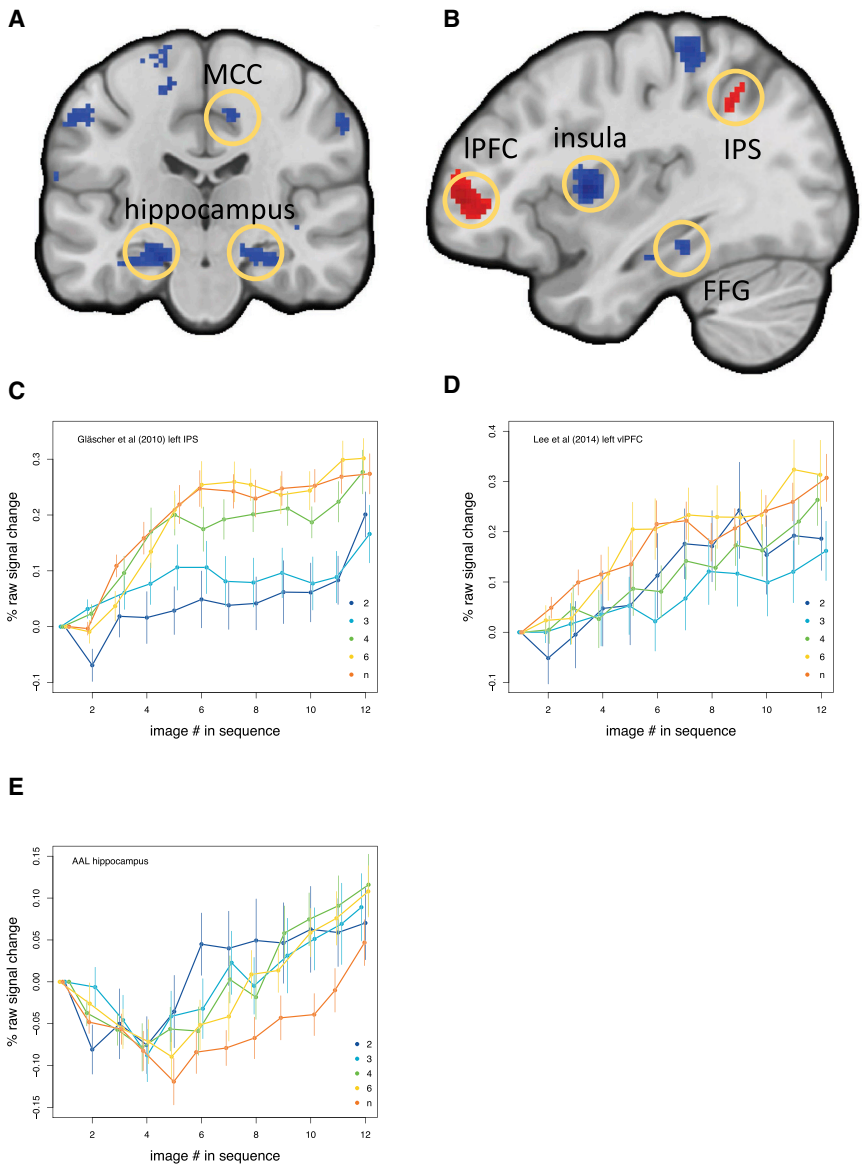

Figure 4. BOLD Effects of State-Level Entropy

Clusters of BOLD activity correlated with the state-level entropy used as a parametric modulator of the HRF at the time of stimulus onset.

(A) Coronal slice (y = −20) depicting negative correlation clusters in the bilateral hippocampus and midcingulate cortex (MCC).

(B) Sagittal slice (x = −36) depicting positive activation clusters in the left lateral prefrontal cortex (lPFC) and intraparietal sulcus (IPS) and negative activation clusters in the left anterior insula and fusiform gyrus (FFG). All images shown with threshold at $p < 0.05$, FWE-corrected (t = 4.42).

(C) Raw BOLD signal change in the left IPS depending on the pattern length (indicated with colors, "n" denoting no pattern), with the ROI defined as a 10 mm sphere around the peak activation (−27, −54, 45) from Gläscher et al. (2010).

(D) Raw BOLD signal change in the left ventrolateral PFC depending on the pattern length (indicated with the numbers, with "n" denoting no pattern), with the ROI defined as a 10 mm sphere around the peak activation (−54, 38, 3) from Lee et al. (2014).

(E) Raw BOLD signal change in the hippocampus depending on the pattern length, with the ROI defined using a standard AAL atlas. See also Table S4.

z = 14), fusiform gyrus (FFG, x = −36,y = −32, z = −16), anterior cingulate, and midcingulate cortex (ACC and MCC, x = −10, y = 12, z = 30) (Figures 4A and 4B; Table S4). The activity dynamics extracted using the standard hippocampus mask (AAL) suggest that the hippocampus might be involved in encoding the sequence (as early activity reflects the pattern length), but not retrieving information about it (as in the case of patterns we see no differentiation between pattern lengths, Figure 4E).

We found only one cluster positively correlated with the Shannon surprise S_a, or the negative log of current image probability, in the left inferior frontal gyrus (IFG, also referred to in the literature as the lPFC, peak x = −48, y = 14, z = 20, Figure 5A). Lower surprise showed increased activity in the right superior frontal gyrus (SFG, peak x = 26, y = 62, z = 24), right medial frontal gyrus (MFG, peak x = 38, y = 34, z = 40), and right inferior parietal lobule (IPL, peak x = 56, y = −50, z = 46) (Figure 5A and 5B; Table S5).

These results suggest that the lPFC, also previously shown to reflect SPE (Gläscher et al., 2010; Lee et al., 2014), can be involved in encoding belief error, while the motor cortex activity likely reflects a fast motor response when the belief about the coming image was high.

Finally, to focus on activity during the transition from not knowing to knowing the pattern, we looked for neural correlates of the change in state-level entropy, defined as the difference between the current and previous trial entropy (GLM4, STAR Methods). This analysis is essentially the model-based version of our earlier analysis looking at the biggest drops in RT. In the case of no pattern, this variable stays roughly constant over the course of the sequence, while in the pattern case it exhibits a U-shape, decreasing as the pattern becomes clear and then increasing as subjects become confident that they know the pattern. Regions that are positively correlated with the change in entropy thus have a decrease in activity in the beginning of the sequence, with an increase toward the end of the sequence, while negatively correlated regions show an increase in the beginning of the sequence and a decrease toward the end.

We found that activity in a number of regions in the prefrontal cortex and the striatum were negatively correlated with the change in entropy, including the posterior-medial frontal cortex and ACC (peak x = 8, y = 4, z = 62), right putamen

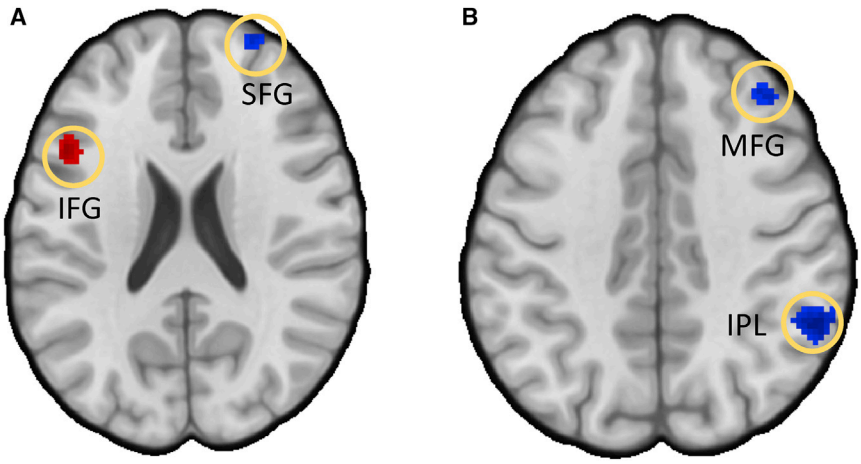

Figure 5. BOLD Effects of State-Level Surprise

Clusters of BOLD activity correlated with the Bayesian surprise as a parametric modulator of the HRF at the time of stimulus onset.

(A) Axial slice (z = 22) depicting a positive correlation cluster in the left inferior frontal gyrus (IFG) and a negative correlation cluster in the superior frontal gyrus (SFG).

(B) Axial slice (z = 40) depicting negative activation clusters in the medial frontal gyrus (MFG) and inferior parietal lobule (IPL). All images shown with threshold at $p < 0.05$, FWE-corrected ($t = 4.42$). See also Table S5.

(peak x = 20, y = 14, z = −8), left insula (peak x = −38, y = 10, z = 4), and right thalamus (peak x = 6, y = −10, z = 6) (Figure S3; Table S6). We did not find any significant clusters with positive correlation.

Neural Signatures of Structure-Level Uncertainty

To investigate neural correlates of the structure-level uncertainty, we included the structure-level entropy time series produced by the pattern-learning model in the same GLM as the state-level time series and surprise. Although there was a correlation between entropy on the two levels ($r = 0.58$, $p < 0.001$), we intended to separate the variance explained by the two variables and turned off the automatic orthogonalization in SPM and controlled for RT (GLM3, STAR Methods). Our analysis revealed a distinct, separate network of regions tracking structure-level entropy (Figures 6A–6C and S7; Table S7).

We found positive correlation with regions typically implicated in estimates of uncertainty and sequence/structure learning: the premotor cortex (peak x = −24, y = 2, z- 56), dorsal striatum (putamen and caudate, peak x = −12, y = 10, z = −2), precuneus (peak x = −10, y = −66, z = 50), left IFG (peak x = −33, y = 30, z = −2), middle frontal gyrus (MFG, x = −24, y = 2, z = 56), and a large cluster in the occipital lobe (primary visual cortex, peak x = 12, y = −82, z = 14). Large activation in the visual area during situations of high structural uncertainty is consistent with subjects relying more on the visual system as they identify the images. The dorsal striatum and the left IFG have been previously shown to be involved in structure learning and sequential prediction.

Regions that are more active when uncertainty is low include clusters in the vmPFC (peak x = 12, y = 46, z = −14) and more inferior visual regions in the occipital lobe (peaks x = −20, y = −90, z = −2, and x = 24, y = −84, z = −4). Activity in the vmPFC reveals a roughly monotonic relationship between the pattern length and the BOLD signal (Figure 6D), with shorter patterns peaking earlier. Notably, vmPFC activity also increases in the case of no pattern, indicating that this activity is not simply due to the higher rewards that subjects could earn once they learned the patterns. Instead it suggests that the vmPFC activity reflects the certainty that there is or is not a pattern. This could be due to subjects feeling a sense of reward from figuring out that there is no pattern, or it could simply reflect learning disengagement.

Analysis of the BOLD signal correlated with the change in structure-level entropy (GLM4, STAR Methods) showed that larger changes in structural uncertainty (i.e., moments when subjects figured out that there was or was not a pattern) were associated with stronger activity in large clusters in mPFC (peak x = −8, y = 42, z = −4), MCC (peak x = −0, y = −12, z = 42), posterior cingulate cortex (peak x = 0, y = −60, z = 24), and left fusiform/parahippocampal gyrus (peak x = −26, y = −38, z = −10) (Figure 7; Table S8), consistent with encoding of the pattern and possible reward anticipation. Conversely, these moments cause a *decrease* of activity in the same frontoparietal network that we found to be associated with higher-state-level uncertainty: the lPFC (left peak x = −32, y = 56, z = 12, right peak x = 46, y = 54, z = −4) and IPS/inferior parietal lobule (IPL, peak x = −44, y = −50, z = 46), as well as the insula (peak x = 30, y = 28, z = −4) and superior medial gyrus (SMG, peak x = 2, y = 22, z = 44) (Figure 7).

Individual Differences in Task Performance and Neural Representations

Finally, we undertook exploratory analyses to investigate correlations between subjects' behavior and BOLD activity in key entropy-related regions. To reduce the number of multiple comparisons, we only used neural betas for state-level and structure-level entropy. We also restricted our analysis to four regions of interest (ROIs): a standard AAL hippocampus mask, the bilateral IPS mask based on the peak activations in Gläscher et al. (2010), the bilateral vlPFC mask based on the peak activations in Lee et al. (2014), and the vmPFC mask from Bartra et al. (2013). From each ROI we extracted one GLM beta per subject.

For behavior we focused on two RT-based measures, since accuracy was close to ceiling. The first measure is the subject-specific difference in mean RT between the sequences with patterns and sequences without patterns. The second measure is the speed of pattern learning, which we constructed as follows. Using the pooled data, we identified an RT threshold, below

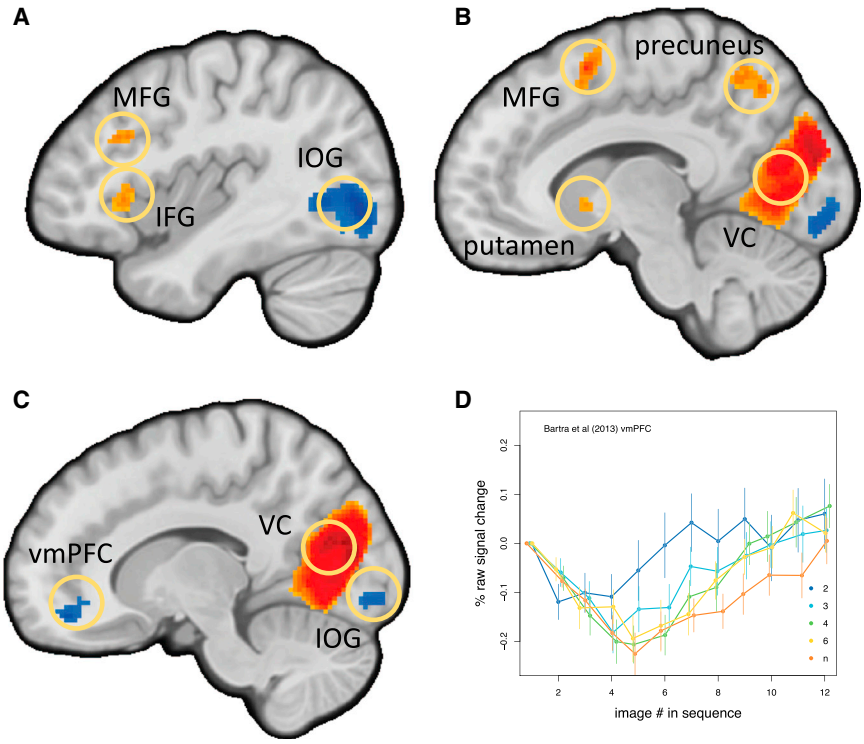

Figure 6. BOLD Effects of Structure-Level Entropy

Clusters of BOLD activity correlated with the structure-level entropy used as a parametric modulator of the HRF at the time of stimulus onset.

(A) Sagittal slice (x = −40) depicting positive correlation clusters in the left middle frontal gyrus (MFG) and left inferior frontal gyrus (IFG), and a negative correlation cluster in the inferior occipital gyrus (IOG).

(B) Sagittal slice (x = −10) depicting positive activation clusters in the striatum (putamen), middle frontal gyrus (MFG), precuneus, and primary visual cortex (V1).

(C) Sagittal slice (x = 12) depicting a negative activation cluster in vmPFC and the inferior occipital gyrus, and positive activation cluster in the visual cortex (VC). All images shown with threshold at $p < 0.05$, FWE-corrected ($t = 4.42$).

(D) Raw BOLD signal change during the full sequence in the vmPFC depending on the pattern length (indicated with colors, "n" denoting no pattern), with the ROI defined using a mask from Bartra et al. (2013). See also Table S7.

which the chance that the sequence had a pattern was 5% or less; this threshold was 710 ms. Then, for each pattern sequence and subject, we identified the first image in the sequence when this threshold was crossed. We averaged these time points within each subject, creating a measure ("learning onset") of how quickly each subject learned the patterns.

Four ROIs, two sets of betas, and two behavioral measures resulted in 16 correlation tests between the subject-specific betas and behavioral performance. We found that the structure-level representations in the hippocampus correlated with both behavioral measures ($r = 0.52$, $p = 0.01$ for the difference in RTs, and $r = 0.37$, $p = 0.078$ for the average timing of learning; Figure 8). Additionally, structure-level betas for the vmPFC were correlated with the RT differences ($r = 0.41$, $p = 0.046$; see Figure S4).

DISCUSSION

Learning of hidden patterns is ubiquitous in decision-making. Underlying transition structures in sequences can easily be modeled using the popular Markov model approach; however, our reaction time data demonstrate that people learn faster than this class of models predicts. Combining ideas from the reinforcement-learning and structure-learning literatures, we showed that our subjects' RTs are consistent with a Bayesian model that assumes that the brain divides all possible patterns into specific groups (structures, or pattern types), assigns beliefs to each group, and updates them during the learning process. This pattern-learning model implies that the brain tracks uncertainty of both the predicted state and the type of the underlying pattern; our fMRI results show that these two levels of uncertainty are indeed correlated with activity in two separate networks of regions, not including regions that reflect anticipatory activity related to particular image types (Figure S8).

Of particular note are our findings regarding the structure-level network. While prior research has studied belief updating in a variety of settings, our structure-level results indicate that the brain tracks more than just beliefs about states. In particular, we identified a network including the dorsal striatum, IFG, MFG, precuneus, and (primary) visual cortex, which tracked the uncertainty about the type of pattern. We also found activity in vmPFC and a different region of visual cortex (the inferior occipital gyrus), which showed the opposite correlation, tracking the certainty about the type of pattern. That being said, we cannot definitively conclude that these regions are exactly computing entropy. More work is needed to clearly disentangle the link between RT, uncertainty, and timing, and to compare alternative models and measures of uncertainty to better understand the mechanistic roles of these regions in learning; in this section we discuss some of the potential computational roles of the areas we identified.

Changes in structure level uncertainty are also of note, as they index when subjects learn something about the type of pattern. Several of the regions tracking the change in structure-level uncertainty corresponded with regions also tracking the state-level uncertainty, since in the pattern sequences, learning about the type of pattern means knowing what image is coming next. Regions showing the most activity coinciding with the steepest learning were IPS and lPFC, while regions showing the opposite pattern included cingulate cortex and the hippocampus. There were, however, other regions that did not also track state-level uncertainty, including the insula and MFG/SFG during steep learning, and mPFC during periods of no learning. Of particular note here is the hippocampus, which showed increased activity after learning whether there was a pattern, and whose activity correlated with the speed and degree of learning, at the subject level.

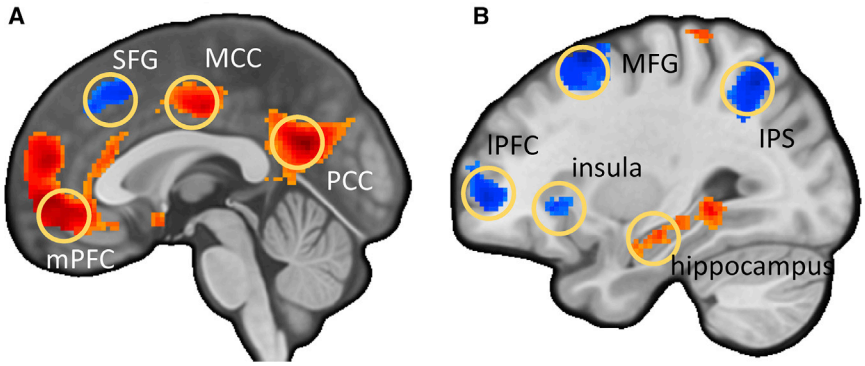

Figure 7. BOLD Effects of Change in Structure-Level Entropy, $(E_s^t - E_s^{t-1})$
Clusters of BOLD activity correlated with the change in state-level entropy used as a parametric modulator of the HRF at the time of stimulus onset. (A) Sagittal slice (x = 0) depicting a negative correlation cluster in the superior frontal gyrus (SFG) and positive correlation clusters in the medial prefrontal cortex (mPFC), midcingulate cortex (MCC), and posterior cingulate cortex (PCC). (B) Sagittal slice (x = 30) depicting a positive activation cluster in the right hippocampus and negative activation clusters in the lPFC, insula, medial frontal gyrus (MFG), and intraparietal sulcus (IPS). All images shown with threshold at $p < 0.05$, FWE-corrected (t = 4.42). See also Table S8.

Frontoparietal Network and Attention

Some of the areas we found to be associated with state-level entropy are a part of the frontoparietal network associated with spatial attention and attentional control, which includes the intraparietal sulcus (IPS), the frontal eye fields (FEF), the supplementary motor area (SMA), the dorsolateral PFC, and other regions. Specifically, a number of studies have shown that the IPS is involved in feature-based or object-based attentional control (for instance, detecting luminance changes). Some studies have shown functional differences in sub-regions of the IPS, with its topographical organization reflecting multimodal representation of attention (Anderson et al., 2010). Since we found the IPS (along with the lPFC) to be more active in high state-entropy situations, one explanation for this result could be increased attentional control with higher uncertainty. However, time courses of the IPS activity demonstrate that it leveled off after subjects learned the patterns, and these levels were roughly monotonic with the length of the pattern (Figure 4C). This might be an argument against the attentional account, since once the pattern is learned, attention would presumably decrease and reach a common level for all pattern lengths.

Working Memory

Notably, the IPS is also a "core node" of the network that is involved in working memory, showing increased activity during N-back tasks (Rottschy et al., 2012) and with set size in visual working memory tasks (Luck and Vogel, 2013; Silk et al., 2010). While sequence learning and motor chunking tasks activate the IPS (Kumaran and Maguire, 2006a; Wymbs et al., 2012), some argue that the IPS is not required for learning all ordered sequences, and instead that the representation of the sequence can depend on the type of the stimuli (with the IPS processing numbers and letters) (Van Opstal et al., 2009).

Given that we observed a dissociation between different pattern lengths in the IPS by the end of the sequence, as well as its sustained activity (Vilberg and Rugg, 2012), we conjecture that in our experiment the IPS is involved in retrieval of the stored sequence representations (Figure 4C). Note that activity in the hippocampus, another area important for storing sequential representations (Bornstein and Daw, 2012; Vilberg and Rugg, 2012), did not differentiate between the pattern lengths at the end of the sequence (Figure 4E), and thus it is unlikely to be involved in retrieval.

Another potential candidate region for storing sequential information is the left IFG (Doeller et al., 2006; van der Graaf et al., 2006; Huettel et al., 2002); however, some prior work has demonstrated that activity in this region could be simply related to the use of verbal encoding strategies (Van Opstal et al., 2009). We have found it to be primarily correlated with Bayesian surprise, which could be an indication that the IFG encodes sequential violations. Finally, another important part of the memory network is the vlPFC, which was also active in situations with higher uncertainty in our task. Previous evidence has suggested that the vlPFC is involved in the cognitive control of memory, with anterior vlPFC accessing stored representations, and mid-vlPFC operating after retrieval to select among competing representations (Badre and Wagner, 2007). Others, however, argue that the vlPFC is mainly involved in semantic processing during memory retrieval (Han et al., 2012). Finally, increased retrieval demand produces stronger BOLD response in a network that is functionally connected and includes the vlPFC and the hippocampus, with the vlPFC potentially serving as a hub that connects the dorsal fronto-parietal (the IPS-dlPFC) and the ventral memory networks (Barredo et al., 2015).

Prediction Error Account

The caudate and putamen have consistently been shown to be involved in sequence learning in a variety of SRT tasks, artificial grammar-learning problems, and motor chunking, with higher activity often reflecting a violation in a repeating pattern (Huettel et al., 2002; Kalm et al., 2013; Kumaran and Maguire, 2006a; Lungu et al., 2014; Schendan et al., 2003; Schwarb and Schumacher, 2009; Werheid et al., 2003; Wymbs et al., 2012; Yang and Li, 2012). These violations produce prediction errors, and the caudate has been shown to encode prediction error at the level of sequences of actions (Schiffer et al., 2012) and to reflect uncertainty in category learning (Daniel et al., 2011). Our results, showing higher activity in the dorsal striatum in situations of higher structure-level uncertainty, are consistent with these accounts. In a paradigm similar to ours, but employing stochastic transitions between states, the caudate encoded uncertainty about state prediction (forward entropy) (Bornstein and Daw, 2013).

SPEs (difference between expected belief and realized outcome) have also been shown to be encoded in the IPS and the lPFC (in the dorsal part and the IFG) (Gläscher et al., 2010;

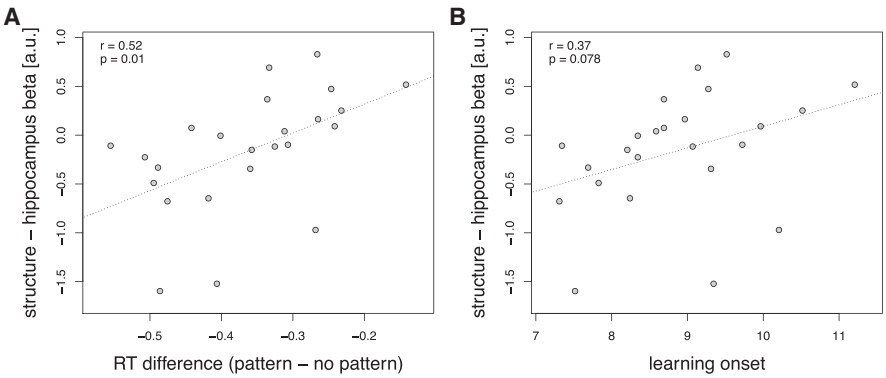

Figure 8. Correlations between Neural Betas and Behavioral Performance

Each point is an individual subject. The beta coefficients were extracted from GLM3 within a standard AAL hippocampus mask. RT difference is equal to mean RT in the pattern condition minus the mean RT in the no-pattern condition. Learning onset reflects the average image position when the subject learned a pattern. The dotted line shows the best-fitting regression line. See also Figure S4.

Lee et al., 2014). Consistent with those results, we also observed increased activity in the IPS with higher uncertainty (and thus higher SPE), while activity in the lPFC was correlated with the Bayesian surprise measure (additionally, we observed that these regions were associated with longer RTs, which we included as a control regressor; previous studies did not control for RT). Interestingly, the more ventral (and anterior) part of the vlPFC that we identified was previously associated with reliability of model-based/model-free control in a reinforcement learning paradigm (Lee et al., 2014) and computing evidence in favor of switching between courses of action (Boorman et al., 2009).

Structure Learning

In order to act in a complex environment, the brain has to form and store its state representation. It has been argued that the brain can potentially infer underlying structure of the environment, and store this causal information as a separate representation (Gershman et al., 2015). Two regions can potentially form the core of a network that implements this process, with the hippocampus serving as an encoder, and the OFC (or vmPFC) as a storage for structural representations. In many settings, the hippocampus has been shown to encode a spatial or structural map of the environment (Aggleton et al., 2007; Garvert et al., 2017; Onuki et al., 2015; Stachenfeld et al., 2017), and animals with lesions to the hippocampus lose the ability to distinguish between different contexts (Ji and Maren, 2007). Similarly, the OFC is vital for higher-order structure representations and reasoning, serving as a cognitive map of the task space (McDannald et al., 2011; Wilson et al., 2014) and implementing mechanisms of probabilistic reasoning (Donoso et al., 2014). In line with these ideas, we observed activity in the vmPFC and hippocampus associated with structure-level uncertainty. Our results showed that activity in the vmPFC was able to dissociate between different pattern lengths (Figure 6D); it is unlikely that this BOLD signal purely reflects reward (Bartra et al., 2013), as we observed an increase in activity not only in sequences with patterns but also in the sequences with no patterns, where subjects received no additional payment as the sequence progressed.

Taken together, our results provide neurocomputational validation of the idea that structure learning in the brain is organized as a complex Bayesian updating process that selects among many classes of potential structures, using prior beliefs about empirical frequencies of these structures. In our model, we simplified the process and assumed that the structure space corresponds to the patterns we used in the experiment. In the real world, the pattern space can be much larger (and potentially infinite), and so more work is needed to clarify the structure space selection process and hypothesis generation. In terms of the neural underpinnings of structure learning, experimental separation of attention, memory, and uncertainty remain important factors to separate in future research.

STAR★METHODS

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

- KEY RESOURCES TABLE
- CONTACT FOR REAGENT AND RESOURCE SHARING
  - Participants
- METHOD DETAILS
  - Task
  - Imaging methods
  - Behavioral analysis
  - Imaging analysis
  - Data and software availability

SUPPLEMENTAL INFORMATION

Supplemental Information includes eight tables and eight figures and can be found with this article at https://doi.org/10.1016/j.neuron.2018.05.013.

ACKNOWLEDGMENTS

We thank Kareem Soliman for research assistance, Aidan Makwana for input on the scanning procedures and data analysis, Julie Golomb for advice on the task design, and Erie Boorman and Todd Hare for their feedback on the manuscript. The project was supported by the OSU CCBBI Neuroimaging Student award.

AUTHOR CONTRIBUTIONS

A.K. and I.K. designed the experiment and analyses. A.K. programmed and conducted the experiment, performed the data analysis, and co-wrote the paper. I.K. co-wrote the paper and supervised the project.

DECLARATION OF INTERESTS

The authors declare no competing financial interests.

Received: December 12, 2017
Revised: March 26, 2018
Accepted: May 7, 2018
Published: May 31, 2018

REFERENCES

Aggleton, J.P., Sanderson, D.J., and Pearce, J.M. (2007). Structural learning and the hippocampus. Hippocampus *17*, 723–734.

Anderson, J.S., Ferguson, M.A., Lopez-Larson, M., and Yurgelun-Todd, D. (2010). Topographic maps of multisensory attention. Proc. Natl. Acad. Sci. USA *107*, 20110–20114.

Badre, D., and Wagner, A.D. (2007). Left ventrolateral prefrontal cortex and the cognitive control of memory. Neuropsychologia *45*, 2883–2901.

Barredo, J., Öztekin, I., and Badre, D. (2015). Ventral fronto-temporal pathway supporting cognitive control of episodic memory retrieval. Cereb. Cortex *25*, 1004–1019.

Bartra, O., McGuire, J.T., and Kable, J.W. (2013). The valuation system: a coordinate-based meta-analysis of BOLD fMRI experiments examining neural correlates of subjective value. Neuroimage *76*, 412–427.

Bates, D., Maechler, M., and Bolker, B. (2012). lme4: Linear mixed-effects models using S4 classes.

Beierholm, U.R., Anen, C., Quartz, S., and Bossaerts, P. (2011). Separate encoding of model-based and model-free valuations in the human brain. Neuroimage *58*, 955–962.

Boorman, E.D., Behrens, T.E.J., Woolrich, M.W., and Rushworth, M.F.S. (2009). How green is the grass on the other side? Frontopolar cortex and the evidence in favor of alternative courses of action. Neuron *62*, 733–743.

Bornstein, A.M., and Daw, N.D. (2012). Dissociating hippocampal and striatal contributions to sequential prediction learning. Eur. J. Neurosci. *35*, 1011–1023.

Bornstein, A.M., and Daw, N.D. (2013). Cortical and hippocampal correlates of deliberation during model-based decisions for rewards in humans. PLoS Comput. Biol. *9*, e1003387.

Brainard, D.H., and Vision, S. (1997). The psychophysics toolbox. Spat. Vis. *10*, 433–436.

Corbetta, M., and Shulman, G.L. (2002). Control of goal-directed and stimulus-driven attention in the brain. Nat. Rev. Neurosci. *3*, 201–215.

Daniel, R., Wagner, G., Koch, K., Reichenbach, J.R., Sauer, H., and Schlösser, R.G. (2011). Assessing the neural basis of uncertainty in perceptual category learning through varying levels of distortion. J. Cogn. Neurosci. *23*, 1781–1793.

Davis, T., Love, B.C., and Preston, A.R. (2012). Striatal and hippocampal entropy and recognition signals in category learning: simultaneous processes revealed by model-based fMRI. J. Exp. Psychol. Learn. Mem. Cogn. *38*, 821–839.

Daw, N.D., Gershman, S.J., Seymour, B., Dayan, P., and Dolan, R.J. (2011). Model-based influences on humans' choices and striatal prediction errors. Neuron *69*, 1204–1215.

Dehaene, S., Meyniel, F., Wacongne, C., Wang, L., and Pallier, C. (2015). The neural representation of sequences: from transition probabilities to algebraic patterns and linguistic trees. Neuron *88*, 2–19.

Doeller, C.F., Opitz, B., Krick, C.M., Mecklinger, A., and Reith, W. (2006). Differential hippocampal and prefrontal-striatal contributions to instance-based and rule-based learning. Neuroimage *31*, 1802–1816.

Doll, B.B., Simon, D.A., and Daw, N.D. (2012). The ubiquity of model-based reinforcement learning. Curr. Opin. Neurobiol. *22*, 1075–1081.

Doll, B.B., Duncan, K.D., Simon, D.A., Shohamy, D., and Daw, N.D. (2015). Model-based choices involve prospective neural activity. Nat. Neurosci. *18*, 767–772.

Donoso, M., Collins, A.G.E., and Koechlin, E. (2014). Human cognition. Foundations of human reasoning in the prefrontal cortex. Science *344*, 1481–1486.

FitzGerald, T.H.B., Hämmerer, D., Friston, K.J., Li, S.-C., and Dolan, R.J. (2017). Sequential inference as a mode of cognition and its correlates in fronto-parietal and hippocampal brain regions. PLoS Comput. Biol. *13*, e1005418.

Garvert, M.M., Dolan, R.J., and Behrens, T.E. (2017). A map of abstract relational knowledge in the human hippocampal-entorhinal cortex. eLife *6*, e17086.

Gershman, S.J., and Niv, Y. (2010). Learning latent structure: carving nature at its joints. Curr. Opin. Neurobiol. *20*, 251–256.

Gershman, S.J., and Niv, Y. (2015). Novelty and inductive generalization in human reinforcement learning. Top. Cogn. Sci. *7*, 391–415.

Gershman, S.J., Norman, K.A., and Niv, Y. (2015). Discovering latent causes in reinforcement learning. Curr. Opin. Behav. Sci. *5*, 43–50.

Gheysen, F., Van Opstal, F., Roggeman, C., Van Waelvelde, H., and Fias, W. (2010). Hippocampal contribution to early and later stages of implicit motor sequence learning. Exp. Brain Res. *202*, 795–807.

Gläscher, J., Daw, N., Dayan, P., and O'Doherty, J.P. (2010). States versus rewards: dissociable neural prediction error signals underlying model-based and model-free reinforcement learning. Neuron *66*, 585–595.

Han, S., O'Connor, A.R., Eslick, A.N., and Dobbins, I.G. (2012). The role of left ventrolateral prefrontal cortex during episodic decisions: semantic elaboration or resolution of episodic interference? J. Cogn. Neurosci. *24*, 223–234.

Harrison, L.M., Duggins, A., and Friston, K.J. (2006). Encoding uncertainty in the hippocampus. Neural Netw. *19*, 535–546.

Huettel, S.A., Mack, P.B., and McCarthy, G. (2002). Perceiving patterns in random series: dynamic processing of sequence in prefrontal cortex. Nat. Neurosci. *5*, 485–490.

Huys, Q.J.M., Eshel, N., O'Nions, E., Sheridan, L., Dayan, P., and Roiser, J.P. (2012). Bonsai trees in your head: how the pavlovian system sculpts goal-directed choices by pruning decision trees. PLoS Comput. Biol. *8*, e1002410.

Ji, J., and Maren, S. (2007). Hippocampal involvement in contextual modulation of fear extinction. Hippocampus *17*, 749–758.

Jiménez, L., and Mendez, C. (1999). Which attention is needed for implicit sequence learning? J. Exp. Psychol. Learn. Mem. Cogn. *25*, 236.

Jiménez, L., Vaquero, J.M., and Lupiáñez, J. (2006). Qualitative differences between implicit and explicit sequence learning. J. Exp. Psychol. Learn. Mem. Cogn. *32*, 475–490.

Kalm, K., Davis, M.H., and Norris, D. (2013). Individual sequence representations in the medial temporal lobe. J. Cogn. Neurosci. *25*, 1111–1121.

Konovalov, A., and Krajbich, I. (2016). Gaze data reveal distinct choice processes underlying model-based and model-free reinforcement learning. Nat. Commun. *7*, 12438.

Kool, W., Cushman, F.A., and Gershman, S.J. (2016). When does model-based control pay off? PLoS Comput. Biol. *12*, e1005090.

Kumaran, D., and Maguire, E.A. (2006a). The dynamics of hippocampal activation during encoding of overlapping sequences. Neuron *49*, 617–629.

Kumaran, D., and Maguire, E.A. (2006b). An unexpected sequence of events: mismatch detection in the human hippocampus. PLoS Biol. *4*, e424.

Kumaran, D., and Maguire, E.A. (2007). Which computational mechanisms operate in the hippocampus during novelty detection? Hippocampus *17*, 735–748.

Lee, S.W., Shimojo, S., and O'Doherty, J.P. (2014). Neural computations underlying arbitration between model-based and model-free learning. Neuron *81*, 687–699.

Linden, D.E.J., Bittner, R.A., Muckli, L., Waltz, J.A., Kriegeskorte, N., Goebel, R., Singer, W., and Munk, M.H.J. (2003). Cortical capacity constraints for visual working memory: dissociation of fMRI load effects in a fronto-parietal network. Neuroimage *20*, 1518–1530.

Luck, S.J., and Vogel, E.K. (2013). Visual working memory capacity: from psychophysics and neurobiology to individual differences. Trends Cogn. Sci. *17*, 391–400.

Lungu, O., Monchi, O., Albouy, G., Jubault, T., Ballarin, E., Burnod, Y., and Doyon, J. (2014). Striatal and hippocampal involvement in motor sequence chunking depends on the learning strategy. PLoS ONE *9*, e103885.

Martini, M., Furtner, M.R., and Sachse, P. (2013). Working memory and its relation to deterministic sequence learning. PLoS ONE *8*, e56166.

McDannald, M.A., Lucantonio, F., Burke, K.A., Niv, Y., and Schoenbaum, G. (2011). Ventral striatum and orbitofrontal cortex are both required for model-based, but not model-free, reinforcement learning. J. Neurosci. *31*, 2700–2705.

Nichols, T.E., and Holmes, A.P. (2002). Nonparametric permutation tests for functional neuroimaging: a primer with examples. Hum. Brain Mapp. *15*, 1–25.

O'Doherty, J.P., Hampton, A., and Kim, H. (2007). Model-based fMRI and its application to reward learning and decision making. Ann. N Y Acad. Sci. *1104*, 35–53.

O'Reilly, J.X., Schüffelgen, U., Cuell, S.F., Behrens, T.E.J., Mars, R.B., and Rushworth, M.F.S. (2013). Dissociable effects of surprise and model update in parietal and anterior cingulate cortex. Proc. Natl. Acad. Sci. USA *110*, E3660–E3669.

Onuki, Y., Van Someren, E.J.W., De Zeeuw, C.I., and Van der Werf, Y.D. (2015). Hippocampal-cerebellar interaction during spatio-temporal prediction. Cereb. Cortex *25*, 313–321.

Otto, A.R., Gershman, S.J., Markman, A.B., and Daw, N.D. (2013). The curse of planning: dissecting multiple reinforcement-learning systems by taxing the central executive. Psychol. Sci. *24*, 751–761.

Rottschy, C., Langner, R., Dogan, I., Reetz, K., Laird, A.R., Schulz, J.B., Fox, P.T., and Eickhoff, S.B. (2012). Modelling neural correlates of working memory: a coordinate-based meta-analysis. Neuroimage *60*, 830–846.

Schendan, H.E., Searl, M.M., Melrose, R.J., and Stern, C.E. (2003). An FMRI study of the role of the medial temporal lobe in implicit and explicit sequence learning. Neuron *37*, 1013–1025.

Schiffer, A.-M., Ahlheim, C., Wurm, M.F., and Schubotz, R.I. (2012). Surprised at all the entropy: hippocampal, caudate and midbrain contributions to learning from prediction errors. PLoS ONE *7*, e36445.

Schwarb, H., and Schumacher, E.H. (2009). Neural evidence of a role for spatial response selection in the learning of spatial sequences. Brain Res. *1247*, 114–125.

Sebastiani, P., Ramoni, M., and Cohen, P. (2000). Sequence learning via Bayesian clustering by dynamics. Sequence Learning (Springer), pp. 11–34.

Seidler, R.D., Purushotham, A., Kim, S.-G., Ugurbil, K., Willingham, D., and Ashe, J. (2005). Neural correlates of encoding and expression in implicit sequence learning. Exp. Brain Res. *165*, 114–124.

Shanks, D.R., Wilkinson, L., and Channon, S. (2003). Relationship between priming and recognition in deterministic and probabilistic sequence learning. J. Exp. Psychol. Learn. Mem. Cogn. *29*, 248–261.

Shanks, D.R., Rowland, L.A., and Ranger, M.S. (2005). Attentional load and implicit sequence learning. Psychol. Res. *69*, 369–382.

Silk, T.J., Bellgrove, M.A., Wrafter, P., Mattingley, J.B., and Cunnington, R. (2010). Spatial working memory and spatial attention rely on common neural processes in the intraparietal sulcus. Neuroimage *53*, 718–724.

Stachenfeld, K.L., Botvinick, M.M., and Gershman, S.J. (2017). The hippocampus as a predictive map. Nat. Neurosci. *20*, 1643–1653.

Strange, B.A., Duggins, A., Penny, W., Dolan, R.J., and Friston, K.J. (2005). Information theory, novelty and hippocampal responses: unpredicted or unpredictable? Neural Netw. *18*, 225–230.

Terada, S., Sakurai, Y., Nakahara, H., and Fujisawa, S. (2017). Temporal and rate coding for discrete event sequences in the hippocampus. Neuron *94*, 1248–1262.e4.

van der Graaf, F.H.C.E., Maguire, R.P., Leenders, K.L., and de Jong, B.M. (2006). Cerebral activation related to implicit sequence learning in a Double Serial Reaction Time task. Brain Res. *1081*, 179–190.

Van Opstal, F., Verguts, T., Orban, G.A., and Fias, W. (2008). A hippocampal-parietal network for learning an ordered sequence. Neuroimage *40*, 333–341.

Van Opstal, F., Fias, W., Peigneux, P., and Verguts, T. (2009). The neural representation of extensively trained ordered sequences. Neuroimage *47*, 367–375.

Vandenberghe, M., Schmidt, N., Fery, P., and Cleeremans, A. (2006). Can amnesic patients learn without awareness? New evidence comparing deterministic and probabilistic sequence learning. Neuropsychologia *44*, 1629–1641.

Vilberg, K.L., and Rugg, M.D. (2012). The neural correlates of recollection: transient versus sustained FMRI effects. J. Neurosci. *32*, 15679–15687.

Werheid, K., Zysset, S., Müller, A., Reuter, M., and von Cramon, D.Y. (2003). Rule learning in a serial reaction time task: an fMRI study on patients with early Parkinson's disease. Brain Res. Cogn. Brain Res. *16*, 273–284.

Wilkinson, L., and Shanks, D.R. (2004). Intentional control and implicit sequence learning. J. Exp. Psychol. Learn. Mem. Cogn. *30*, 354–369.

Wilson, R.C., Takahashi, Y.K., Schoenbaum, G., and Niv, Y. (2014). Orbitofrontal cortex as a cognitive map of task space. Neuron *81*, 267–279.

Woo, C.-W., Krishnan, A., and Wager, T.D. (2014). Cluster-extent based thresholding in fMRI analyses: pitfalls and recommendations. Neuroimage *91*, 412–419.

Wymbs, N.F., Bassett, D.S., Mucha, P.J., Porter, M.A., and Grafton, S.T. (2012). Differential recruitment of the sensorimotor putamen and frontoparietal cortex during motor chunking in humans. Neuron *74*, 936–946.

Yang, J., and Li, P. (2012). Brain networks of explicit and implicit learning. PLoS ONE *7*, e42993.

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.

STAR★METHODS

KEY RESOURCES TABLE

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Software and Algorithms		
MATLAB R2017B	MathWorks	https://www.mathworks.com/
Psychtoolbox-3	Brainard and Vision, 1997	http://psychtoolbox.org/
Neurosynth	Yarkoni et al., 2011	http://neurosynth.org/
R version 3.4.0	R Development Core Team	https://www.r-project.org/
LME4	Bates et al., 2012	http://cran.r-project.org/web/packages/lme4/
SPM12	Wellcome Trust	http://www.fil.ion.ucl.ac.uk/spm/software/spm12/
SnPM12	Nichols and Holmes, 2002	https://warwick.ac.uk/fac/sci/statistics/staff/academic-research/nichols/software/snpm
bspmview	Bob Spunt	http://www.bobspunt.com/bspmview/
Marsbar	Jean-Luc Anton, Matthew Brett, Jean-Baptiste Poline, Romain Valabregue	http://marsbar.sourceforge.net/
ArtRepair	Paul Mazaika (CIBSR)	http://cibsr.stanford.edu/tools/human-brain-project/artrepair-software.html

CONTACT FOR REAGENT AND RESOURCE SHARING

As Lead Contact, Ian Krajbich is responsible for all reagent and resource requests. Please contact Ian Krajbich at Department of Psychology, The Ohio State University, 1827 Neil Avenue, 200E Lazenby Hall, Columbus, OH 43210, USA, with requests and inquiries.

Participants

We recruited 26 adult subjects (14 female, mean age 21, range 18 – 28) from the Department of Economics subject pool at the Ohio State University (OSU). Participants were paid based on their overall performance in the task, earning $35 on average, including $10 as a show-up fee. One subject's brain was identified by a physician as abnormal, and one subject had an exceptionally large head, which resulted in high overall motion (> 5 mm) over the course of the experiment and missing imaging data. We excluded these two participants' data, leaving 24 subjects for the analysis. All participants were free of neurological or psychiatric disease, right-handed, with normal or corrected-to-normal vision. The OSU Internal Review Board approved the experiment, and all subjects provided written informed consent.

METHOD DETAILS

Task

We wrote the experimental code in MATLAB (MathWorks), using the Psychophysics Toolbox (Brainard and Vision, 1997).

Before entering the scanner, subjects went through a 15-minute tutorial and learned the correspondence between index, middle, and ring fingers of their right hand, and three images (a face, a hand, and a landscape), which were the same for all subjects. We used three grayscale photos that were normalized for size, contrast, and luminance. During behavioral training, subjects used three arrow keys on a laptop keyboard (left, down, and right). We randomized the mappings between the fingers and the images at the subject level; these correspondences remained the same for each subject during training and the whole duration of the experiment.

Inside the scanner, subjects used an fMRI-compatible 4-button box to make responses, operating the three leftmost buttons. We instructed them to respond with the correct button as soon as they knew which image was presented. We showed each image as an animation, with all images initially scrambled together with Gaussian noise and one image emerging while the noise decreased over time (under a linear rule) (Figure 1). Each animation took 3 s, plus or minus uniformly distributed pseudorandom jitter, up to 486 ms in increments of 54 ms (the length of one slice acquisition, following previous studies [Bornstein and Daw, 2012, 2013]). The speed of animation did not depend on presentation duration: once an image was fully revealed (3 s after the stimulus onset), it stayed that way until the end of the trial. Correct responses triggered a white frame around the image for the remaining stimulus time; if the first response was incorrect, the white frame did not appear even if one of the subsequent responses was correct.

We organized images in sequences, with 12 images in each sequence. Following previous literature (Bornstein and Daw, 2012, 2013), we separated images by a black screen that lasted for 260 ms (1/10 TR). After each 12-image sequence, we presented a black screen for 3 s, followed by the question "Are you sure which image is coming next"?, which required a "yes/no" response using the index and ring fingers, with RT limited to 3 s. After another 3 s of black screen, subjects predicted the next image in the sequence by pressing the corresponding button, with RT again limited to 3 s, regardless of the response.

At the end of the experiment we randomly selected one sequence for payment. For that sequence, we rewarded each (first) correct button press with $2, while also incentivizing subjects to respond quickly, with payment equal to $1 – 0.3RT, with RT in seconds, so subjects could earn a $1 bonus if they responded immediately, and 10c if they responded in 3 s. We also rewarded the correct final image prediction with a $4 bonus.

Each subject observed the same set of 50 sequences. Thirty sequences had a repeating pattern of length 2, 3, 4, and 6 (e.g., HFHFHFHFHFHF and HHFFLLHHFFLL), while 20 sequences were constructed pseudorandomly by manually selecting random sequences that displayed no obvious patterns and had conditional probabilities of roughly 1/3 for each image. During the tutorial, we told subjects that they would encounter repeating patterns and provided a few examples, but we did not describe the full pattern space (length or composition of the patterns) or the true distribution of patterns and non-patterns within the sequence set. We divided the 50 sequences into 5 scanning runs, with each run having 6 sequences with patterns and 4 sequences without patterns (subjects were unaware of this structure). Within a run, we presented all sequences in random order.

We were forced to remove one sequence out of 50 from all of the analyses due to an error in the experimental code. This sequence was intended to have a pattern of length 4, but was erroneously coded with a violation of the pattern in the middle of the sequence. This makes interpretation of this sequence's data ambiguous. This omission did not qualitatively affect the results.

Imaging methods

We collected fMRI data on a 3T Siemens Trio scanner with a 32-channel head coil at the Center for Cognitive and Behavioral Brain Imaging (CCBBI) at the Ohio State University. First, we obtained a T1-weighted high-resolution anatomical image (MPRAGE); second, we collected functional data in 5 runs with 213 acquisitions each. Each acquisition used a gradient-echo EPI sequence with 48 slices of 3mm thickness and 3mm in-plane resolution rotated 30° up from the AC-PC axis (TR = 2.6 s, flip angle = 80°, TE = 28ms). At the end of the experiment, we collected resting state data with 150 brain volume acquisitions using the same parameters, with subjects looking at a white fixation cross.

We preprocessed the data using SPM12 (Wellcome Trust, London) and corrected for motion artifacts using the ArtRepair toolbox (http://cibsr.stanford.edu/tools/human-brain-project/artrepair-software.html). We corrected images for slice time acquisition with the first slice as reference, realigned scans to the first scan within each run, co-registered the anatomical image to the mean of the functional images, and normalized to the standard MNI152 template image. Then we normalized the functional images and smoothed with a FWHM Gaussian kernel of 8mm. We added a nuisance regressor for outlier scans with average motion over 1.5mm in any direction and one more for scans with global BOLD signal changes of more than 5 standard deviations.

Behavioral analysis

Bayesian pattern-learning model

To fit subjects' RTs, we used the following model. Let a_i be the state of the world (i = 'hand', 'face', or 'landscape') and $s_j \in \{s_1, s_2, \ldots, s_6, s_{np}\}$ be the underlying pattern structure, where $s_1, \ldots, s_6$ denote repeating patterns of length 1-6, and s_{np} is a structure that includes all other possible sequences ("no pattern"). We assume that each repeated pattern structure contains all possible sequences of these states given the repetition restriction (i.e., length 1 contains 3 possible sequences, length 2 contains 6 possible sequences, and so on). We assume that subjects have a common prior belief about the structure space:

$$p(s_j) = (p_1, \ldots, p_6, p_{np}),$$

with $\sum_j p(s_j) = 1$. We treat these priors as 6 free parameters that are fit to the data at the group level (note that p_{np} is derived from the other parameters).

Given any structure s_j, a Bayesian observer can compute the probability of an observed history of states h given that structure, or $p(h|s_j)$. For instance, the probability of observing three faces in the beginning of a sequence is 1/3 under p_1 and 0 under p_2. Given the priors $p(s_j)$ and these probabilities, at each step in the sequence the observer can compute a posterior probability of each structure given observed history h:

$$p(s_j|h) = \frac{p(s_j) \cdot p(h|s_j)}{\sum_{k=\{1\ldots7\}} p(s_k) \cdot p(h|s_k)}.$$

Using the posterior and the probabilities of the states conditional on the structure and history, the observer can calculate the probabilities of each state of the world a_i given previous history h:

$$p(a_i|h) = \sum_j p(a_i|s_j, h) \cdot p(s_j|h).$$

Following previous literature (Bornstein and Daw, 2012, 2013; Davis et al., 2012; Harrison et al., 2006; Schiffer et al., 2012), we use Shannon entropy as the standard measure of uncertainty (reflecting the value of additional information about the sequence provided by each image), both on the structure level:

$$E_s = -\sum_j p(s_j|h)\log p\ (s_j|h),$$

and the state level:

$$E_a = -\sum_i p(a_i|h)\log p(a_i|h).$$

Additionally, we calculate surprise, or "negative evidence" about the state as a negative natural log of the current image's probability (Harrison et al., 2006; O'Reilly et al., 2013; Schiffer et al., 2012; Strange et al., 2005):

$$S_a = -\log\ (p(a_i|h)).$$

To fit subjects' responses, we regressed their RTs on the state level entropy and surprise in a linear mixed effects model using the lme4 R package (Bates et al., 2012), with subjects treated as random effects ($RT \sim E_a + S_a + \ (1 + E_a + S_a|id)$), and used Akaike Information Criterion (AIC) as a fitting criterion to estimate the free parameters (6 prior beliefs and the mixed effects model coefficients). Structure-level entropy was not a significant predictor of RTs when we added it to this model, and provided a quantitatively worse fit without state-level entropy in the model. We also explored several models with non-linear (polynomial, log, and exponential) relationship between entropy and RT, but they did not yield a better fit. We used the actual distribution of patterns in the experiment as the starting point and searched over a grid centered around these probabilities.

The exact model fitting procedure was implemented as follows: (1) We created a grid of initial priors built around the true distributions of structures in the experiment, and simulated trial-by-trial predictions for entropy and surprise for each set of priors; (2) We used the resulting entropy and surprise as linear predictors of RT in a hierarchical model (a mixed effects regression that treated subjects as random effects and allowed for individual variability in the relationship between RT and model variables); (3) We calculated the model likelihood (AIC) from the mixed effects regression and used it as the model fitting criterion. (4) We repeated this procedure for each set of parameters in the grid to identify the best fitting set of priors.

To provide a robustness check and reduce the degrees of freedom, we also fit two alternative models: one with priors that were equal to the actual distribution of patterns, and one with flat priors that subjects update as they work through the experiment. Both models provided comparable results, with slightly worse performance. Notably, using the model that has the actual distribution of patterns as the prior does not qualitatively affect the fMRI results. In this model, a sequence's entropy is identical for every subject, confirming that the neural correlates of entropy that we observed were not driven by RTs (Figure S6).

Finally, we also fit the model to each subject individually, allowing for individual differences in prior beliefs; this did not provide a significant improvement in the model fits, and did not affect the neural results (Figure S5).

Markov matrix model

We compared our model to the standard transition matrix model used in the literature (Konovalov and Krajbich, 2016; Otto et al., 2013). Here we assume that the subject updates a k-th order $(k\ x\ 3)$ Markov transition matrix with a uniform categorical prior using the Dirichlet posterior:

$$p(a_{it}|h_k) = \frac{\theta_i + N\ (a_i|h_k)}{\sum_i(\theta_i + N(a_i|h_k))},$$

where θ_i is a concentration hyperparameter of each specific state, and $N\ (a_i|h_k)$ is the empirical number of transitions to state a_i from the previously observed history h of length k. The case of $k = 1$ represents the standard Markov transition process where the next state is only defined by the previous state (history h including only that one state), while higher k's imply beliefs conditional on more observed states (images).

We observed no improvement in model predictions above $k\ =\ 3$, so we report results for models with $k\ =\ \{1, 2, 3\}$. We treat the hyperparameters θ_i as free parameters and assume $\theta_H\ =\ \theta_F\ =\ \theta_L\ :\ =\ \theta$. Although it is common in the literature to assume a flat Dirichlet distribution $(\theta\ =\ 1)$, which provides more gradual changes in beliefs, we found that this restriction yielded a worse fit to the data (see Figure S1). As with the Bayesian pattern-learning model, we calculated state entropy at each step in the sequence and fit it to subjects' RTs using a mixed effects model.

Imaging analysis

We included 5 runs in the same analysis with constants added to the GLM to account for run differences (mean activation and scanner drift). We used SPM12 to run the first-level analyses, and SnPM13 (Nichols and Holmes, 2002) (using the default settings: 5,000 permutations, cluster forming threshold of 0.0001, cluster-wise family-wise error rate of 0.05) for the second-level contrasts (we obtained similar results using the standard SPM12 second-level analyses). We only report clusters that fell below FWE-corrected value of 0.05, as reported by SnPM, with extent over 40 voxels. We visualized the data using the bspmview toolbox for SPM (http://www.bobspunt.com/bspmview/) and used the Harvard-Oxford Atlas and Anatomy Toolbox distributed with bspmview to label the regions.

For each GLM, we locked stimulus onsets to each image presentation, with durations set to variable stimulus presentation lengths (using Dirac functions at stimulus onsets instead yielded similar results) and convolved the BOLD signal with the canonical HRF. We included 6 motion parameters as well as nuisance regressors for images with head motion > 1.5mm compared to the previous volume and signal outliers (> 5 sigma). In each case, we set the regressors as parametric modulators of the HRF at the time of stimulus presentation. For all models, we included after-sequence questions as separate boxcars modulated by corresponding responses. Automatic orthogonalization was switched off.

GLM1 included the following regressors: RT, drop in RT dummy variable (= 1 if the trial has the largest RT decrease in the sequence of 12 images), position in the sequence (1-12), 4 dummy variables for pattern lengths (2, 3, 4, and 6), a dummy variable for repeated images (= 1 if the previous image in the sequence is repeated), and nuisance regressors (incorrect response, no response, head motion). The goal of this analysis was to contrast the trials with the largest drop in RT to all other trials.

GLM2 included the following regressors: RT, change in RT from the previous trial to the current trial, position in the sequence (1-12), 4 dummy variables for pattern lengths (2, 3, 4, and 6), a dummy variable for repeated images (= 1 if the previous image in the sequence is repeated), and nuisance regressors (incorrect response, no response, head motion). The goal of this analysis was to identify the neural correlates of learning, as indexed by changes in RT.

GLM3 included the following regressors: RT, current trial state-level entropy (E_a), current trial structure-level entropy (E_s), surprise (S_a), position in the sequence (1-12), 4 dummy variables for pattern lengths (2, 3, 4, and 6), a dummy variable for repeated images (= 1 if the previous image in the sequence is repeated), and nuisances regressors (incorrect response, no response, head motion). The goal of this analysis was to identify the main correlates of uncertainty on two levels (state and structure).

GLM4 included the following regressors: RT, change in state-level entropy (E_a), change in structure-level entropy (E_s), position in the sequence (1-12), 4 dummy variables for pattern lengths (2, 3, 4, and 6), a dummy variable for repeated images (= 1 if the previous image in the sequence is repeated), and nuisances regressors (incorrect response, no response, head motion). Changes in entropy were calculated as entropy on the current trial minus entropy on the previous trial. The goal of this analysis was to identify the differential dynamics of entropy-related activity.

Region of interest (ROI) extraction. For ROI illustrations, we extracted raw BOLD signal using the Marsbar toolbox (http://marsbar.sourceforge.net/) and adjusted the signal for 6 s (standard hemodynamic response delay). We used one standard mask from the AAL atlas (for the hippocampus), one ROI defined as a 10 mm sphere around a peak (−27, −54, 45) in the intraparietal sulcus (IPS) found in a previous study on state prediction (Gläscher et al., 2010), one ROI defined as a 10 mm sphere around a peak (−54, 38, 3) in the ventrolateral PFC from another study on state learning (Lee et al., 2014), and one ventromedial prefrontal cortex (vmPFC) mask from a meta-study on value signal (Bartra et al., 2013).

Data and software availability

The data that support the findings of this study are available from the corresponding author upon request.