



ORIGINAL ARTICLE

A neural model of schemas and memory encoding

Tiffany Hwu^{1,2} · Jeffrey L. Krichmar^{1,3}

Received: 21 April 2019 / Accepted: 24 September 2019 / Published online: 4 November 2019
© Springer-Verlag GmbH Germany, part of Springer Nature 2019

Abstract

The ability to rapidly assimilate new information is essential for survival in a dynamic environment. This requires experiences to be encoded alongside the contextual schemas in which they occur. Tse et al. (Science 316(5821):76–82, 2007) showed that new information matching a preexisting schema is learned rapidly. To better understand the neurobiological mechanisms for creating and maintaining schemas, we constructed a biologically plausible neural network to learn context in a spatial memory task. Our model suggests that this occurs through two processing streams of indexing and representation, in which the medial prefrontal cortex and hippocampus work together to index cortical activity. Additionally, our study shows how neuromodulation contributes to rapid encoding within consistent schemas. The level of abstraction of our model further provides a basis for creating context-dependent memories while preventing catastrophic forgetting in artificial neural networks.

Keywords Memory consolidation · Schemas · Catastrophic forgetting · Spatial navigation

1 Introduction

Despite the large amount of information in a dynamic world, humans develop a structured understanding of the environment, learning to recognize scenarios and apply the

This article is part of the Special Issue entitled ‘Complex Spatial Navigation in Animals, Computational Models and Neuro-inspired Robots’.

This material is based upon work supported by the United States Air Force and DARPA under contract no. FA8750-18-C-0103, and other support in part by Toyota Motor North America and HRL Laboratories, LLC. Any opinions, findings and conclusions or recommendations expressed in this material are those of the author(s) and do not necessarily reflect the views of the United States Air Force and DARPA.

Electronic supplementary material The online version of this article (<https://doi.org/10.1007/s00422-019-00808-7>) contains supplementary material, which is available to authorized users.

✉ Tiffany Hwu
hwut@uci.edu

Jeffrey L. Krichmar
jkrichma@uci.edu

¹ Department of Cognitive Sciences, University of California, Irvine, Irvine, CA, USA

² HRL Laboratories, LLC, Malibu, CA, USA

³ Department of Computer Science, University of California, Irvine, Irvine, CA, USA

appropriate behaviors. A longstanding goal in neuroscience is to understand how the brain learns these structures. The stability–plasticity dilemma asks how the brain is plastic enough to acquire memories quickly and yet stable enough to recall memories over a lifetime (Abraham and Robins 2005; Mermillod et al. 2013). A related question is how the brain avoids catastrophic forgetting, which is when a neural network forgets previously learned skills after being trained on new skills (French 1999; Kirkpatrick et al. 2017; Soltoggio et al. 2017). We believe that the brain avoids catastrophic forgetting and balances stability and plasticity by storing information in schemas, or memory items bound together by common contexts.

Our work builds on existing theories of memory and learning, including complementary learning systems (CLS), which states that memories are acquired through rapid associations in the hippocampus and then gradually stored in long-term connections within the neocortex (Kumaran et al. 2016; McClelland et al. 1995). This aligns with hippocampal indexing theory (Teyler and DiScenna 1986), which states that memories in the form of neocortical activation patterns are stored as indices in the HPC, which are later used to aid recall. Tse et al. (2007) later showed that new memories are acquired much more quickly when consistent with a pre-existing schema. In their experiment, rats learned a schema of spatially arranged food wells, each containing a different food. Once familiar with the schema, two new food wells

were introduced. The rats rapidly encoded this new information into the neocortex within a day. Lesion studies showed a dependency on the HPC for this short time span of learning. Further studies by Tse et al. (2011) show activation of plasticity-related genes in the medial prefrontal cortex (mPFC) and related regions, suggesting their involvement in rapid encoding.

Studies of connectivity between the mPFC and HPC have yielded theories of how these areas process schemas. van Kesteren et al. (2012) introduced the SLIMM framework (schema-linked interactions between medial prefrontal and medial temporal regions), which states that when a stimulus is familiar to a schema, the mPFC inhibits the HPC. However, if the stimulus is novel, the HPC activates to encode the information. In this way, the two brain areas enhance memory acquisition via different pathways. As for how the mPFC and HPC communicate, Eichenbaum (2017) hypothesized that theta phase synchronization between the mPFC and MTL is controlled by the thalamic nucleus reuniens (Re) to change the directional flow of information when encoding or retrieving. As theta oscillations control many aspects of timing and control in the HPC, they may play an important role in schema learning. Furthermore, Eichenbaum (2017) notes the different roles of the ventral hippocampus (vHPC) in providing contextual cues and the dorsal hippocampus (dHPC) in retrieving specific memories. Since the specificity of encoding increases along the dorsal–ventral axis of the HPC (Jung et al. 1994), the HPC may be representing memories at multiple levels of specificity, from general context to specific episode.

In addition to the HPC–mPFC pathways, supporting areas modulate the speed of encoding. The neuromodulatory system is important for detecting salience and making quick adaptations (Krichmar 2008). The basal forebrain (BF) modulates attention and cortical information processing (Baxter and Chiba 1999), and the locus coeruleus (LC) modulates network activity in response to environmental changes (Aston-Jones and Cohen 2005). The LC also drives single-trial learning of new information via inputs to the HPC (Wagatsuma et al. 2018). Moreover, activity from the LC selectively increases oscillations in the HPC in the theta and gamma range according to novelty (Berridge and Foote 1991; Walling et al. 2011), suggesting that it may be involved in single-trial learning.

This type of processing is likely necessary for the fast learning that occurs when novel stimuli are introduced within a familiar schema. Based on this background, we created a neural network architecture with the simulated brain regions and pathways comparable to those described above. The network portrays the involved neurobiological mechanisms at a level of abstraction that transfers easily to addressing the problem of avoid catastrophic forgetting in artificial neural networks.

2 Methods and tasks

2.1 Summary of Tse et al. (2007)

In Tse et al. (2007), rats were trained for 20 days on a schema consisting of paired associations between food well location and food flavor. After training, the rats were well familiarized with the schema, as measured by dig time in the correct well corresponding to a cued food. This is referred to as the preexisting schema. On the 21st day, two food wells were removed and replaced with two new food wells with new foods, close to the locations of the original wells. These are referred to as the new paired associations within a preexisting schema. The rats learned the locations of the two new food wells rapidly, suggesting that the presence of the wells within a familiar schema increased the rate of learning. After this, the HPC was lesioned in one group of rats, and the two new food wells were replaced with yet another two food wells. Without the HPC, the rats were unable to learn this second set of new food wells. However, they were still able to retain the original schema as well as the first set of novel food wells. In the next task, the same set of rats was trained on a second schema, which is referred to as a novel schema, with entirely new food well placements and food flavors. The group with the HPC lesion was unable to learn the second schema, but still retained the original information. The control group was able to learn the second schema while maintaining the first, despite the similarity of tasks. We use these experiments as a baseline test of whether our model is able to capture the same behaviors of schema-consistent learning. A summary of these tasks is described in Fig. 1.

In both Tse et al. (2007) and tests of our model, a schema is a spatial arrangement of food wells. The presence of novelty within a familiar schema occurs when some of the food well and flavor paired associations are replaced with food wells in new locations with new flavors. Some of the terminology in our experiments differs from the original rat experiments. The 24-h periods of training in Tse et al. are referred to as “trials” in our experiment. While each 24-h period in Tse et al. only presents each paired association once officially, the trials in our experiment select paired associations at random for several hundred epochs in one trial, representing interleaved replays of episodic memories. There is no distinction between waking and sleeping activity in our model, instead modeling replays as general replay during quiet periods.

2.2 Contrastive Hebbian learning

To model representations of tasks, a multilayer network can store information of increasing levels of abstraction from input to output layers. Backpropagation is commonly used to train such networks and has had many successful applications in artificial neural networks (LeCun et al. 2015; Rumelhart

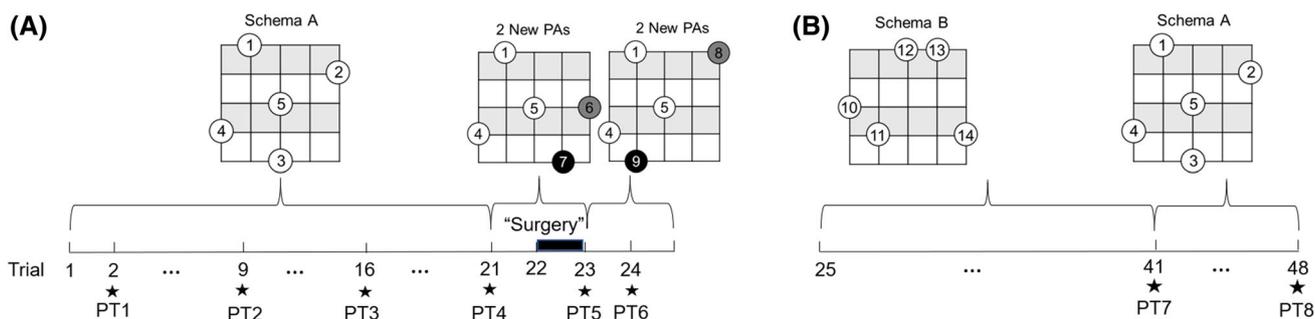


Fig. 1 Timeline of experiments. Each grid represents a square arena, with circles representing locations of food wells and numbers representing different flavors of food. **a** Experiment 1 timeline. Schema A is trained for 20 days, which are called trials in our model. Stars represent days in which a probe test (PT) was performed. Two new paired asso-

ciations (PAs) are introduced on day 21, represented by black and gray circles. Surgical lesioning occurs on day 22, and yet another two PAs are introduced on day 23, also represented by black and gray circles. Probe tests are performed throughout. **b** Experiment 2 timeline. Schema B is trained for 16 days; then Schema A is retrained for 7 days

et al. 1986). However, many view backpropagation as biologically implausible, as there is no widely accepted mechanism in the brain to account for sending error signals backwards along one-way synapses. An alternative account for developing representations in the brain may be contrastive Hebbian learning (CHL) (Movellan 1991), which uses a local Hebbian learning rule and does not require explicit calculations of an error gradient.

Given a multilayered network with layers 0 through L , neuron activations of the k th layer are denoted as vector x_k and weight matrices from the $k-1$ to k th layer are denoted as W_k . Each weight matrix W_k has a feedback matrix of γW_k^T such that every weight has a feedback weight of the same value but scaled by γ . The learning process consists of cycling between three modes of the network. The first mode is known as the free phase of the network, in which the input layer x_0 is fixed and the following equation is applied to update neurons in layer k from $k = 1$ to L at time t :

$$x_k(t) = f_k(W_k x_{k-1}(t-1) + \gamma W_{k+1} x_{k+1}(t-1)), \quad (1)$$

where f is any monotonically increasing transfer function. This equation is applied for T_s time steps, which is when network activity converges to a fixed point. The resulting settled activity for x_k is noted as \check{x}_k , which is the final neural activity for the free phase. The network then transitions to the clamped mode, in which the input layer is fixed as before and the output layer is fixed to the desired target value, as in supervised learning. Again neuron activities are updated using Eq. 1 for T_s time steps which allows the network activity to converge. The settled activity for x_k is noted as \hat{x}_k , which is the final neural activity for the clamped phase. The third mode combines an anti-Hebbian update rule for neurons in the free phase and Hebbian update rule for neurons in the clamped phase:

$$\Delta W_k = \eta_{\text{CHL}} (\hat{x}_k \hat{x}_{k-1}^T - \check{x}_k \check{x}_{k-1}^T), \quad k = 1, \dots, L. \quad (2)$$

Although the local learning rule is more biologically plausible than backpropagation, the use of symmetric weights is viewed as less plausible. To address this, versions of CHL that do not use symmetric weights have been implemented. For instance, Detorakis et al. (2018) provide feedback signals via random feedback weights during the clamped mode, avoiding any use of bidirectional symmetric weights. However, we use the original version of CHL for simplicity.

3 Neural model

Our model consists of two main information streams, the indexing stream and the representation stream, in a network that is trained on context-dependent tasks such as the one found in Tse et al. (2007). Figure 2 shows an overview of the network.

3.1 Indexing stream

The indexing stream begins with a context pattern, which can be any encoding of context using patterns of neuron activity. In our case, we used a 2D grid with inputs of 1 if a food well existed in that grid location or 0 otherwise. This input projects to the mPFC for the schema to be learned from sensory input. The dynamic mPFC neuron activity is calculated by the following synaptic input equation at time t :

$$x_k(t) = f_k(W_k x_{k-1}(t-1)), \quad (3)$$

where layer k is the mPFC layer, layer $k-1$ is the context pattern, and f_k is the rectified linear unit (ReLU) transfer function:

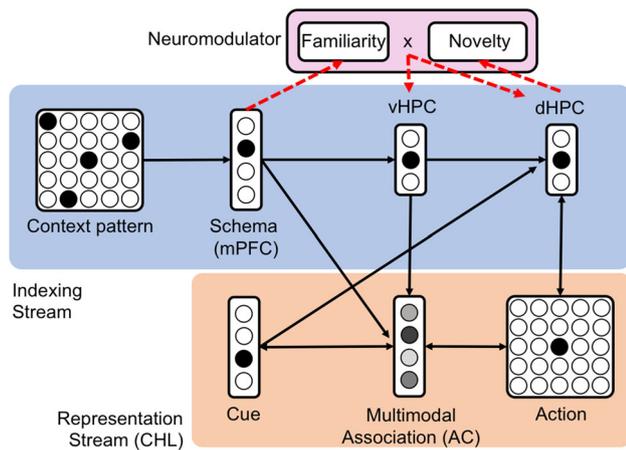


Fig. 2 Overview of network. The light blue box contains the indexing stream, including the ventral hippocampus (vHPC) and dorsal hippocampus (dHPC). The light orange box contains the representation stream, including the cue, medial prefrontal cortex (mPFC), multimodal layer (AC), and action layer. Bidirectional weights between layers in the representational stream are learned via contrastive Hebbian learning (CHL). Weights from the indexing stream are trained using the standard Hebbian learning rule. Dotted lines indicate influences of the neuromodulatory area, which contains submodules of novelty and familiarity. Weights extend to these modules from the mPFC and dHPC. Neuro-modulator activity impacts how often the vHPC and dHPC are clamped and unclamped while learning the task via contrastive Hebbian learning (CHL)

$$f(x) = \max(x, 0). \quad (4)$$

A hard winner-take-all selection is then applied, in which all activations are set to zero except for the one with maximum value. Weights from the context pattern to the mPFC are then trained by the standard Hebbian learning rule:

$$\Delta W_k = \eta_{pattern} x_k x_{k-1}^T, \quad (5)$$

where x_k is the mPFC layer, $\eta_{pattern}$ is the learning rate, and x_{k-1} is the context pattern layer. The weights are normalized such that the norm of weight vectors going to each postsynaptic neuron i is 1:

$$w_i = \frac{w_i}{\|\mathbf{w}\|}, \quad (6)$$

where \mathbf{w} is the vector of weights going to one postsynaptic neuron and w_i is an individual weight in \mathbf{w} . The indexing stream continues on to the dHPC and vHPC. The vHPC learns an index of mPFC activity using the same synaptic input function and learning rule described in Eqs. 3–6. The dHPC learns indices in the same way as the vHPC, except that it uses a learning rate of $\eta_{indexing}$ and has weights from the vHPC, cue, and action selection layer. The term $\eta_{indexing}$ is a separate learning rate used for weights that index activity from the representation stream. Rather than indexing context,

the dHPC indexes triplets of context, cue, and action. This agrees with how context is encoded in the vHPC and specific experiences are encoded in the dHPC in Eichenbaum (2017). It also aligns with the fact that selectivity of encoding increases from the ventral to dorsal end of the HPC (Jung et al. 1994).

3.2 Representation stream

The representation stream is a multilayered network with sensory cue input areas and the mPFC that encodes the current schema or context. The middle layer of the representation stream acts as the association cortex (AC) and makes multimodal associations of its inputs and conjoins context and cue information (AC). The output layer selects an action response to the sensory cue, which is a spatial grid with activation of 1 at the locations of food cues and 0 otherwise. To train the correct actions, the multilayered network uses CHL as described in the background section, with a transfer function as in Eq. 4. The alternation of clamped and free phases is controlled by the indexing stream. The dHPC alternates between clamping and unclamping the action layer, providing input to the action layer during the clamped stage of contrastive Hebbian learning. During clamping, the winning neuron in the vHPC gates neurons in the AC. This is done via a static weight matrix of strong inhibitory weights from the vHPC to AC layer, with sparse random excitatory weights that allow only some neurons in the AC to be active. All weights in this matrix are first initialized by a strong negative value of w_{inh} , and then a random selection of the weights in this matrix is set to 0. The number of weights selected to be 0 is determined by P , which is a number in the range of 0 and 1 that defines the proportion of randomly selected weights. This is meant to mimic hippocampal indexing. While there is little evidence that the HPC projects widespread inhibition to the representation areas of the brain, the mix of inhibition and zeroed weights allows specified patterns of neurons to be active with their usual activity levels, which is meant to mimic an attentional sharpening effect in the model.

Taken as a whole, the representation stream models how the neocortex learns representations (Hawkins et al. 2017), with the indexing stream driving the learning process and preventing catastrophic forgetting by allocating different sets of AC neurons for each task. By using CHL for the representation stream, we form an equivalence with backpropagation methods that allows us to expand our model to help improve traditional neural networks in the future.

3.3 Novelty and schema familiarity

In the SLIMM framework, the encoding strength combines schema familiarity and cue novelty. Novelty is defined as how infrequently a stimulus has been experienced before, whereas

schema familiarity is how frequently a schema has been experienced. Furthermore, the SLIMM framework proposes that resonance occurs in the presence of schema familiarity. However, SLIMM also suggests that the mPFC inhibits the HPC, whereas we suggest that a combination of schema familiarity and novelty from the mPFC and dHPC, respectively, affects learning by controlling oscillatory activity in the HPC.

In our model, a neuromodulatory area detects novelty in the presence of a familiar schema and modulates the strength of learning that occurs within the representation stream. To detect novelty, each neuron in the dHPC projects to a novelty submodule with $w_{novelty}$ as the starting weight. $w_{novelty}$ represents the baseline level of surprise when a new stimulus is presented. Whenever the activity of the dHPC is updated, the activity of the novelty submodule is the rectified weighted sum of inputs from dHPC after applying winner-take-all, as in Eqs. 3 and 4. The weights from the dHPC to the novelty submodule are then updated with an anti-Hebbian learning rule:

$$\Delta W = -\eta_{indexing} x_{novelty} x_{dHPC}^T, \tag{7}$$

where W is the matrix of weights from the dHPC to novelty submodule, $\eta_{indexing}$ is the learning rate, $x_{novelty}$ is the activity of the novelty submodule, and x_{dHPC} is the activity of neurons in the dHPC. Therefore, the weight between an active dHPC neuron and novelty submodule will experience long-term depression and decrease the novelty score of a stimulus after each epoch of training. Since the dHPC uses winner-take-all, each weight from dHPC represents an individual novelty score for each possible triplet. The activity of the familiarity module, $x_{familiarity}$, is similarly calculated through the weighted sum of inputs from the mPFC after winner-take-all, as in Eq. 3. However, rather than a rectified linear unit, we use a shifted sigmoidal transfer function:

$$f(x) = \frac{1}{1 + e^{-s(x-x_{shift})}}, \tag{8}$$

where s is the sigmoidal gain and x_{shift} is the amount of input shift. The shifted sigmoidal transfer function ensures that the familiarity module requires a baseline amount of training on a schema to be considered familiar with it and that familiarity does not continue to increase in an unbounded manner with extended exposure. The activity of the familiarity module is thus mostly bimodal, with a very low activity if the schema is unfamiliar and a high activity if the schema is familiar. The effect of the sigmoidal function will be seen in “Results.” The weights from the mPFC start with the same value of w_{fam} , which is a very small value close to zero that represents low familiarity of schemas prior to training. These weights are updated after mPFC activity is updated, using the Hebbian learning rule from Eq. 5 with $\eta_{pattern}$ as the learning

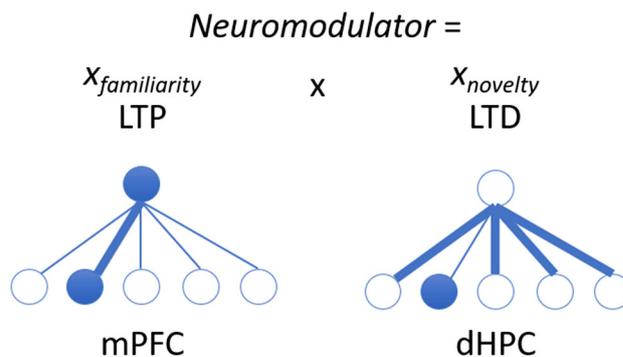


Fig. 3 Explanation of neuromodulation in the model. Neuromodulator activity is the product of familiarity and novelty. Familiarity comes from long-term potentiation (LTP) of weights from the mPFC. As a schema is exposed to the network multiple times from each epoch of training in a trial, the mPFC neuron encoding that schema potentiates its own weight to increase the familiarity. Novelty works in a similar way with the dHPC, except that weights start as high values and undergo long-term depression (LTD) with repeated exposure

rate. $\eta_{pattern}$ is small to model the long-term consolidation of schemas. The neuromodulator activity is then set to the simple product of activity from the familiarity and novelty modules:

$$\text{neuromodulator} = x_{novelty} * x_{familiarity}. \tag{9}$$

See Fig. 3 for a visual explanation of how this value is calculated. This value determines the number of times the vHPC and dHPC will clamp and unclamp the representation layer in a single trial and thus determine the number of extra epochs in a trial that are added to a default number of epochs, $e_{default}$:

$$\text{epochs} = e_{default} + \text{neuromodulator} * e_{boost}. \tag{10}$$

The model therefore suggests that the rapid encoding that occurred in Tse et al. was due to increased replay of important information during quiet waking periods. Rewarding experiences are known to be prioritized in hippocampal replay (Mattar and Daw 2018; Pfeiffer and Foster 2013). We reflect this by using neuromodulator levels to determine the number of training epochs. In this case, it is not just the rewarding experiences that get replayed more, but also experiences that are novel and familiar. This also follows the idea from the SLIMM framework, suggesting that resonance occurs in the presence of schema familiarity. In the SLIMM framework, resonance is defined as the neural co-activity of two or more brain regions. The resonance in our model does not occur within the mPFC, but instead occurs with the HPC via the repeated clamping and unclamping of layers in the representation stream. As in the case of the rat experiments, the replay of novel experiences within familiar schemas leads to faster acquisition of reward. Since hippocampal replay involves the firing of cell sequences

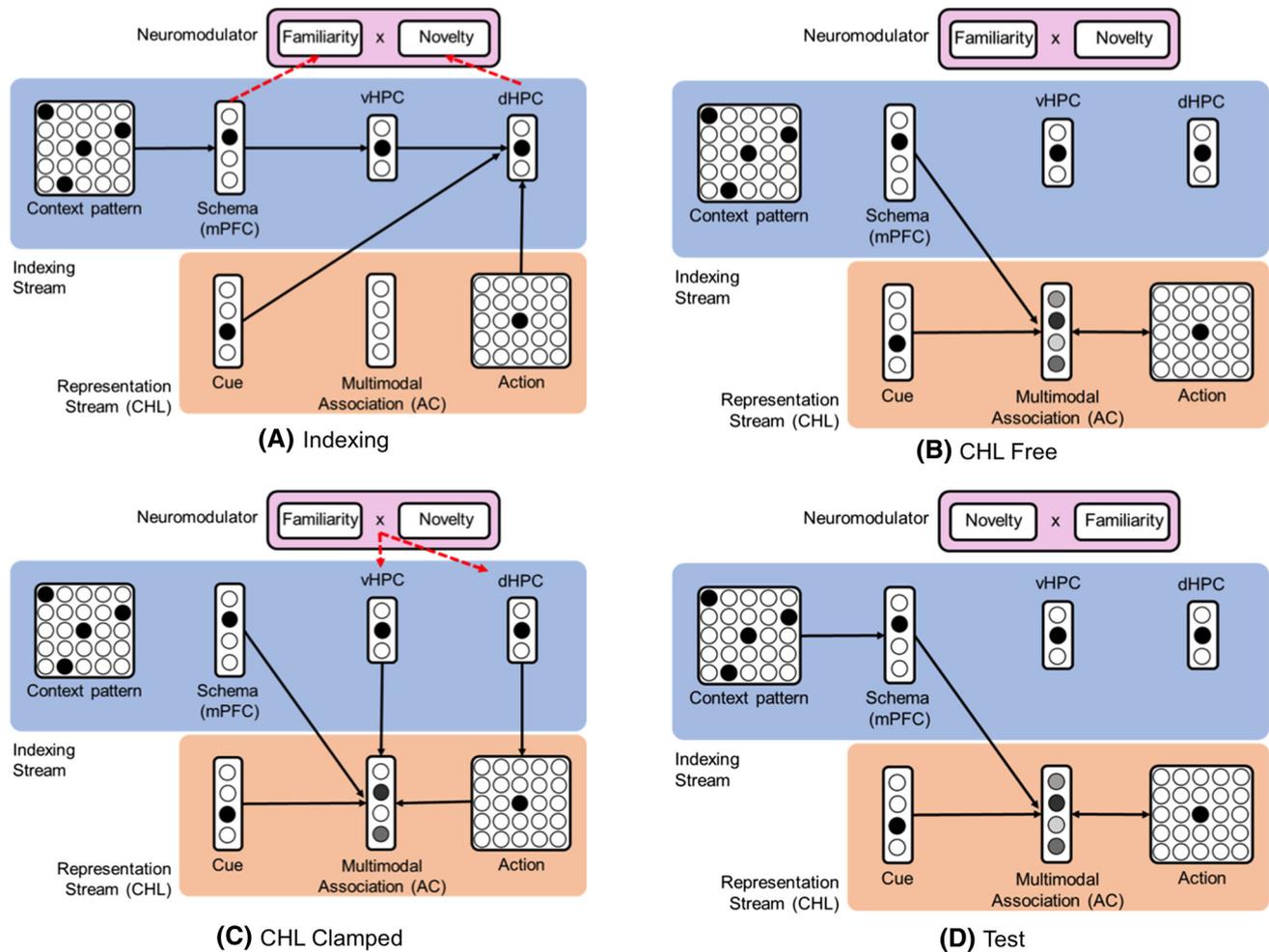


Fig. 4 The four phases of a trial of training. **a** In the indexing phase, the indexing stream forms indices of activity. Additionally, the activities of the novelty and familiarity modules of the neuromodulator are calculated, setting the ultimate activity of the neuromodulator to the product of these values. **b** The free phase of CHL. **c** The clamped phase

of CHL. After this phase, the CHL update rule is applied. **d** The test phase of a network for measuring performance during training and unrewarded probe tests. No learning occurs in this phase, and weight values are static. This is the same as the CHL free phase, except that mPFC activity is also calculated

in the hippocampus, we expect indexing to occur at this time. While it is less plausible that the cue and context patterns would continue to receive input as depicted in the model, these areas may also be reactivated during sleep and quiet waking replays by connections via the hippocampus.

When training on a task, the network runs trials of training that consist of four phases as shown in Fig. 4. Each trial consists of many epochs of training, the number of which is determined by Eq. 10. This mimics the high resonance caused by schema familiarity, as postulated in the SLIMM framework. Each epoch consists of an indexing phase, a CHL free phase, and a CHL clamped phase. During indexing, the mPFC, vHPC, and dHPC form indices using the unsupervised Hebbian rule and winner-take-all rule described previously. During the free phase (Fig. 4b), the

representation stream runs freely. During the clamped phase (Fig. 4c), the representation stream is clamped to input from the dHPC using Eq. 3. After free and clamped phases, the CHL learning rule is applied. Also during clamping, the AC receives input from the vHPC also using Eq. 3, which effectively inhibits most of the AC neurons except for the few that have a weight of 0 from the winning vHPC neuron. The clamping function of random inputs to the representation stream reflects the specific role of HPC as a sparse indexer of neocortical activity. As opposed to modeling the clamping of the AC using another part of the network such as the mPFC, the HPC requires an alternation between clamped and unclamped states, whereas the mPFC activity stays constantly clamped during CHL training. At the beginning of a trial, the number of training epochs is undetermined, but tentatively set at e_{settle} epochs (Fig. 4a). During

```

numTrials = 20 numEpochs = 20;
nmMax = 0;
i = 0;
trial = 0;
while trial < numTrials do
  while i < numEpochs do
    Pick random PA from schema;
    Indexing Phase (Eqns 3,4,8,9);
    if neuromodulator > nmrMax then
      | nmrMax = neuromodulator;
    end
    Indexing Weight Update (When training, Eqns 5,6,7);
    CHL Free (Eqn 3);
    CHL Clamped (Eqn 3);
    CHL Weight Update (When training, Eqn 2);
    if i == esettle then
      | numEpochs = edefault + eboost * nmMax (Eqn 10);
    end
    i++;
  end
  trial++;
end

```

Algorithm 1: Schema Training

this time, activity levels of the neuromodulator are tracked, and the maximum neuromodulator activity found within this period is used to determine the ultimate number of training epochs within the trial. After each trial, the performance of the network is measured during the test phase (Fig. 4d) by presenting a cue to the network and allowing the network to settle on an action. For a pseudocode description of the schema training process, refer to Algorithm 1. The network parameters used in our network are listed in Table 1.

Table 1 Parameters used in experiment

Population sizes		Learning parameters	
N_{cue}	18	T_s	5
N_{mPFC}	10	$\eta_{indexing}$.1
$N_{multimodal}$	40	$\eta_{pattern}$.0001
$N_{context}$	25	η_{CHL}	.001
N_{vHPC}	5	g	.001
N_{dHPC}	40	e_{boost}	1000
		$e_{default}$	600
		e_{settle}	20
		w_{min}	0.3
		w_{max}	0.8
		w_{inh}	-10
		w_{fam}	.0001
		$w_{novelty}$	1
		s	200
		x_{shift}	.03
		P	0.3

3.4 Experimental design

We validated our model by simulating the experiments in Tse et al. (2007). We simulated a population of 20 rats. For each individual, weights from vHPC to AC were sparsely connected with a probability P and set to w_{inh} . Weights to the novelty and familiarity module were all set to $w_{novelty}$ and $w_{familiarity}$, respectively. Remaining network weights were initialized along a uniform distribution on the range w_{min} and w_{max} , fully connected. The arena was discretized into a 5x5 grid, corresponding to location in the arena for the context pattern and action layers. Each trial consisted of multiple epochs to account for replays during sleep and quiet waking periods. In Tse et al. (2007), performance was measured during training by counting the errors in a trial, and non-rewarded probe tests (PT) were performed intermittently by recording the amount of time spent searching the correct well when given a food cue. In our model, our test of performance was to present each flavor in a schema during the test phase. Upon presenting the cue, the network ran in free phase until the activity converged. In the action selection layer, the activity of the neuron corresponding to the correct location of the food was divided by the sum of all of the neurons corresponding to the wells in the arena. This value corresponds to the amount of time a rat would spend digging in a well given a food cue. We used this value to simulate performance during both training and unrewarded probe tests. In the Tse et al. (2007) experiment, the unrewarded probe test was conducted with a food cue and no food reward in the correct well, thus limiting learning for those trials. Although our experiment does not model reward representation in the brain, we approximate unrewarded probe tests by not updating weights after cueing and running the network.

A timeline of the two replicated experiments is shown in Fig. 1. The first experiment was to train the network on Schema A layout for 20 trials with paired associations (PAs) between food cue and location (Fig. 1a; PTs 1–3). On the 21st trial, two of the PAs were replaced by two new ones (Fig. 1a). The original schema with two new PAs was trained for one trial, and then a probe test was performed by cuing one of the new foods and examining the output of the well locations of the new cued PA, new non-cued PA, and other wells (Fig. 1a; PT 4). After that, the network was split into a control network and an HPC-lesioned group. The HPC-lesioned group was copied with weights from the original network and had all connections to and from the vHPC and dHPC removed. Another probe test was performed to see whether both groups could still recall the original schema as well as the schema with two new PAs (Fig. 1a; PT 5). Next, the two new PAs were replaced with yet another two new PAs and trained on both groups for one trial. Another probe test was performed after this (Fig. 1a; PT 6). The second experiment was performed on the resulting networks from the

first experiment (Fig. 1b). For 16 trials, both conditions were trained on an entirely new schema, Schema B, and a probe test was conducted (Fig. 1b; PT 7). After this, the groups returned to train and test on the original schema for 7 days, with yet another probe test at the end (Fig. 1b; PT8).

3.5 Code availability

The code used to produce our results can be accessed at <https://github.com/fitany/SchemaNetworkBICY/>.

4 Results

In our experiment, we modeled the selection of locations associated with food wells. To perform the location selections, we presented an odor cue to the network (see Cue area in Fig. 2), and chose a location by selecting the action neuron with the highest activation level (see action area in Fig. 2). Since the trajectories were not models, we did not need to designate starting locations for the simulated rats.

To measure the model's performance, we compared the samples of "dig time" for cued and non-cued wells. A cued well was defined as the well containing the food cued at the beginning of the trial. In tests where two new wells were introduced, the non-cued well was defined as the new well that did not contain the cued food. This was to examine whether the model simply preferred novel food wells, as opposed to the correct food well. Original wells were defined as the wells that existed in the original schema, before the new PAs were introduced. To test the differences between HPC-lesioned and control groups, we compared the cued samples of both groups. While our numerical results varied from the rat experiments, all major trends and key findings of our model were consistent with Tse et. al.

Statistical significance was determined through the Wilcoxon rank sum test with Bonferroni's correction. For performance results involving cued versus non-cued dig time, we compared cued and non-cued samples. Unless otherwise noted, all findings are significant at $p < .001$.

4.1 Experiment 1

The goal of the first experiment was to show that new information matching a schema can be quickly learned and that the HPC is necessary for this learning. Figure 5a shows that the model was able to gradually learn Schema A over 20 trials of training. Figure 5b confirms this with probe tests, which show the proportion of neuron activity corresponding to the correct food location given a cue. Figure 5c shows the probe test results after training for one day on Schema A with two new PAs added. During the test phase, the longer dig time in the correct well shows that the new PAs were learned in just

one trial, with much more dig time in the cued wells than non-cued wells. This matches the finding in Tse et al. that finding that new information consistent with a preexisting schema is learned rapidly. The Tse et al. experiment showed approximately a 200% increase in dig time of the cued wells versus non-cued wells, which is equivalent to the results of our model.

Next, the role of the HPC was examined. In Fig. 5d, e, upon splitting into two conditions of HPC-lesioned and control, both groups were still able to recall the original schema, as well as the two new PAs. This suggests that the new information had been consolidated within a short period and no longer required the HPC for retrieval. This was due to the neuromodulator detecting when new information was present within a familiar schema and increasing the rate of learning accordingly. When training yet another two PAs on both conditions in Fig. 5f, the HPC-lesioned group was unable to learn at all, while the control group was still able to learn. This suggests that the HPC was responsible for driving the index-based clamping and unclamping of the output layer, as CHL mechanism was unable to update the weights properly. The same finding of HPC-dependent learning was observed in the original rat experiments. However, the original rat experiments showed dig times of cued, non-cued, and original wells at chance level, whereas our model shows dig times of far less than chance, still largely preferring the original food wells.

Figure 6 shows the weights of the network after training on the first experiment. Our results for the first experiment were able to capture most of the effects seen in Tse et al. (2007). We were able to show that information is acquired rapidly when consistent with a prior schema and that the HPC is necessary for acquisition.

4.2 Experiment 2

The purpose of the second experiment was to test whether multiple schemas could be learned by the same network and whether the HPC was necessary for this. Figure 7A shows that the control group was able to learn Schema B. However, the HPC-lesioned group was unable to learn, staying at chance levels the entire time. This is shown again in the probe tests in Fig. 7b. When the HPC-lesioned model returned to retrain on Schema A in Fig. 7c, it had good performance the entire time and retained the information learned prior to surgery, but did not improve performance beyond what it had learned in Experiment 1. The control group displayed a minute decrease in performance of Schema A at the beginning, but quickly regained prior performance. The decrease was due to small overlaps in the gating patterns from vHPC. This is confirmed again in the probe tests in Fig. 7d.

Figure 8 shows the weights of the network in the control condition after training on the second experiment. The results show that our network is able to learn multiple schemas in

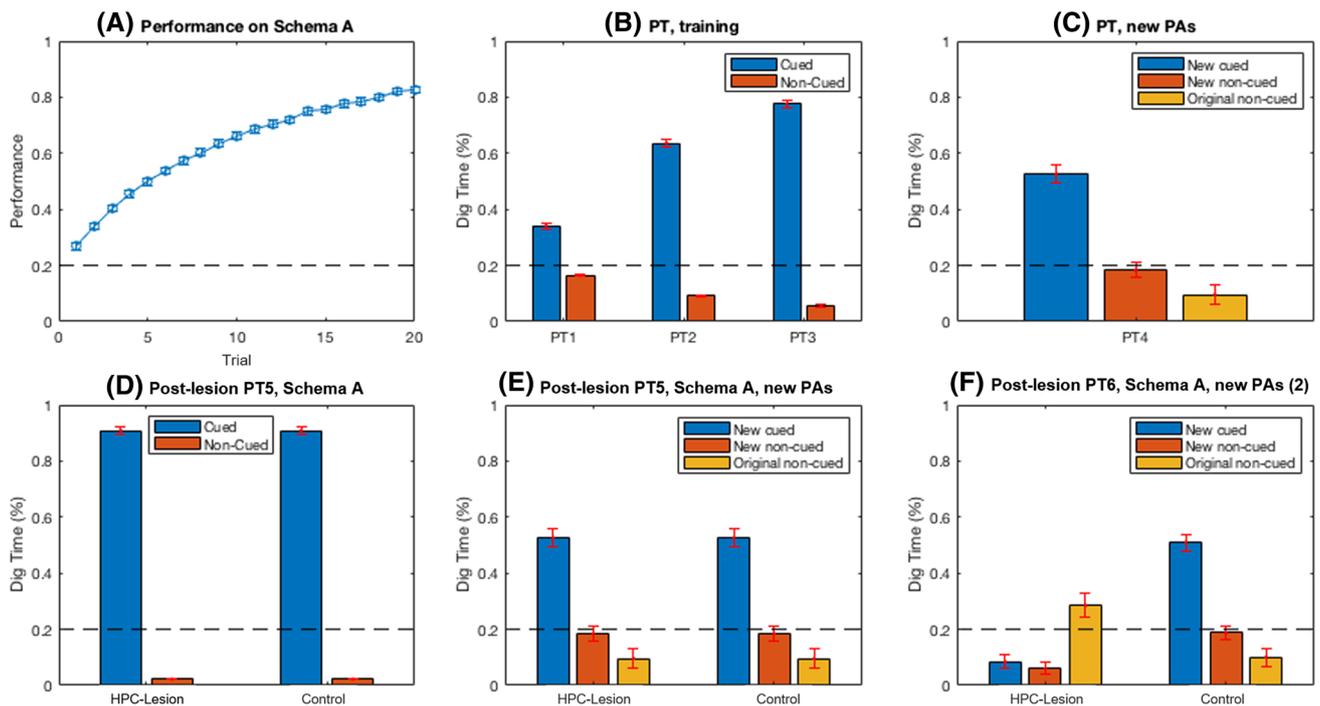


Fig. 5 Results of replicating the first experiment of Tse et al. Our model equates dig time in a food well from the original Tse et al. experiment with the proportion of activity in the action neuron corresponding to the location of the correct food well after cueing the network with a food flavor. **a** The dig time performance over 20 trials, showing a gradual increase. **b** Probe tests of trials 2, 9, 4, and 16, showing the proportion of activity of the correct well given a food cue compared to activity of the incorrect wells. In the case of our simulations, these values are the same as in **a**, but at specific trials to best correspond with the original rat experiments. **c** Probe test after training the new PAs for 1 day, in

which one of the new pairs is cued and the activities of the correct well, well of the other new pair, and original Schema A wells are compared. The new PAs were learned within one trial, as the dig time for the cued pair was significantly higher than the rest. **d** Probe tests of Schema A after splitting into HPC-lesioned and control groups. Both conditions retained knowledge of the schema. **e** Probe tests of the new PAs after splitting into HPC-lesioned and control groups. Both groups recalled the new PAs equally. **f** Probe tests of Schema A after training another two new PAs. The HPC group could not learn

succession, without forgetting prior schemas. We were thus able to match the effect seen in Tse et al. (2007). A small difference is that our network retained information about Schema A much better than their experiments for both of their HPC-lesioned and control groups. While the rat experiments had dig times of 50 percent for the correct well in HPC-lesioned and control conditions, our model maintained performance of closer to the original 90% from before training on the new schema.

4.3 Neural activity

To gain a better understanding of the network activity, we plotted the neuron traces of the mPFC, vHPC, and dHPC before winner-take-all was applied. Figure 9 shows the neuron activities for the first two experiments, with Schema A and new PAs for Experiment 1 and Schema B with a retraining of Schema A for Experiment 2. Each colored line in the figure represents the activity of a single neuron in a single simulated rat. The black vertical lines separate epochs into

trials. The neuron traces show that the mPFC chooses a single neuron to represent Schema A and continues to strengthen its weights as it is trained. When new PAs are introduced, the same neuron is still active, but decreases in activity. When Schema B is introduced at the beginning of Experiment 1, a different neuron becomes active, and when Schema A is reintroduced immediately after training Schema B, the original winning neuron returns. By viewing the spacing between black vertical lines, we see that the large spacing for trials with new PAs indicates that more epochs of training occur during those times. This is due to the novelty and familiarity detection from the neuromodulator. vHPC activity in Fig. 9B shows the same effect, except that the weights of all of the neurons together increase and decrease, as they are all affected by the rising and falling activity levels of mPFC. dHPC works similarly, as shown in Fig. 9, although different winners are chosen at every epoch. For the dHPC, we display neural traces for the first 100 epochs in the first trial epochs, due to the frequent switching of winners.

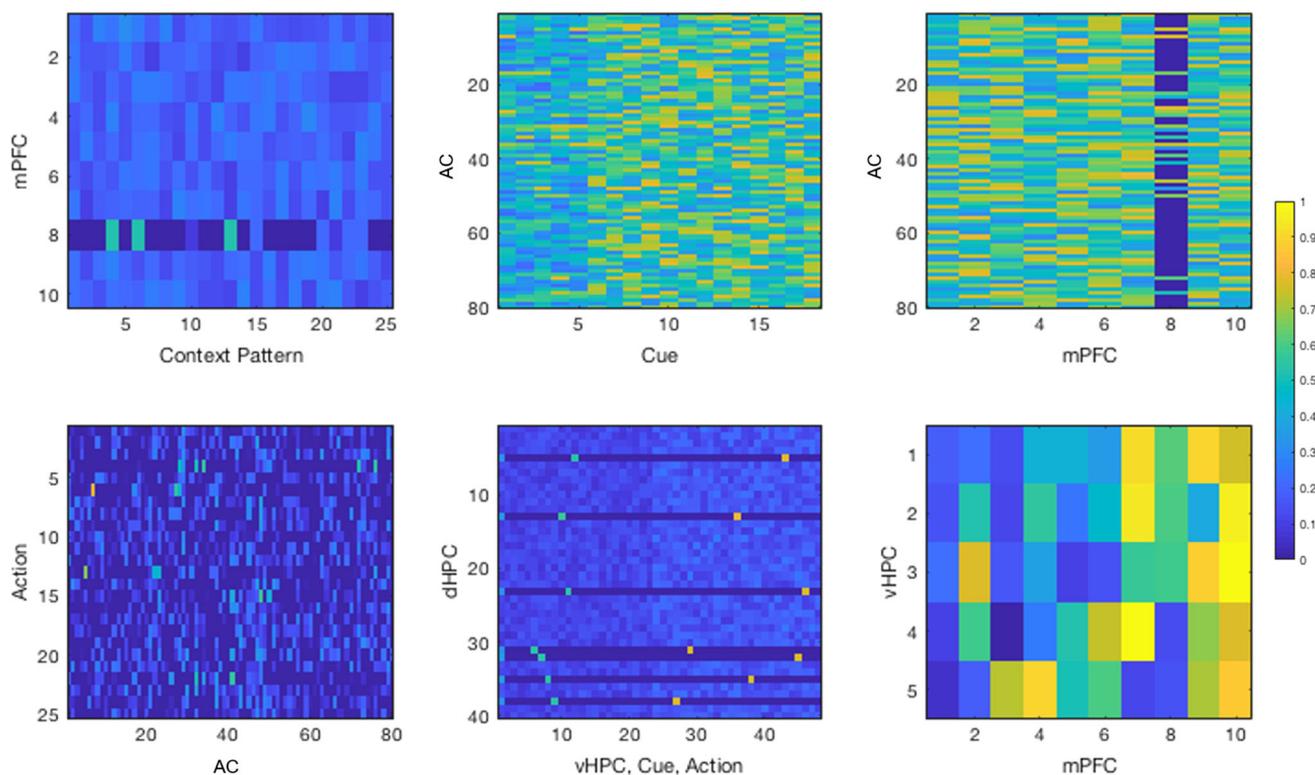


Fig. 6 Weight matrices of the network of one simulated rat after simulating all of the first experiment. Rows represent postsynaptic layers and columns represent presynaptic layers. Weights from the context pattern to mPFC show the development of a distinct schema pattern encoded with stronger weights where the wells are located. Weights from the mPFC to the AC show the effect of the gating, in which the mPFC neuron representing Schema A is associated with a set of neurons in the AC. Weights from the AC to action layer show that the

actions are dependent on activity from select AC neurons, suggesting that the AC neurons are learning specific features useful for action selection. Weights going from the vHPC, cue, and action to the dHPC are displayed in one matrix to show clear encodings of context, cue, and action triplets with one neuron from each of the three areas. Weights from mPFC to vHPC show that there is not necessarily a one-to-one mapping from a winning mPFC neuron to a vHPC neuron, but that the schema information is transferred in a distributed manner

To show the effects of schema familiarity and novelty on neuromodulator activity, Fig. 10 displays the activities of the familiarity module, novelty module, and neuromodulator over the first 21 trials of Experiment 1. Due to the Hebbian learning of connections from the mPFC to the familiarity module, familiarity increases over time. Since it uses a sigmoidal transfer function, the increase is step-like, increasing from 0 to 1 swiftly. Novelty starts high when Schema A is introduced, and quickly drops to 0 due to the anti-Hebbian learning rule. When two new PAs are introduced at trial 21, novelty returns to a high state. Taking the product of novelty and familiarity, the activity of the neuromodulator increases only when familiarity and novelty are high. Since the activity of the neuromodulator is proportional to the number of training epochs in a trial, this leads to the desired behavior of increased learning when new PAs are introduced to an existing schema.

We also tracked the activity of the neuromodulator for all experiments. Figure 11A shows the number of training epochs in each trial of the first and second experiments. As

a reminder, each epoch within a trial consists of an indexing phase, a free phase, and a clamped phase. The first 21 trials of the first experiment are on the left of the black vertical line, and the remaining trials to the right of the vertical line are for the second experiment. Each individual colored line represents the neuromodulator activity of a single run, with 20 runs total. When new PAs were introduced on trial 21 of the first experiment, the familiarity of Schema A multiplied by the novelty of the new stimuli caused a sharp spike in neuromodulator activity, increasing the number of epochs for those trials. The neuromodulator activity for the second experiment remained low, as there were no new PAs introduced. Figure 11B shows the index of the winning mPFC neuron in each trial, with a different colored line for each individual run. Every time the schema changes, the index of the winning neuron changes accordingly.

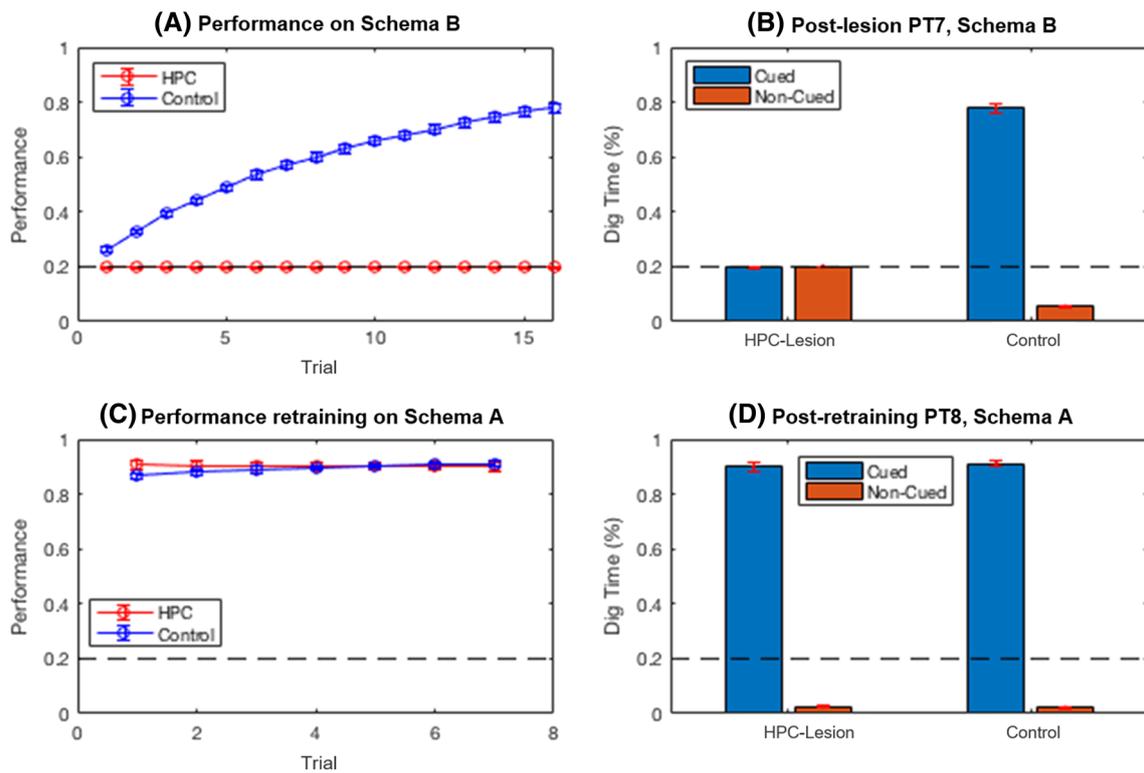


Fig. 7 Results of replicating the second experiment of Tse et al. **a** The performance over 16 trials of training Schema B. The control group was able to learn the new schema, while the HPC group was not. **b** Probe test after training Schema B, confirming part A. **c** Performance over

seven trials of retraining on Schema A. Both groups retained Schema A, though the control group recovered from a very minimal forgetting of the schema initially. **d** Probe test after retraining on Schema A, confirming part C

4.4 Effects of neuromodulation

To study how the neuromodulator influences learning, we removed connections to and from the novelty and familiarity modules in the network. Rather than boosting the number of epochs using neuromodulator activity, we used a constant number of epochs for each trial. We repeated the first part of Experiment 1 with the same number of epochs for every trial, trying different values for the constant number of epochs. As before, the network was trained on Schema A for 20 trials and training a new PA was introduced on the 21st trial. As shown in Fig. 12, training with more epochs per trial leads to faster learning and better performance on the new PA. It is therefore not required to have a neuromodulator for rapid learning of new information. However, the number of total epochs over all trials can be greatly reduced if the model is able to detect novelty within the schema and adjust the learning accordingly. The smaller number of epochs reduces the overall learning time of the experiment. Compared to the conditions with a flat number of epochs per trial, we found that our original network with the neuromodulator had better performance on the new PAs than the other successful

conditions, despite having fewer total epochs. It was therefore able to conserve network training time overall.

4.5 Effects of HPC sparsity on catastrophic forgetting

We hypothesized that catastrophic forgetting could be avoided due to the sparsity of weights between the vHPC and AC. Therefore, we investigated the effects of varying the sparsity, from a value of $P=0$, which would be effectively lesioning all connections, to $P=1$, a dense all-to-all connectivity. We trained the network on Schemas A, B, and C in succession for ten trials each and recorded performance on each schema throughout all 30 trials. Results are presented in Fig. 13. For low values of P , there are insufficient neurons to learn any representations. For high values of P , each schema is learned with high performance, but quickly forgotten when a new schema is introduced. Furthermore, as more schemas are learned, the recall performance of all previous schemas drops. This shows that the capacity of schemas able to be stored may also be dependent on sparsity level. Intermediate sparsity levels strike a balance, learning schemas to a satisfactory performance without catastrophic forgetting.

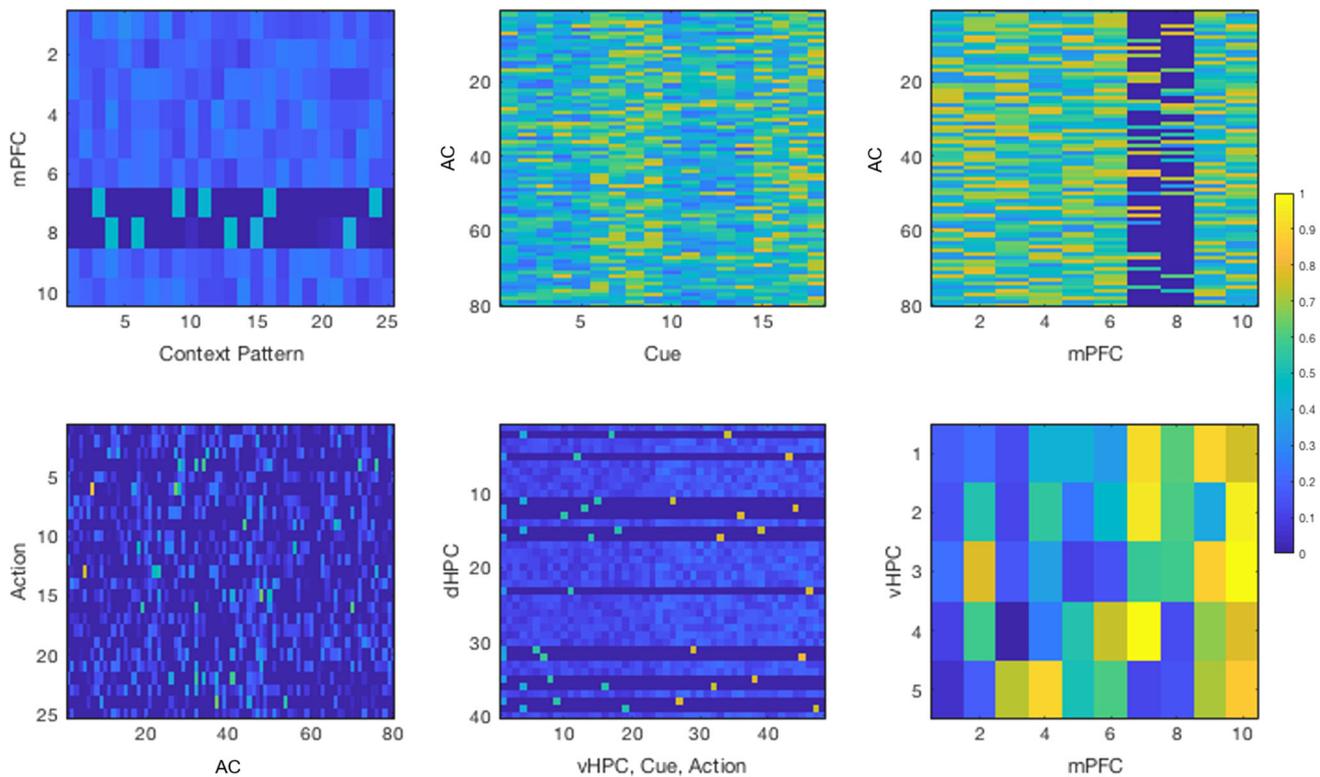


Fig. 8 Weight matrices of the control network after simulating all of the second experiment. Rows represent postsynaptic layers and columns represent presynaptic layers. Weights from the context pattern to the mPFC now show two distinct schemas, with stronger patterns than seen

in the first experiment. Weights from mPFC to the AC show two sets of gating patterns. Weights to the dHPC now show twice the amount of triplets as before, reflecting that triplets from two schemas are being encoded

4.6 Effects of mPFC size on network capacity of schemas

The mPFC recognizes schemas, providing contextual information to the vHPC in the indexing stream and AC in the representation. We thus examine the effects of varying the number of neurons in the mPFC when training three schemas in succession as before. Results from Fig. 14 show that if the number of mPFC neurons exceeds the number of schemas being encoded, the model performs as in previous experiments. However, if there are an insufficient number of mPFC neurons, we get the surprising result that catastrophic forgetting decreases. When there is only one neuron in the mPFC, there is no catastrophic forgetting at all. This is due to the fact that the three schemas have no overlap with each other, such that storing all of the information in a single schema would not destroy any information. We then repeated this experiment with three schemas having significant overlaps. In this case, all three schemas shared the same set of flavors, and the second and third schemas each had one overlapping food well location with the first schema. Figure 15 shows that this causes catastrophic forgetting of the first schema due to the overlaps. Taken together, this suggests that increasing

the number of mPFC neurons decreases the amount of catastrophic forgetting that occurs.

5 Discussion

Accounting for the results of Tse et al. (2007), we showed that our biologically plausible neural network was able to learn schemas over time and quickly assimilate new information if it was consistent with a prior schema. Furthermore, the components of our network were consistent with the functionality of brain areas in the Tse et al. experiments, with HPC dependence for schema learning and updating, and mPFC dependence for recognizing context and modulating learning accordingly. The learning was done through an indexing stream in which the mPFC and HPC projected context-dependent patterns onto the representation stream, and the rapid encoding was done by enhancing replay activity of novel and familiar information. The network was also able to learn multiple schemas without catastrophic forgetting, by maintaining separate sets of AC neurons for tasks within different schemas.

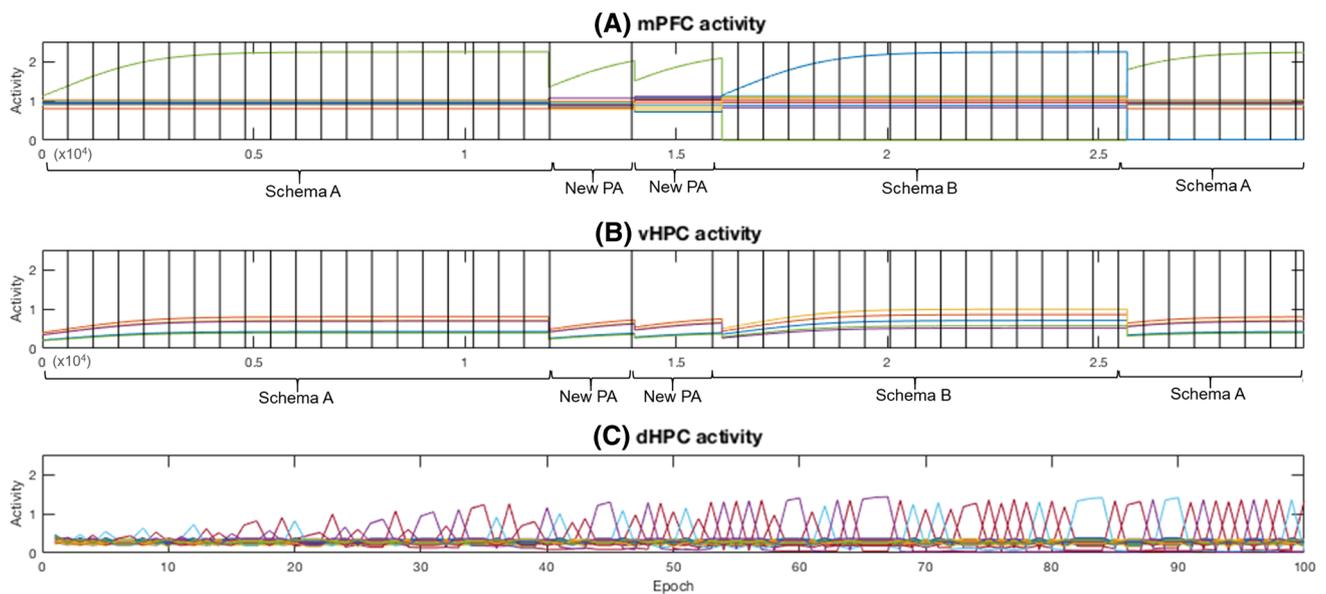


Fig. 9 Neural activity for the network of a single simulated rat while performing the first and second experiments combined. Each colored line represents the activity of one neuron. Each vertical black line represents the separation of epochs into trials. Activity is measured before winner-take-all is applied. **a** The mPFC activities for the first and second experiments are shown in sequence, with the training of Schema A and two instances of new PAs for the first experiment and the training of Schema B and return of Schema A for the second experiment. When training on Schema A, one mPFC neuron is consistently chosen as winner, as its weight values increase over time. When new PAs are introduced, the winner remains the same, but decreases in activity. At the start of Experiment 2, Schema B is introduced and a new mPFC

neuron wins. The number of epochs greatly increases due to the novelty. When Schema A is retrained, the neuron previously associated with schema A becomes active again. **b** The neurons in vHPC follow the same general trend as the mPFC neurons. However, all neurons increase and decrease together, as they all take input from the winning mPFC neuron. **c** Activity of the dHPC neurons for the first 100 epochs of the first trial is shown. As the density of switching of dHPC winners occurs at every epoch rather than at every trial, it is necessary to display the activity at the epoch level. At each epoch, a different dHPC neuron is selected to have its weights increase, as a different triplet is present each time. The activities of winning neurons gradually increase for 40 epochs, remaining stable afterward

5.1 Hippocampal indexing

The network highlighted diverse roles of indexing by the HPC. Eichenbaum (2017) proposed that specific memories are represented in the dHPC whereas contextual information is represented in the vHPC. Combined with indexing theory, our model showed that indexing separates representations of objects, spatial layouts and tasks by the contexts in which they belong. This modularity makes it less likely that learning new information in new contexts would overwrite previously learned information in old contexts. The indexing of the dorsal HPC is necessary for driving the processes that transfer information to long-term storage.

The indexing behavior is comparable to a recent approach to avoiding catastrophic forgetting by Nakano and Hattori (2017), in which the intermediate layers of a deep neural network are gated by patterns that differ by context. A paper by Masse et al. (2018) has the similar idea of using CHL as a plausible deep representation of information, and applying “pseudopatterns” alongside their regular training patterns for better separation. Our experiments explain in more depth how these patterns are formed and employed throughout dif-

ferent stages of the learning process. The central location of the hippocampus makes it a likely candidate for effecting context-dependent gating within the network. Its connections to the mPFC allow the formation of context-dependent indices, and its wide connectivity to the whole neocortex gives it the ability to gate information at many levels of representation.

5.2 Neuromodulation and novelty detection

The gating of patterns by the HPC is not employed evenly at all times, but depends heavily on external factors such as novelty, uncertainty, and reward. Therefore, our model predicts that gating of context information is controlled by neuromodulatory areas. This could explain how the brain controls phases of learning in such a flexible manner. Our model suggests that schema familiarity could be detected by the mPFC and novelty could be detected by the HPC. We propose that neuromodulation monitors these signals and amplifies learning and encoding of new (i.e., novel or unfamiliar) information while sparing old information. Our simulated neuromodulator may have biological correlates in

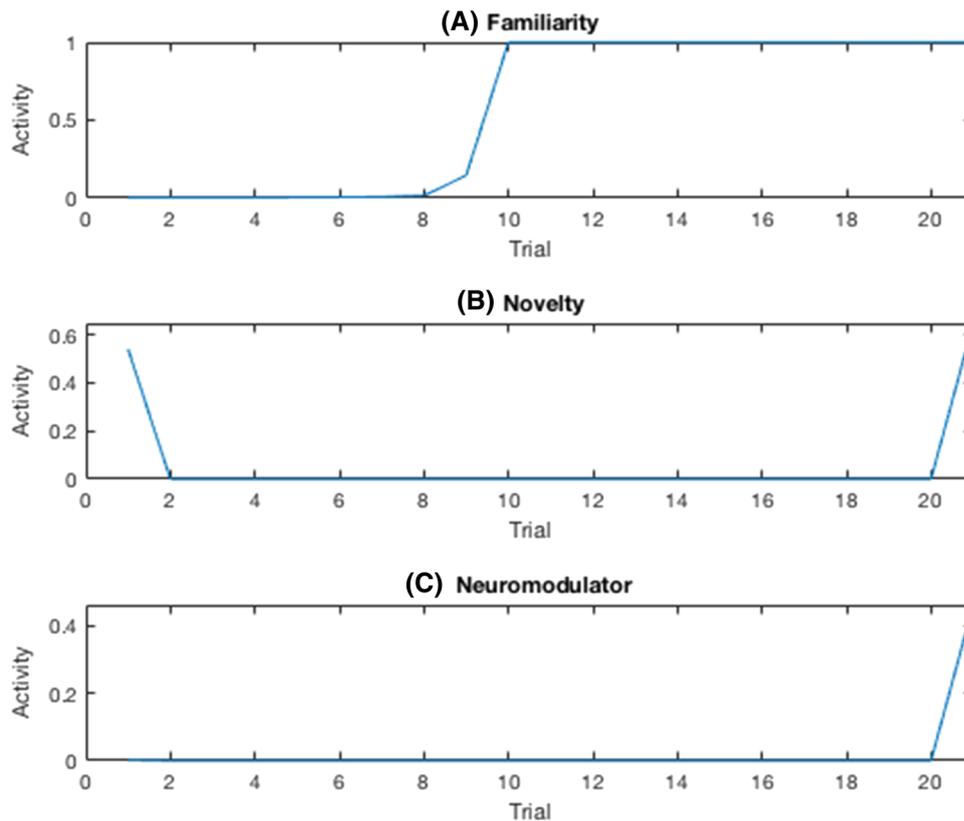


Fig. 10 Combining familiarity and novelty for neuromodulation. Familiarity, novelty, and neuromodulator are updated after each epoch of training, with their final levels at the end of each trial shown in this graph. **a** Familiarity increases as the network is trained on a schema. Due to the sigmoidal transfer function, familiarity activity is step-like, going from unfamiliar to familiar over some training. **b** Novelty starts

at a high value whenever there are new PAs, which occurs on the first trial when the schema is introduced, and on the 21st trial when two PAs are replaced. **c** At each trial, the product of familiarity and novelty leads to an increase in activity when schema familiarity is high and novelty is high

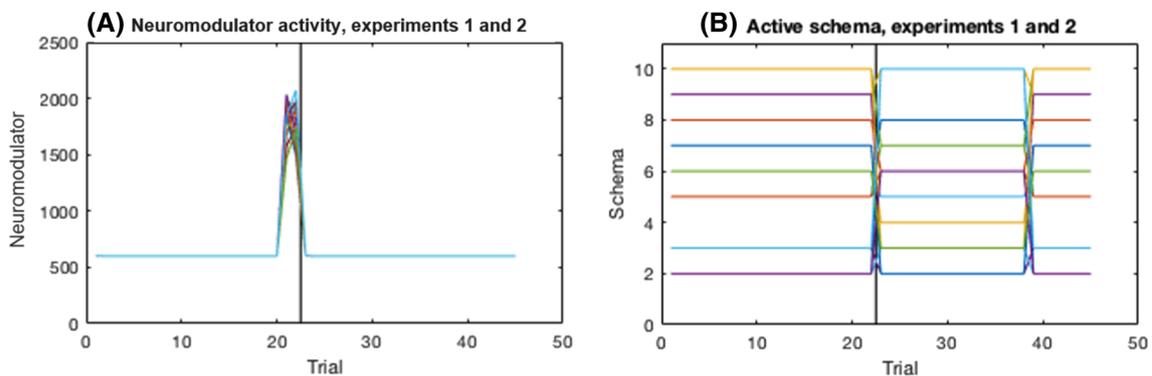


Fig. 11 Neuromodulator activity and winning schemas. Each colored line represents the activity of the neuromodulator from an individual simulated rat out of 20. **a** Neuromodulator activity expressed by the number of epochs trained per trial. The vertical black line separates experiments 1 and 2. On trial 21, the number of epochs rises sharply due to the combination of familiar schema and novel stimulus. **b** Index of winning mPFC neuron in each trial of experiments 1 and 2. The y-axis represents the indices of the neurons in mPFC. By tracing a single colored line, the winning mPFC neuron can be tracked throughout the

entirety of the two experiments. Since only one neuron is active in the winner-take-all selection, the indices also correspond to the encoding of the current context pattern as a single schema. The winning neuron was determined as the neuron with the highest activity in the mPFC. The index switches clearly each time the context pattern switches. In experiment 2, when returning to the original schema to check for retention at trial 38, the winning schema returns to the original index from experiment 1

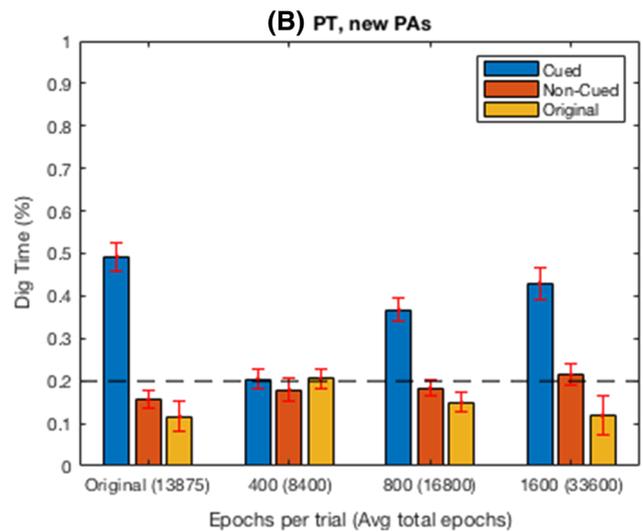
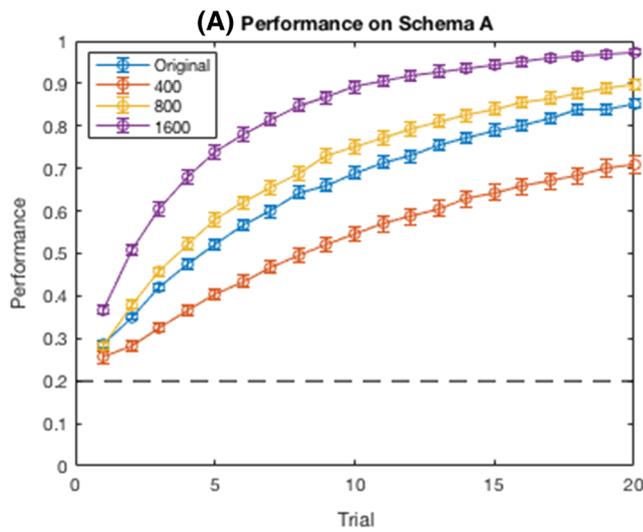


Fig. 12 **a** Performances of each condition. The blue line represents the original network and the remaining lines use a flat number of epochs as indicated in the legend. All conditions are able to learn the schema, but with different learning rates. **b** Probe tests of each condition after introducing a new PA. The average number of epochs for each trial is displayed in parentheses. The performance of the probe test increases

as more epochs are trained per trial. However, the original network can get a performance equivalent to the other conditions, but with only 13,875 epochs of training as opposed to 16,800 for the condition with 800 epochs per trial and 33,600 for the condition with 1600 epochs per trial

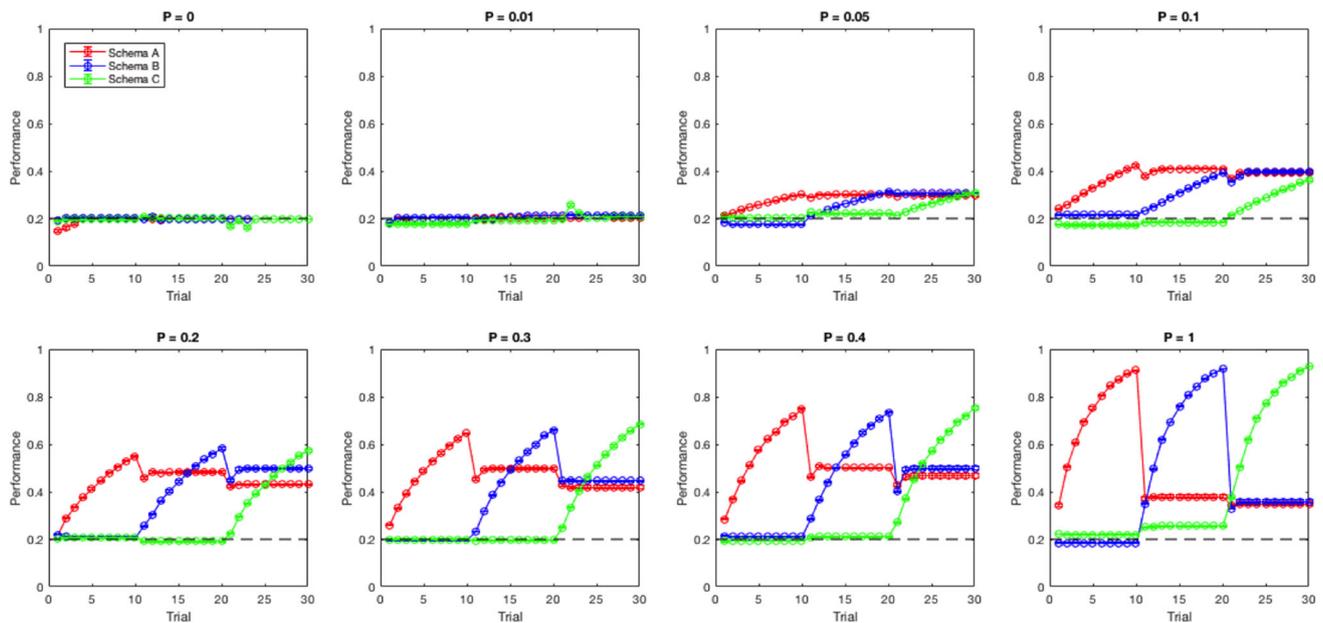


Fig. 13 Recall performance when training three schemas are learned in succession for ten trials each, with different vHPC sparsity values of P . Each colored line represents the performance of one schema. Performance for each schema starts at 0 and remains low until the schema is

actually trained on the network. When $P=0$, no learning occurs. When $P=1$, each schema is learned but forgotten when a new schema is introduced. Intermediate levels of sparsity make a trade-off between these two extremes

the locus coeruleus, as the LC reacts to sudden changes in schemas and causes changes in theta oscillations within the HPC. However, the combination of novelty and familiarity is also reminiscent of the basal forebrain functionality

suggested by Yu and Dayan (2005), in that the level of uncertainty is framed within a specific context. Other areas such as the dopaminergic ventral tegmental area (VTA) may be involved in neuromodulatory gating as well, as it has inbound

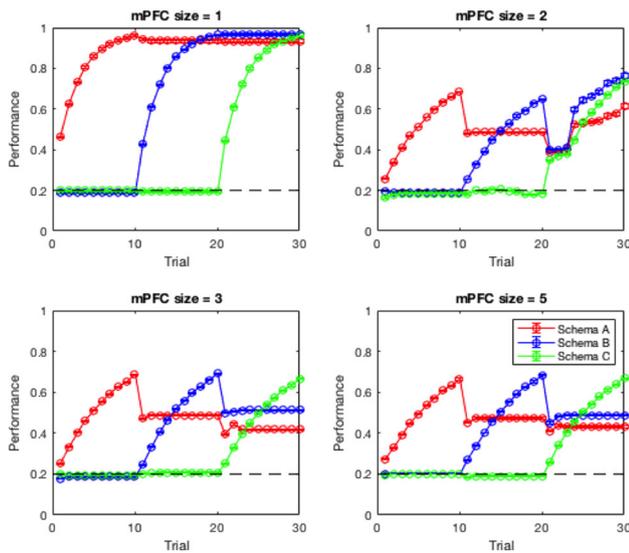


Fig. 14 Performance when training three distinct schemas in succession for 10 trials each, with different sizes of mPFC. When mPFC size = 1, no catastrophic forgetting occurs, as seen by how performance of each schema remains high even as other schemas are introduced. When there are sufficient mPFC neurons to represent the schemas, the network behaves as in previous experiments

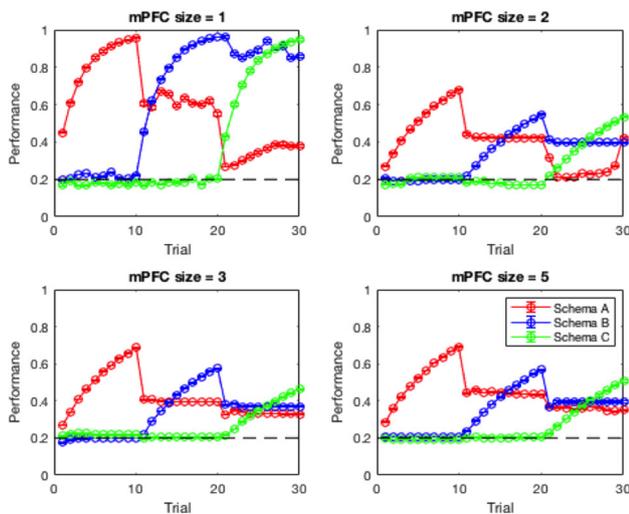


Fig. 15 Performance when training three schemas in succession for ten trials each, with different sizes of mPFC. The second and third schemas use the same flavors as the first schema, and each has one overlapping food well location with the first schema. Now with overlap, having an mPFC size of 1 is inadequate for learning the overlapping schemas

and outbound connections with the hippocampus that control the speed of learning according to reward and novelty (Otmakhova et al. 2013). It is important to note that the speed of learning in our model is defined not by the number of epochs used for training, but by the number of epochs that occur within a day of learning. This is meant to approximate how hippocampal replays may increase for salient information.

5.3 Interactions between the medial prefrontal cortex and hippocampus

There are differing opinions on how interaction between the mPFC and HPC is involved in context-dependent tasks. The SLIMM model suggests a competitive relationship, with activation of the mPFC inhibiting the HPC when a stimulus is congruent with a prior schema. On the other hand, Preston and Eichenbaum (2013) suggest a more cooperative interaction, with the mPFC drawing specific memories from the vHPC, and in turn influencing the dHPC via entorhinal inputs. As the mPFC is also known to mediate attention shifting in context-dependent tasks (Birrell and Brown 2000), it is likely that shifts in schemas cause the mPFC to change the activity of the HPC. In our experiment, the presence of new PAs within an existing schema should require both the mPFC and HPC to express the familiarity of the schema and novelty of the new PAs. In our model, the main roles of the mPFC are to provide contextual input to the representation stream and apply top down control of the HPC to change which specific neurons are active in the HPC to effectively separate tasks by schema. To further align with mPFC functions, future work should consider a distributed encoding in the mPFC, which could better represent overlapping information between schemas. It is also important to note that no direct anatomical connections exist from the mPFC and HPC and instead are routed through the thalamus. Future models including the thalamus could further test the theory proposed by Eichenbaum (2017) that the thalamus controls information flow between the mPFC and HPC.

5.4 Relevance to complex spatial navigation

Our model builds spatial maps of the environment that incorporate contextual information. As place cells are highly sensitive to context (Smith and Mizumori 2006), we demonstrate how associations of context and place change navigational behavior. Furthermore, experimental literature shows that the spatial selectivity of place cells decreases along the dorsal-ventral axis of the HPC (Jung et al. 1994). By extending the role of the dorsal-ventral axis as a hierarchical indexing area, we see how navigational decisions are affected by different schemas.

In addition, the use of the neuromodulatory area to increase the training of novel information matches observations that neuromodulation shapes and prioritizes replay events during consolidation (Atherton et al. 2015; Hasselmo 1999). This consolidation is important for decision making in complex spatial navigation tasks and is reflected in the model results. By combining the familiarity of context and novelty of index neurons in the hippocampus, the model quickly learns the locations associated with the cues for better navigation.

5.5 Applications for future studies

By uniting the theories of hippocampal indexing and interactions between the mPFC and HPC, we generate new and testable hypotheses that can be validated experimentally. By deactivating the LC or BF, we can test the effects of neuromodulation on the time it takes to learn the Tse et al. (2007) task, and discover which specific areas are applying the neuromodulation. We also expect that severing connections from the mPFC to the HPC would cause catastrophic forgetting of the tasks, as intermediate representations would not be properly gated. It may also be possible to lesion the HPC and project artificial gating patterns on the neocortical areas storing intermediate representations to see if this prevents catastrophic forgetting. In fMRI and behavioral experiments, activity in the HPC and vmPFC is observed to correlate with different stages of schema-consistent and schema-inconsistent encoding (van Kesteren et al. 2010). Following this idea, we may be able to study whether hippocampal activity along the dorsal–ventral axis supports the idea of hierarchical indexing based on hierarchical schemas. fMRI experiments may also yield insights on what brain areas may be implementing the context pattern, cue, AC, and action. The context pattern could be another region of the prefrontal cortex, upstream of the mPFC. The cue represents a main input and could correspond to the thalamus. The AC could be separated into more specific areas in the parietal lobe. The action could be an area controlling movement, such as the striatum or premotor cortex. Yet another possible benefit of behavioral and fMRI studies could be to learn how human brain prevents catastrophic forgetting. For instance, human subjects could be trained on multiple overlapping tasks to observe whether training each within a separate schema leads to better separation of tasks.

5.6 Applications to artificial intelligence and machine learning

In addition to the neurobiological implications of our model, our work could have practical applications to a range of tasks in artificial intelligence, machine learning, and robotics. We have recently applied our model architecture to a robotic task. Rather than learning the locations of food, a robot runs the model to learn the general layout of objects in a household. Different schemas, which are associated with different rooms, are formed by groups of objects typically found in a room. This is used to aid in human robot interaction by facilitating efficient context-dependent behavior. For example, if a person wants a banana, the robot knows that bananas are typically found in the kitchen, rather than in an office. Unlike our current model, which does not plan paths to objects, our robot demonstration shows how schemas lead to efficient trajectories when retrieving items.

In the future, we hope to test our model on a variety of other datasets, such as word corpora for the formation of semantic schemas in natural language processing. The increased complexity of the environment may require the addition of more multimodal association layers in the network, which would test the scalability of context-based gating. Rather than having just a vHPC and dHPC, the model would include multiple HPC layers along the dorsal–ventral axis, one for each of the layers in the representation stream. The use of more layers presents the possibility of starting from raw visual input as opposed to labeled objects, for an entirely end-to-end approach to context-based task learning.

Acknowledgements We thank the participants of the 2017 Telluride Neuromorphic Cognition Workshop, especially Xinyun Zou, Brent Komer, Georgios Detorakis, and Scott Koziol, who worked on a preliminary project leading to the creation of this model.

References

- Abraham WC, Robins A (2005) Memory retention—the synaptic stability versus plasticity dilemma. *Trends Neurosci* 28(2):73–78
- Aston-Jones G, Cohen JD (2005) Adaptive gain and the role of the locus coeruleus-norepinephrine system in optimal performance. *J Compar Neurol* 493(1):99–110
- Atherton LA, Dupret D, Mellor JR (2015) Memory trace replay: the shaping of memory consolidation by neuromodulation. *Trends Neurosci* 38(9):560–570
- Baxter MG, Chiba AA (1999) Cognitive functions of the basal forebrain. *Curr Opin Neurobiol* 9(2):178–183
- Berridge CW, Foote SL (1991) Effects of locus coeruleus activation on electroencephalographic activity in neocortex and hippocampus. *J Neurosci* 11(10):3135–3145
- Birrell JM, Brown VJ (2000) Medial frontal cortex mediates perceptual attentional set shifting in the rat. *J Neurosci* 20(11):4320–4324
- Detorakis G, Bartley T, Neftci E (2018) Contrastive hebbian learning with random feedback weights. *arXiv preprint arXiv:1806.07406*
- Eichenbaum H (2017) Prefrontal-hippocampal interactions in episodic memory. *Nat Rev Neurosci* 18(9):547
- French RM (1999) Catastrophic forgetting in connectionist networks. *Trends Cognit Sci* 3(4):128–135
- Hasselmo ME (1999) Neuromodulation: acetylcholine and memory consolidation. *Trends Cognit Sci* 3(9):351–359
- Hawkins J, Ahmad S, Cui Y (2017) A theory of how columns in the neocortex enable learning the structure of the world. *Front Neural Circuits* 11:81
- Jung MW, Wiener SI, McNaughton BL (1994) Comparison of spatial firing characteristics of units in dorsal and ventral hippocampus of the rat. *J Neurosci* 14(12):7347–7356
- van Kesteren MT, Fernández G, Norris DG, Hermans EJ (2010) Persistent schema-dependent hippocampal-neocortical connectivity during memory encoding and postencoding rest in humans. *Proc Natl Acad Sci* 107(16):7550–7555
- van Kesteren MT, Ruitter DJ, Fernández G, Henson RN (2012) How schema and novelty augment memory formation. *Trends Neurosci* 35(4):211–219
- Kirkpatrick J, Pascanu R, Rabinowitz N, Veness J, Desjardins G, Rusu AA, Milan K, Quan J, Ramalho T, Grabska-Barwinska A, et al (2017) Overcoming catastrophic forgetting in neural networks. In: *Proceedings of the national academy of sciences*, p 201611835

- Krichmar JL (2008) The neuromodulatory system: a framework for survival and adaptive behavior in a challenging world. *Adapt Behav* 16(6):385–399
- Kumaran D, Hassabis D, McClelland JL (2016) What learning systems do intelligent agents need? complementary learning systems theory updated. *Trends Cognit Sci* 20(7):512–534
- LeCun Y, Bengio Y, Hinton G (2015) Deep learning. *Nature* 521(7553):436
- Masse NY, Grant GD, Freedman DJ (2018) Alleviating catastrophic forgetting using context-dependent gating and synaptic stabilization. arXiv preprint [arXiv:1802.01569](https://arxiv.org/abs/1802.01569)
- Mattar MG, Daw ND (2018) Prioritized memory access explains planning and hippocampal replay. *Nat Neurosci* 21(11):1609
- McClelland JL, McNaughton BL, O'Reilly RC (1995) Why there are complementary learning systems in the hippocampus and neocortex: insights from the successes and failures of connectionist models of learning and memory. *Psychol Rev* 102(3):419
- Mermillod M, Bugajska A, Bonin P (2013) The stability-plasticity dilemma: investigating the continuum from catastrophic forgetting to age-limited learning effects. *Front Psychol* 4:504
- Movellan JR (1991) Contrastive Hebbian learning in the continuous Hopfield model. In: *Connectionist models*. Elsevier, pp 10–17
- Nakano S, Hattori M (2017) Reduction of catastrophic forgetting in multilayer neural networks trained by contrastive Hebbian learning with pseudorehearsal. In: *2017 IEEE 10th International Workshop on computational intelligence and applications (IWCIA)*. IEEE, pp 91–95
- Otmakhova N, Duzel E, Deutch AY, Lisman J (2013) The hippocampal-VTA loop: the role of novelty and motivation in controlling the entry of information into long-term memory. In: *Intrinsically motivated learning in natural and artificial systems*. Springer, pp 235–254
- Pfeiffer BE, Foster DJ (2013) Hippocampal place-cell sequences depict future paths to remembered goals. *Nature* 497(7447):74
- Preston AR, Eichenbaum H (2013) Interplay of hippocampus and prefrontal cortex in memory. *Curr Biol* 23(17):R764–R773
- Rumelhart DE, Hinton GE, Williams RJ (1986) Learning representations by back-propagating errors. *Nature* 323(6088):533
- Smith DM, Mizumori SJ (2006) Hippocampal place cells, context, and episodic memory. *Hippocampus* 16(9):716–729
- Soltoggio A, Stanley KO, Risi S (2017) Born to learn: the inspiration, progress, and future of evolved plastic artificial neural networks. arXiv preprint [arXiv:1703.10371](https://arxiv.org/abs/1703.10371)
- Teyler TJ, DiScenna P (1986) The hippocampal memory indexing theory. *Behav Neurosci* 100(2):147
- Tse D, Langston RF, Kakeyama M, Bethus I, Spooner PA, Wood ER, Witter MP, Morris RG (2007) Schemas and memory consolidation. *Science* 316(5821):76–82
- Tse D, Takeuchi T, Kakeyama M, Kajii Y, Okuno H, Tohyama C, Bito H, Morris RG (2011) Schema-dependent gene activation and memory encoding in neocortex. *Science* 333(6044):891–895
- Wagatsuma A, Okuyama T, Sun C, Smith LM, Abe K, Tonegawa S (2018) Locus coeruleus input to hippocampal CA3 drives single-trial learning of a novel context. *Proc Natl Acad Sci* 115(2):E310–E316
- Walling SG, Brown RA, Milway JS, Earle AG, Harley CW (2011) Selective tuning of hippocampal oscillations by phasic locus coeruleus activation in awake male rats. *Hippocampus* 21(11):1250–1262
- Yu A, Dayan P (2005) Uncertainty, neuromodulation, and attention. *Neuron* 46(4):681–692

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