



Modeling Contextual Modulation of Memory Associations in the Hippocampus

Praveen K. Pilly*, Michael D. Howard and Rajan Bhattacharyya

Center for Human-Machine Collaboration, Information and Systems Sciences Laboratory, HRL Laboratories, Malibu, CA, United States

We present a computational model of how memories can be contextually acquired and recalled in the hippocampus. Our adaptive contextual memory model comprises the lateral entorhinal cortex (LEC), the dentate gyrus (DG) and areas CA3 and CA1 in the hippocampus, and assumes external inputs about context that originate in the prefrontal cortex (PFC). Specifically, we propose that there is a top-down bias on the excitability of cells in the DG of the hippocampus that recruits a sub-population of cells to differentiate contexts, independent of experienced stimuli, expanding the “pattern separation” role typically attributed to the DG. It has been demonstrated in rats that if PFC is inactivated, both acquisition and recall of memory associations are impaired. However, PFC inactivation during acquisition of one set of memory associations surprisingly leads to subsequent facilitation of the acquisition of a conflicting set of memory associations in the same context under normal PFC operation. We provide here the first computational and algorithmic account of how the absence or presence of the top-down contextual biases on the excitability of DG cells during different learning phases of these experiments explains these data. Our model simulates PFC inactivation as the loss of inhibitory control on DG, which leads to full or partial activation of DG cells related to conflicting memory associations previously acquired in different contexts. This causes context-inappropriate memory traces to become active in the CA3 recurrent network and thereby the output CA1 area within the hippocampus. We show that these incongruous memory patterns proactively interfere with and slow the acquisition of new memory associations. Further, we demonstrate that pattern completion within CA3 in response to a partial cue for the recall of previously acquired memories is also impaired by PFC inactivation for the same reason. Pre-training the model with interfering memories in contexts different from those used in the experiments, simulating a lifetime of experiences, was crucial to reproduce the rat behavioral data. Finally, we made several testable predictions based on the model that suggest future experiments to deepen our understanding of brain-wide memory processes.

Keywords: hippocampus, contextual coding, pattern completion, dentate gyrus, memory encoding, memory retrieval, memory interference, catastrophic forgetting

OPEN ACCESS

Edited by:

Michael A. Yassa,
University of California, Irvine,
United States

Reviewed by:

Avi Mendelsohn,
University of Haifa, Israel
James B. Aimone,
Sandia National Laboratories (SNL),
United States

*Correspondence:

Praveen K. Pilly
pkpilly@hrl.com

Received: 22 May 2018

Accepted: 12 October 2018

Published: 09 November 2018

Citation:

Pilly PK, Howard MD and
Bhattacharyya R (2018) Modeling
Contextual Modulation of Memory
Associations in the Hippocampus.
Front. Hum. Neurosci. 12:442.
doi: 10.3389/fnhum.2018.00442

1. INTRODUCTION

A defining characteristic of our daily lives is our ability to recall memories of experiences that occurred in arbitrary places, even from the distant past. So, a fundamental question is what enables the brain to represent these seemingly countless memories? It has been known for a long time that the hippocampus is critical for the formation of long-term declarative memories. Its main function is to rapidly bind together multi-modal cortical signals representing a current event into a memory engram (Poo et al., 2016), such that a partial cue presented later can trigger the reactivation of the corresponding engram and thereby the episodic memory recall (Marr, 1971; Squire and Alvarez, 1995). Distributed regions beyond the medial temporal lobe, such as the prefrontal cortex (PFC), are also known to be involved in memory processing, but their roles are computationally less clear. In this regard, the main goal of this article is to advance a computational account of how external contextual signals can modulate various aspects of associative memory encoding and recall in the entorhinal-hippocampal system. In support of our computational model, we present simulation results that replicate experimental data from rats related to regulating interference in the encoding and recall of context-based associative memories. That data showed that memory associations can be not only impaired but also facilitated by the inactivation of PFC under various conditions (Navawongse and Eichenbaum, 2013; Peters et al., 2013). In particular, if PFC is inactivated, both acquisition and recall of memory associations are impaired. However, PFC inactivation during acquisition of one set of memory associations surprisingly leads to subsequent facilitation of the acquisition of a conflicting set of memory associations in the same context under normal PFC operation. Furthermore, the typical advantage for learning a conflicting set of memory associations in a different context is lost under PFC inactivation.

Context is a high-level functional concept, more than simply a particular grouping of spatial locations. A set of items that need to be memorized together can signify a context, and so can a common set of task rules governing behavior in a given situation (Preston and Eichenbaum, 2013). In the Navawongse and Eichenbaum (2013) and Peters et al. (2013) experiments, contextual cues are supplied by placing the rats in rooms with different wallpapers or ambient odors that serve to functionally distinguish the varied memory-based responses. When a familiar context is recognized, PFC is thought to bias the retrieval of context-appropriate memories in the hippocampus to guide current behavior (Preston and Eichenbaum, 2013). PFC is generally agreed to be the brain region responsible for top-down executive control, able to filter out arbitrary task-irrelevant neural representations in distributed cortical areas with its widespread connectivity (Miller and Cohen, 2001). Several studies support the involvement of PFC in memory processing, with coordinated interactions between PFC and hippocampus (Siapas and Wilson, 1998; Jones and Wilson, 2005; Peyrache et al., 2009; Kim et al., 2011; Brincat and Miller, 2015).

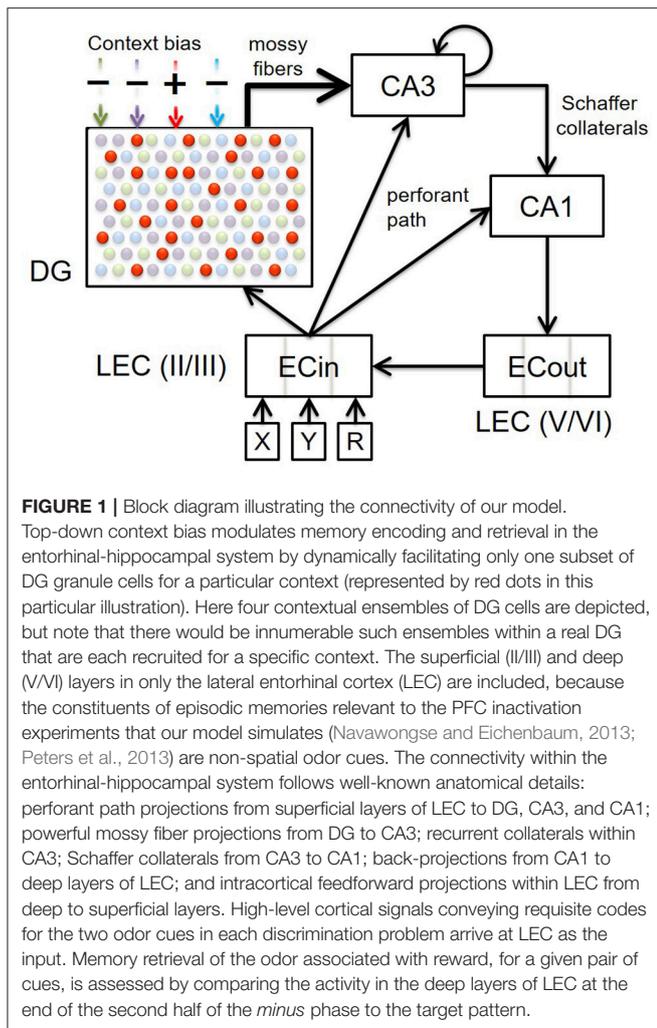
It is well known that different subsets of hippocampal neurons are assigned to encode memories experienced in different contexts (Kubie and Muller, 1991; Leutgeb et al., 2004, 2005).

Specifically, DG neurons encoding various memories in one context are not only sparse and distributed but also statistically independent from those encoding memories in a different context (Markus et al., 1995; Doboli et al., 2000; Chawla et al., 2005; Liu et al., 2012; Ramirez et al., 2013), and involve a local inhibitory circuit for context-based retrieval (Raza et al., 2017). Our model makes a direct link between these two sets of data, positing coordinated interactions between PFC and DG for contextualizing memories. In particular, we demonstrate that the experimental data from the PFC inactivation experiments can be simulated by adding an external biasing signal over DG granule cells to exert inhibitory control of their excitability, whereby an independent subset of cells are recruited for the acquisition and recall of memory associations experienced in a given context. Based on the Navawongse and Eichenbaum (2013) and Peters et al. (2013) experiments, we propose that this external biasing signal is present only when the PFC is online. We believe our computational work adds to the emerging understanding of brain-wide memory processes at the neural, circuit, and network levels.

2. MODEL

Our memory model (see **Figure 1**) provides a computational account of how episodic memory associations can be encoded and retrieved in a context-sensitive manner. It employs rate-coded point neurons in multiple layers and subfields within the entorhinal-hippocampal system; namely, superficial (ECin) and deep (ECout) layers of lateral entorhinal cortex (LEC), dentate gyrus (DG), and areas CA3 and CA1 within the hippocampus. CA3 is modeled as a dense recurrent neural network where active cells comprising the current episodic experience learn to rapidly “auto-associate” (i.e., learn projections to themselves), which allows for pattern completion of the corresponding “engram” pattern when only a subset of them are activated subsequently. The lateral entorhinal cortical signals traverse the hippocampal circuitry from LEC to CA1 along two streams; namely, the direct (monosynaptic) pathway: ECin → CA1, and the indirect (trisynaptic) pathway: ECin → (DG and CA3), DG → CA3, CA3 → CA1 (Amaral and Witter, 1995). Our model builds on a previous hippocampal model proposed by Ketz et al. (2013) in the *emergent* neural network simulator (O’Reilly and Munakata, 2000; Aisa et al., 2008). The Ketz model differentially modulates these hippocampal pathways in different phases of the hippocampal theta rhythm as suggested by some prior computational models (Hasselmo et al., 2002; Kunec et al., 2005); see Douchamps et al. (2013) for pertinent data.

We have made several changes to the Ketz et al. (2013) model. Our key innovation is to instantiate different subsets of DG cells in different situational contexts. This is critical as it relates to our main hypothesis that the contextual recruitment of DG cells is governed by a top-down inhibitory control mechanism. We also made a number of technical changes to conform to known anatomy and physiology. First, synaptic plasticity in a majority of the connectivity (perforant path: ECin → {DG and CA3}; mossy fibers: DG → CA3; Schaffer collaterals: CA3 → CA1)



is fully Hebbian (i.e., $k_{hebb} = 1.0$ in Equation 6), and not a combination of Hebbian and error-driven learning that is heavily biased toward the latter, as in Ketz et al. (2013). Second, there are no EC-like “slots” (or groupings of cells) in CA1 coding individual input streams, as CA1 has conjunctive/episodic cells like CA3 (Leutgeb et al., 2004) that can develop and sustain even without CA3 inputs (Brun et al., 2002) in response to divergent perforant path projections from superficial layers of entorhinal cortex (Naber et al., 2001). Lastly, we removed ECout \rightarrow CA1 backprojections, as there is little supporting evidence. Details of the model are described below in the Model Equations section.

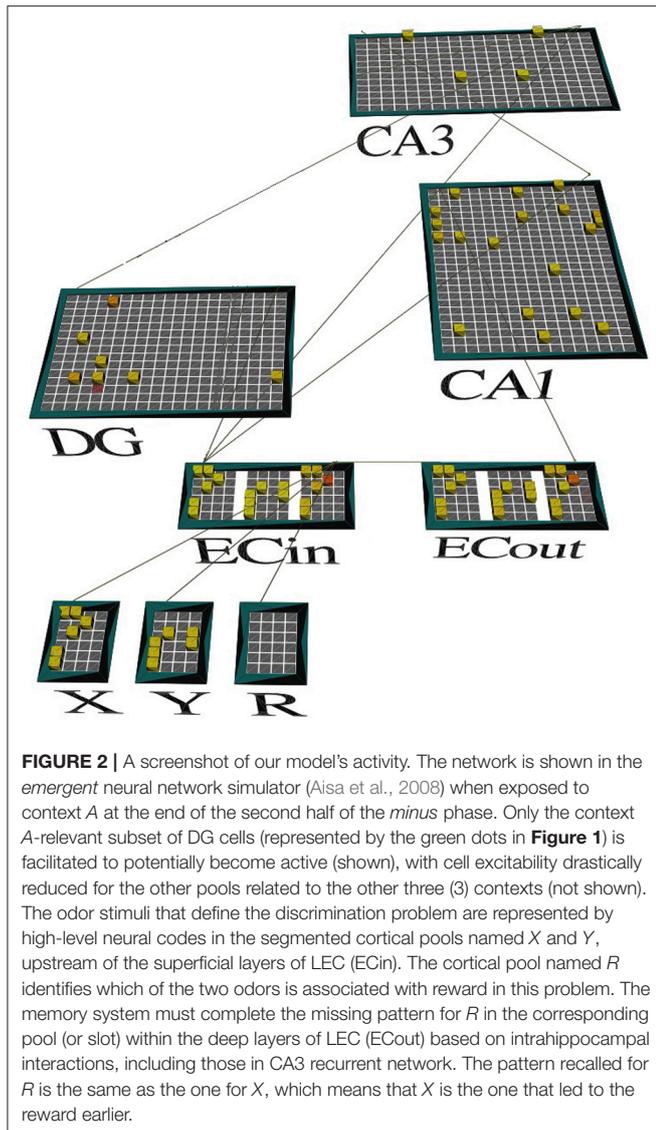
3. EXPERIMENTAL DETAILS

We now briefly describe how a contextual memory association is formed and recalled in our model. Suppose a rat enters context *A* for the first time and encounters two dishes X_1 and Y_1 with distinct odors. The experimenter has placed a reward only in one dish, say X_1 , and the rat is allowed to dig for the reward only from one dish in each trial. If by chance the rat chooses the baited dish and thereby obtains the reward, then an episodic memory is

formed for context *A* that is composed of the association between the two odor cues of X_1 and Y_1 and the presence of reward in X_1 (i.e., $\{X_1, Y_1\} \rightarrow X_1$). As the memory of the rewarded event gets strengthened from several trials in context *A*, the rat gradually makes fewer errors in making the memory-guided choice whenever it encounters odor cues X_1 and Y_1 in context *A*. The rat can learn several other discrimination memories within the same context; e.g., $\{X_2, Y_2\} \rightarrow X_2$, $\{X_3, Y_3\} \rightarrow X_3$, and so on. In Experiment 1A of Peters et al. (2013) the rats were presented with a sequence of eight discrimination problems in random order on each day until they reached a criterion level of performance in terms of the number of correct choices across the eight pairs. The rats can also learn and remember a new set of odor discrimination problems in another context *B*. The context *B*-specific memories can exhibit different levels of overlap in the components of episodes from context *A*. In the experiment of Navawongse and Eichenbaum (2013), the rats were presented with the same pair of odor cues in context *B* (e.g., $\{X_1, Y_1\}$) but the baited dish was the opposite of that in context *A*; i.e., $\{X_1, Y_1\} \rightarrow Y_1$. And in Experiments 2 and 3 of Peters et al. (2013), the rats were required to learn a new set of odor discrimination problems either within the same context or a different context. Here the memories overlapped only in one component. In particular, one odor in each pair was retained and the reward prediction of this odor was reversed compared to the first set. For example, for the episode $\{X_1, Y_1\} \rightarrow X_1$, if X_1 were retained then the new rewarded event would be $\{X_1, Z_1\} \rightarrow Z_1$; whereas if Y_1 were retained then the new rewarded event would be $\{Z_1, Y_1\} \rightarrow Y_1$.

The three components of each odor discrimination memory (e.g., $\{X_1, Y_1\} \rightarrow X_1$) are represented by sparse distributed codes in distinct cortical populations upstream of the entorhinal-hippocampal system; see groups of neurons labeled *X*, *Y*, and *R* in Figure 2. In our simulations, each unique odor is assigned a randomly constructed binary bit pattern on a 6x4 grid with exactly six cells turned on (maximal activation of 1). Our model incorporates four independent contextual subsets of cells in DG (namely, DGa, DGb, DGc, and DGd), whose recruitment is controlled by direct modulation of cell excitability (see Figure 1). In the biological DG, there can be some minimal overlap among contextual ensembles, and new ensembles can constantly be created due to neurogenesis (Luu et al., 2012); so our model is only a very simple test of the concept. In particular, the excitability of context-inappropriate DG cells is temporarily suppressed by raising the maximal conductance of their lateral inhibitory channels (namely, \bar{g}_i in Equation 2); see Table 1. When PFC is inactivated, there is a loss of this contextual inhibitory control over the DG. In other words, the excitability of all DG cells remains at normal levels without any top-down relative bias for a particular contextual ensemble.

Context *A* is used to simulate the data for Experiments 1A and 1B in Peters et al. (2013), while contexts *A* and *B* are both employed for the other experiments. We devised a simple scheme to emulate the innumerable contexts the rats may have experienced in their lifetimes, for which they have formed memories that could potentially interfere with encoding stimuli in these contexts. For this purpose, two odor discrimination lists



with high interference (overlap of one odor in each problem presented in the contexts at hand) were designed. Simulated rats acquired these lists for five blocks each in the other two contexts C and D before the experimental manipulations. See section 3.1 for how each simulated rat was instantiated.

For all experiments, the retrieval performance P (%) in response to each discrimination problem was assessed by using a similarity metric based on the root mean squared error (RMSE), E , between the correct pattern R_{target} and the pattern R_{out} recalled in the third field of the deep layers of LEC (ECout), which represents the cortical read-out of the hippocampal recall process, as follows:

$$P(\%) = \min\left(50\%, 100 \times \frac{(1 - E)^{10}}{0.5^{10} + (1 - E)^{10}}\right), \quad (1)$$

where E is in the range 0 – 1. Note Equation 1 ensures that chance performance is 50%, as the rats have to choose between two odors

TABLE 1 | Model parameters.

Parameter	Equation	LEC (II/III, V/VI)	DG	CA3,CA1
\bar{g}_j	(2)	0.1	0.1	0.1
\bar{g}_e	(2)	1	1	1
\bar{g}_i	(2)	1	1 or 5	1
(Υ, θ)	(4)	(100, 0.5)	(100, 0.5)	(100, 0.5)
(E_j, E_e, E_i)	(2)	(0.3, 1, 0.25)	(0.3, 1, 0.25)	(0.3, 1, 0.25)
V_{rest}	(2)	0.3	0.3	0.3
q	(5)	0.25	0.25	0.25
k in kWTA	(5)	25%	1%	2.5%

Values for various parameters in each sub-region within our entorhinal-hippocampal model that were used to simulate the experimental data related to the effects of PFC inactivation on memory encoding and recall behavior. Parameter values for the connections among various sub-regions are provided in **Table 2**. Note that all sub-regions share the same parameter values, except for the level of sparseness in activity (namely, k in Equation 5). Consistent with data (Jung and McNaughton, 1993; Chawla et al., 2005), DG is the sparsest compared to CA3 and CA1 in the hippocampus (archicortex), which are still less sparse relative to the superficial and deep layers of LEC (neocortex). And note \bar{g}_i is set to 1 for the DG ensemble assigned to a current context and set to 5 for other DG ensembles assigned to other contexts.

for each discrimination problem. We now briefly describe how each of the experiments was simulated.

Navawongse and Eichenbaum (2013): Simulated rats ($N = 10$) learned a list of odor discrimination memories ($\{X_i, Y_i\} \rightarrow X_i/Y_i, i = 1...8$) in context A across several blocks until criterion performance was achieved (i.e., 90% accuracy in two consecutive blocks) and then learned a list of conflicting associations ($\{X_i, Y_i\} \rightarrow Y_i/X_i, i = 1...8$) in context B to criterion. This training phase was followed by blocks of test trials ($\{X_i, Y_i\} \rightarrow ?$, $i = 1...8$) in each context to assess memory retrieval performance with saline or muscimol injections into PFC. As mentioned above, we simulated the muscimol condition by allowing all DG ensembles to potentially become active (with a lower \bar{g}_i), whereas we simulated the saline (control) condition by selectively increasing \bar{g}_i for all DG ensembles except the one that is pertinent to the current context. We also simulated prior experiences that could potentially interfere with the new memories of the experiment by pre-training the rats on interfering lists of discrimination problems in contexts C and D for 10 blocks each ($C:\{X_i, Z_i\} \rightarrow X_i/Z_i, i = 1...8$); ($D:\{Z_i, Y_i\} \rightarrow Z_i/Y_i, i = 1...8$).

Peters et al. (2013) – Experiment 1A: Simulated rats ($N = 10$ for normal PFC [control]; $N = 10$ for inactivated PFC [muscimol]) learned a list of odor discrimination memories ($\{X_i, Y_i\} \rightarrow X_i/Y_i, i = 1...8$) in context A across several blocks until criterion performance was achieved (i.e., 90% accuracy in two consecutive blocks). Each discrimination problem was presented once per block in random order. Muscimol rats had their PFC inactivated only during the first three training blocks. For each rat, once criterion was reached, four test blocks were conducted to assess memory retrieval performance. As in the experiment, PFC inactivation was applied to the first three test blocks for the simulated control rats only. We simulated prior experiences that could potentially interfere with

the new memories of the experiment by pre-training the rats on interfering lists of discrimination problems in contexts C and D for 10 blocks each ($C: \{X_i, Z_i\} \rightarrow X_i/Z_i, i = 1 \dots 8$); ($D: \{Z_i, Y_i\} \rightarrow Z_i/Y_i, i = 1 \dots 8$).

Peters et al. (2013) – Experiment 1B: Simulated rats ($N = 10$ for normal PFC [control]; $N = 10$ for inactivated PFC [muscimol]) were trained to learn one odor discrimination problem at a time to criterion ($\{X_i, Y_i\} \rightarrow X_i/Y_i, i = 1 \dots 10$), unlike the concurrent acquisition in Experiment 1A. The relative order for acquiring the memories was chosen randomly for each rat. For muscimol rats, all training trials occurred under PFC inactivation. We also simulated prior experiences that could potentially interfere with the new memories of the experiment by pre-training the rats on interfering lists of discrimination problems in contexts C and D for 10 blocks each ($C: \{X_i, Z_i\} \rightarrow X_i/Z_i, i = 1 \dots 8$); ($D: \{Z_i, Y_i\} \rightarrow Z_i/Y_i, i = 1 \dots 8$).

Peters et al. (2013) – Experiment 2: Simulated rats ($N = 40$) were first trained in context A to learn to criterion the same list of discrimination problems that was used for Experiment 1A. They were then trained on a new list of memories for five blocks. As mentioned earlier, these two lists conflicted as follows: for each context A memory (say, $\{X_1, Y_1\} \rightarrow Y_1$), one of the two odor cues was randomly replaced by a new cue (say, W_1 instead of Y_1) and the reward prediction of the remaining cue from context A (i.e., X_1) was reversed leading to $\{X_1, W_1\} \rightarrow X_1$. Rats were exposed to List 2 either in the same context A ($N = 10$) or a different context B ($N = 10$), with either a normal ($N = 10$) or an inactivated ($N = 10$) PFC. In other words, this experiment employed a 2x2 design with the following subsets of rats: control – different context, control – same context, muscimol – different context, and muscimol – same context. As in Experiment 1A, muscimol rats had their PFC inactivated only for the first three blocks to learn List 2. We also simulated prior experiences that could potentially interfere with the new memories of the experiment by pre-training the rats on interfering lists of discrimination problems in contexts C and D for 10 blocks each ($C: \{X_i, Z_i\} \rightarrow X_i/Z_i, i = 1 \dots 8$); ($D: \{Z_i, Y_i\} \rightarrow Z_i/Y_i, i = 1 \dots 8$).

Peters et al. (2013) – Experiment 3: Simulated rats ($N = 10$ for normal PFC [control]; $N = 10$ for inactivated PFC [muscimol]) were first trained in context A for five blocks with the same list of discrimination problems used in Experiment 1A, and then trained on the second, conflicting list from Experiment 2 for five blocks in the same context A . PFC inactivation for the muscimol group occurred only during the first three training blocks for List 1. We also simulated prior experiences that could potentially interfere with the new memories of the experiment by pre-training the rats on interfering lists of discrimination problems in contexts C and D for 10 blocks each ($C: \{X_i, Z_i\} \rightarrow X_i/Z_i, i = 1 \dots 8$); ($D: \{Z_i, Y_i\} \rightarrow Z_i/Y_i, i = 1 \dots 8$).

3.1. Model Equations

We now provide structural and functional details of our model, including values for the various parameters. The model simulates associative memory formation and recall using dynamic sparse ensemble codes and activity-dependent synaptic plasticity in the recurrent connections within the hippocampus. Simulations were performed in the *emergent* neural network simulator

(O'Reilly and Munakata, 2000; Aisa et al., 2008), whose underlying equations are described in detail below. Following Ketz et al. (2013), we simulated a cycle of three different phases of hippocampal processing during each trial of memory formation; namely, first and second halves of the *minus* phase, followed by the *plus* phase. In the first half of the *minus* phase, the trisynaptic pathway is suppressed, inhibiting the CA3-based recall of previous memories. But the monosynaptic pathway is still active, which enables the activity pattern in CA1 (decoupled from CA3) to be learned as an auto-encoded representation of the EC inputs. In the second half of the *minus* phase, the monosynaptic pathway is suppressed in favor of the trisynaptic pathway. Memory recall thus coincides with the second half of the *minus* phase, when the system expectation is registered in ECout as a result of pattern completion in CA3. In this way, CA1 activity is controlled by different input streams in the different *minus* phases. The *emergent* framework also posits a *plus* phase that is similar to the first half of the *minus* phase but with ECout clamped to the incoming cortical patterns representing the memory association (in the streams for the input cues and the rewarded cue). The patterns in the cortical areas upstream to ECin are clamped during all phases in each trial.

During memory acquisition, the synaptic weights in the network change in an activity-dependent manner at the end of the *plus* phase using a combination of Hebbian and error-driven learning. For Hebbian learning, the weight changes depend on just the *plus* phase activities. For error-driving learning, the weight changes for connections depend on activities in the *plus* phase and the corresponding *minus* phase. In particular, synaptic weight changes in the monosynaptic pathway (ECin \rightarrow CA1, CA1 \rightarrow ECout, ECout \rightarrow ECin) depend on activities in the *plus* phase and the first half of the *minus* phase. Of particular importance, during memory acquisition, the CA3 \rightarrow CA1 connections are adapted such that the pattern-completing CA3 activity can reproduce the auto-encoded representation in CA1 during subsequent cue-based recall. There are now more biologically plausible implementations of error-driven learning for training multi-layered recurrent neural networks (Liao et al., 2016; Lillicrap et al., 2016; Hassabis et al., 2017; Scellier and Bengio, 2017; Whittington and Bogacz, 2017) without requiring matching feedforward and feedback connectivity and non-local information for updating weights.

The model equations are described next, and values for various model parameters are listed in **Tables 1, 2**. The membrane potential V_m of a given model cell is defined by Equation 2 that obeys shunting dynamics within a recurrent competitive network:

$$\frac{dV_m(t)}{dt} = \tau [\bar{g}_l(E_l - V_m(t)) + g_e(t)\bar{g}_e(E_e - V_m(t)) + g_i(t)\bar{g}_i(E_i - V_m(t))], \quad (2)$$

where τ scales the rate of the cell's temporal integration; \bar{g}_l is the constant conductance of the leak Cl^- channel; E_l is the reversal potential of the leak channel; \bar{g}_e is the maximal conductance of each excitatory channel and E_e is the corresponding reversal

TABLE 2 | Model connectivity.

Projections between brain regions	Connectivity	k_{hebb} (Equation 6)	Weight scale (absolute, relative)
LEC II/III → DG (perforant path)	$\rho = 25\%$	1	(1, 1)
LEC II/III → CA3 (perforant path)	$\rho = 25\%$	1	(1, 1)
LEC II/III → CA1 (perforant path)	$\rho = 25\%$	0.05	(1, 1)
DG → CA3 (mossy fibers)	$\rho = 4\%$	1	(10, 3.5)
CA3 → CA1 (Schaffer collaterals)	full	1	(5, 1)
CA1 → LEC VI/VI (recurrent collaterals)	full	0.05	(1, 2)
LEC VI/VI → LEC II/III (intracortical)	one-to-one	0.05	(1, 1)

Details of the connections among various sub-regions within our entorhinal-hippocampal model, including the type of connectivity [namely, one-to-one, full (each-to-all), random with probability], the proportion of Hebbian learning compared to error-driven learning ($0 \leq k_{hebb} \leq 1$ in Equation 6), and the values for absolute scaling parameter and relative weight (see a_j and r_j in Equation 3). Parameter ϵ , which scales the rate of learning, is set to 0.01 for all connections.

potential; \bar{g}_i is the maximal conductance of each inhibitory channel and E_i is the corresponding reversal potential; g_e is the net synaptic weight of excitatory channels controlled by cells either within the network or from other networks; and g_i is the net synaptic weight of inhibitory channels controlled by cells within the network. Equation 2 was numerically integrated using Euler’s forward method with a fixed time step $\Delta t = 1$ ms. The net excitatory synaptic weight g_e is defined as follows:

$$g_e(t) = \sum_j \sum_h \left(\frac{r_j}{\sum_k r_k} \right) a_j w_h^j, \tag{3}$$

where w_h^j is the synaptic weight of the excitatory projection from the h^{th} cell in the j^{th} network; r_j is the relative weight for the j^{th} network that is normalized by net relative weight for all incoming networks; and a_j is the absolute scaling parameter for the projections from the j^{th} network. The output activity $y(t)$ of the cell is computed using an activation function that operates on the membrane potential as follows:

$$y(t) = \frac{\chi}{1 + \chi}, \text{ with } \chi = \Upsilon [V_m(t) - \theta]_+, \tag{4}$$

where Υ is a scaling parameter; θ is the activity threshold on the membrane potential; and $[]_+$ denotes the half-wave rectifier function. The net inhibitory synaptic weight g_i is a constant for all cells within the network such that only a given proportion of the cells (say, the top k) exhibit non-zero activities at any moment, and is defined as follows:

$$g_\theta^k = \frac{g_e \bar{g}_e (E_e - \theta) + \bar{g}_i (E_i - \theta)}{\theta - E_i}, \text{ and } g_i(t) = g_\theta^{k+1} + q (g_\theta^k - g_\theta^{k+1}), \tag{5}$$

where $0 < q < 1$, g_θ^k is the threshold inhibition for the cell with the k^{th} most activity in the network. These equations to determine g_i , thus, implement k -Winner Take All (kWTA) inhibition leading to distributed representations with variable sparsities (O’Reilly and Munakata, 2000). They provide a simple computational approximation to the function of feedback inhibitory interneurons in real neural networks.

All membrane potentials are initialized to V_{rest} at the start of each trial. But model CA1 cells are reset to V_{rest} at the start of each

of the three theta phases in every trial. Additionally, cell activities in the deep layers of LEC (namely, ECount) are clamped during the *plus* phase (in the encoding trials) to the incoming pattern in the high-level cortical area upstream to LEC. In every encoding trial the three phases proceed in sequence as follows, each with a fixed number of time steps to allow for the activities to settle: first half of *minus* phase (30 steps) → second half of *minus* phase (30 steps) → *plus* phase (30 steps).

The adaptive weights, w_h^j , of the excitatory synaptic connection from the h^{th} cell in the j^{th} network are updated at the end of *plus* phase in each encoding trial using a combination of Hebbian learning and error-driven learning (O’Reilly and Munakata, 2000):

$$\Delta w_h^j = \epsilon [k_{hebb}(\Delta w_{hebb}) + (1 - k_{hebb})(\Delta w_{err})], \tag{6}$$

where $0 \leq k_{hebb} \leq 1$, ϵ scales the rate of learning; and k_{hebb} is a parameter that determines the proportion of Hebbian learning compared to error-driven learning in the mixture. The Hebbian (Δw_{hebb}) and error-driven (Δw_{err}) weight updates are defined as follows:

$$\Delta w_{hebb} = y^+ (x_{h,j}^+ - w_h^j), \tag{7}$$

$$\Delta w_{err} = (x_{h,j}^+ y^+) - (x_{h,j}^- y^-), \tag{8}$$

where $x_{h,j}$ is the activity of the projecting cell, and the superscripts + and - correspond to activities at the end of the *plus* phase and the applicable *minus* phase, respectively. Equation 7 is a variant of Hebbian learning that prevents the weights from growing without bounds (Grossberg, 1976; Rumelhart and Zipser, 1986; O’Reilly and Munakata, 2000). It ensures the selective strengthening of projections from those input cells that are consistently co-active with the post-synaptic cell during the *plus* phase. Equation 8, which is equivalent to the contrastive Hebbian learning rule (Hinton, 1990), ensures gradual matching between the activities at the end of *minus* (expectation) and *plus* (output) phases. The error-driven weight updates and the weights themselves are subject to soft bounding between the limits of 0 and 1. The weights of all present projections for each simulated rat are initialized at the beginning of experience

by random sampling from a uniform distribution with mean 0.5 and variance 0.25; see **Table 2** for probabilities related to sparse connectivity in the perforant path and mossy fiber projections.

4. RESULTS AND EXPLANATION

This section presents the simulation results of our model that explain the data from the odor discrimination experiments (Navawongse and Eichenbaum, 2013; Peters et al., 2013), which reveal the top-down modulation of context-based memory associations in the hippocampus and highlight the distributed nature of memory processes. With independent pools of DG cells that are facilitated one at a time for a given context under normal operation, our model can successfully retrieve context-appropriate memories and correctly discriminate in response to various pairs of odor cues. Even if the same stimulus is experienced in different contexts (say, odor X_1), varied and potentially conflicting associations can nevertheless be learned (e.g., $X_1 \leftrightarrow Y_1$ in context A; $X_1 \leftrightarrow Z_1$ in context B); see **Figure 6** for an illustrated explanation. This is possible despite the direct perforant path projections from the input layers of LEC (ECin) to CA3 and CA1. Our model is capable of forming and recalling contextual memories, because not only are the different contexts distinguished within the DG, but also the mossy fiber projections from DG to CA3 are the strongest relative to other connections to CA3 cells (namely, $EC \rightarrow CA3$ and $CA3 \rightarrow CA3$); see **Table 2**.

4.1. How Contextual Bias Affects Memory

It is well accepted that DG plays a crucial role in mediating the important sub-function of “pattern separation” by virtue of sparse connectivity in its perforant path projections from EC, its much larger size compared to the input EC layers, and its sparse distributed activity (Marr, 1971; Jung and McNaughton, 1993; Treves and Rolls, 1994; McClelland et al., 1995). There is a variety of evidence that the DG helps to distinguish similar memories experienced in different contexts. The degree of overlap among sets of DG cells that get activated for various memories within two different contexts is just 1%, which is much less than the 30% overlap in CA1 (Thompson and Best, 1989; Ramirez et al., 2013). Moreover, the selection of a new ensemble of DG cells can be triggered by exposure to a new environment or change in the behavioral task (Chawla et al., 2005; Leutgeb et al., 2007; Satvat et al., 2011; Schmidt et al., 2012; Deng et al., 2013).

With its random network comprising extensive recurrent collaterals (Amaral and Witter, 1989), CA3 underlies the crucial sub-functions of “auto-association” during encoding and “pattern completion” during cue-triggered retrieval (Marr, 1971; Treves and Rolls, 1994). Recurrent connections among CA3 cells that are simultaneously activated by a current event are selectively strengthened in an activity-dependent manner, which thereby subsequently support pattern completion in response to a partial cue (McNaughton and Morris, 1987; Nakazawa et al., 2002). As noted above, DG projects to CA3 with extremely sparse but potent mossy fibers (Henze et al.,

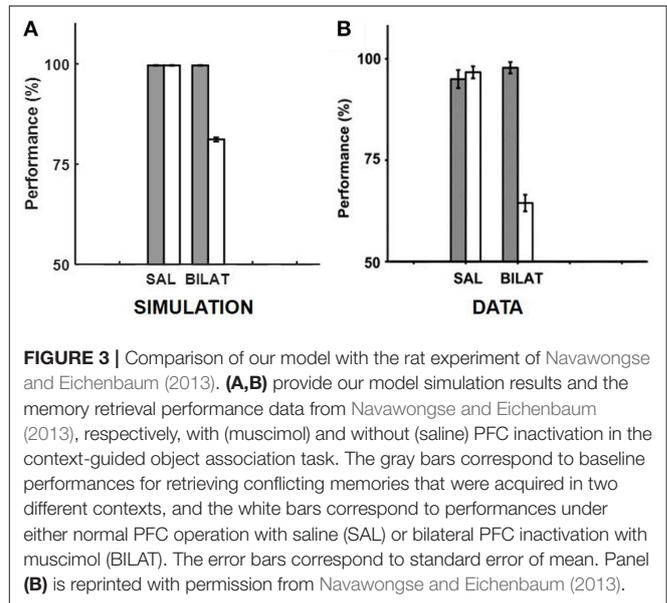
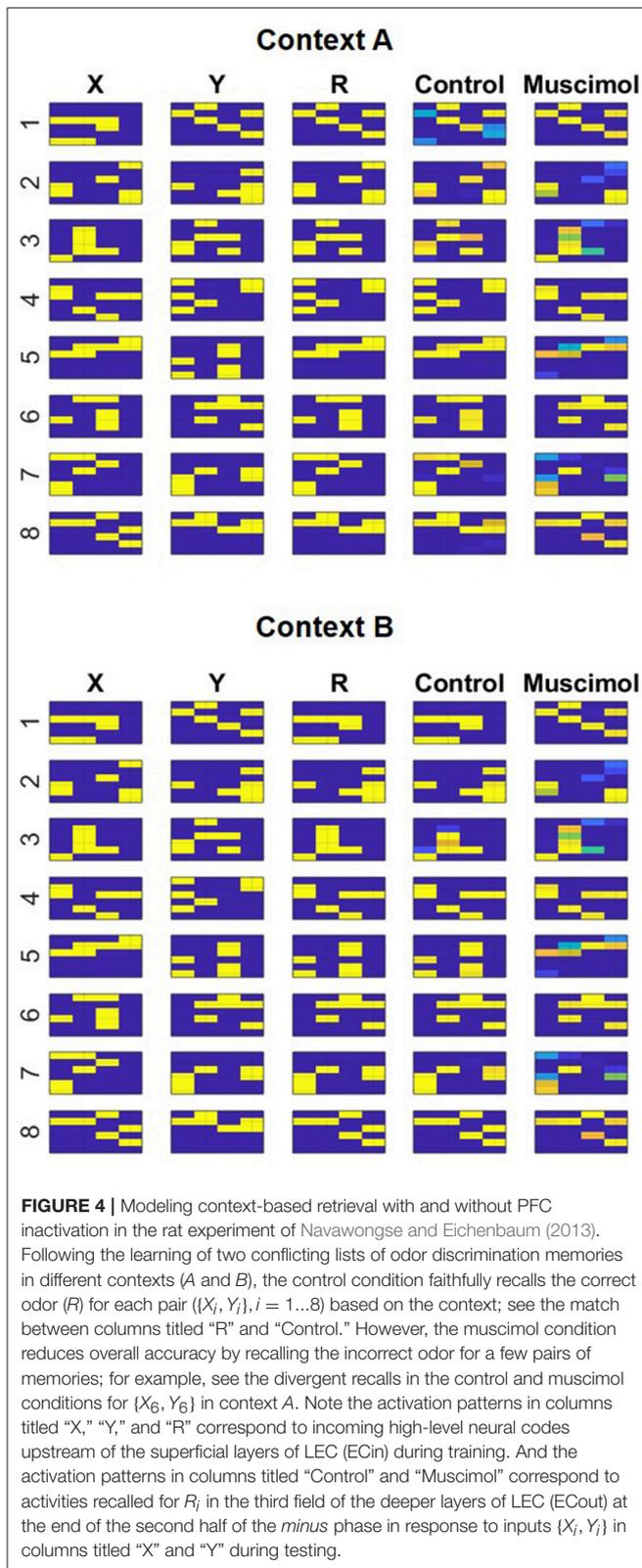


FIGURE 3 | Comparison of our model with the rat experiment of Navawongse and Eichenbaum (2013). **(A,B)** provide our model simulation results and the memory retrieval performance data from Navawongse and Eichenbaum (2013), respectively, with (muscimol) and without (saline) PFC inactivation in the context-guided object association task. The gray bars correspond to baseline performances for retrieving conflicting memories that were acquired in two different contexts, and the white bars correspond to performances under either normal PFC operation with saline (SAL) or bilateral PFC inactivation with muscimol (BILAT). The error bars correspond to standard error of mean. Panel **(B)** is reprinted with permission from Navawongse and Eichenbaum (2013).

1997) contributing to auto-association and pattern completion to minimize interference from similar memories (Leutgeb et al., 2007). We propose that all these properties work in conjunction with, and are further enhanced, by the additional contextual recruitment of DG cells; see also Deng et al. (2013). In other words, the mossy fibers from DG to CA3 bias the selection of the appropriate attractor state within the CA3 recurrent network in response to cues shared between multiple familiar contexts.

Given the specifics of the intra-hippocampal connectivity (see **Figure 1** and **Table 2**), context-specific subsets could also be triggered in CA3 and CA1 downstream from DG. However, CA3 and CA1 would exhibit an appreciable overlap among cells participating in different contexts (Thompson and Best, 1989; Ramirez et al., 2013) owing to their smaller sizes and more densely distributed activities compared to DG. We hypothesize that the contextual mossy fibers from DG to CA3 can bias the selection of the appropriate attractor state within the CA3 recurrent network in response to cues shared among multiple familiar contexts. As in the experiment of Navawongse and Eichenbaum (2013), our simulated rats learned to perform the contextual odor discrimination task almost perfectly. And they exhibited highly impaired performance when their PFC was inactivated [$t_{(9)} = 33.52, p = 0$; see **Figure 3**]. The ability to learn one or more memory associations in the same or different contexts for subsequent robust behavioral recall indicates the presence of stable attractor states in the CA3 recurrent network. **Figure 4** shows the activity patterns recalled by a representative simulated rat for the correct choice in response to each pair of odors in both contexts with and without PFC inactivation during the test blocks. The Control column shows the recalled patterns for normal PFC operation, which faithfully match the corresponding trained R patterns. The Muscimol column shows the recalled patterns when PFC is inactivated, and it appears

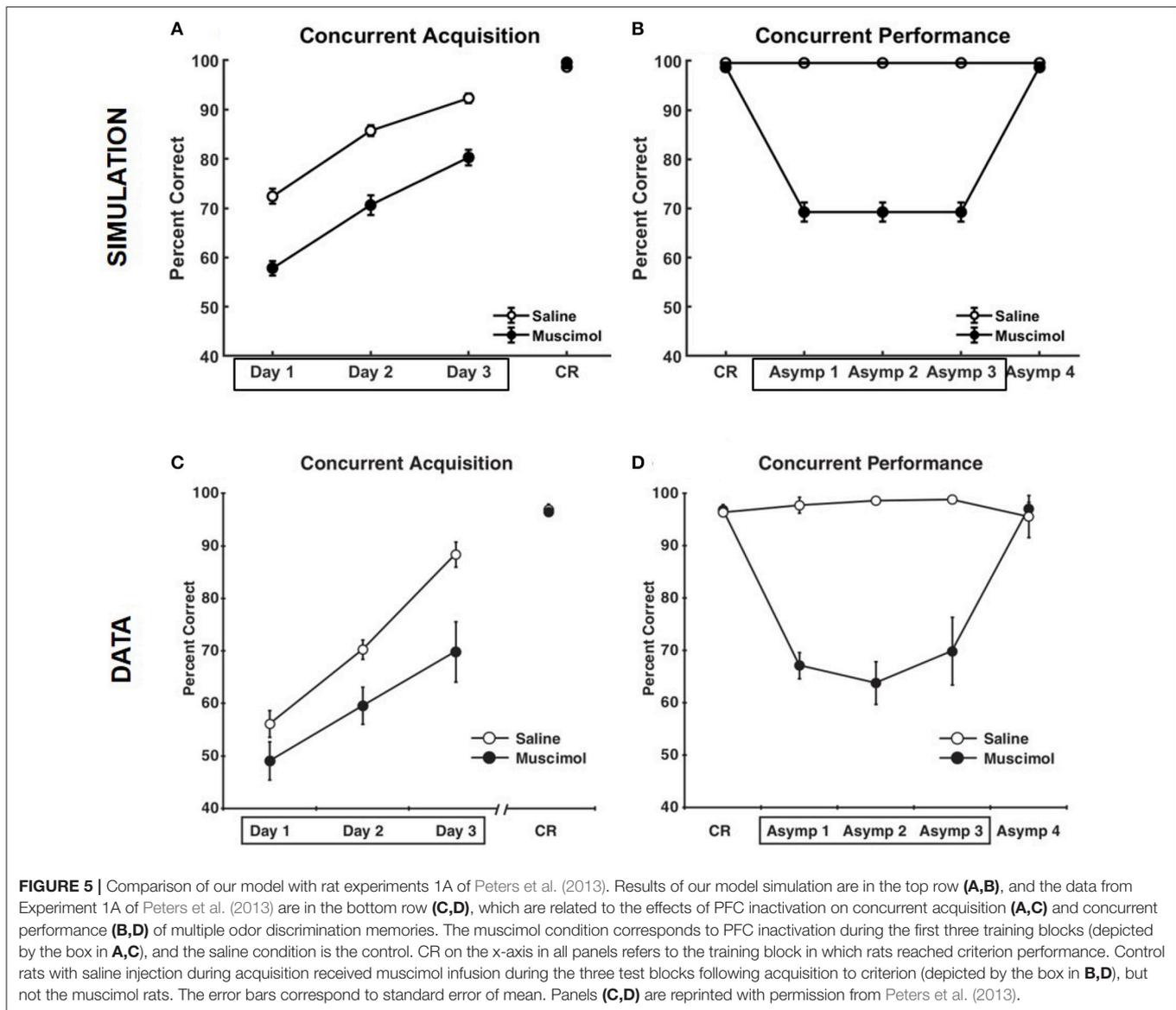


that the recalled patterns are a random choice between the corresponding patterns for contexts *A* and *B*, and not a mixture of the two contexts.

4.2. Concurrent and Blocked Acquisition of Memories

Figure 5 compares the results of our model simulation with the data from Experiment 1A of Peters et al. (2013), which shows that PFC inactivation can affect the ability to discriminate various pairs of odors even if they have only been presented in a single context. The model simulated 10 control rats and 10 muscimol rats, with each simulated rat being presented with a sequence of eight discrimination problems in random order on each epoch until a criterion level of performance was reached. The performance data of the simulated rats from the first three training blocks was analyzed using a two-way repeated measures ANOVA with inactivation condition and training block as factors. Both simulated control and muscimol rats showed evidence of learning, with a significant main effect of training block [$F_{(2, 54)} = 94.49, p = 0$], as in the data (see Figures 5A,C). And similar to the data, simulated muscimol rats learned the memory associations less accurately than simulated control rats initially, with a significant main effect of inactivation condition [$F_{(1, 54)} = 119.64, p = 0$]. Also matching the data, there was no interaction between PFC inactivation and training block in our model [$F_{(2, 54)} = 0.56, p = 0.5737$]. The performance data of the simulated rats from the first three test blocks, after concurrent acquisition to criterion performance, was also analyzed using a two-way repeated measures ANOVA with inactivation condition and training block as factors. As in the data, muscimol infusion for the simulated control rats severely impaired their ability to recall well-formed memories, with a significant main effect of inactivation condition [$F_{(1, 54)} = 654.09, p = 0$; see Figures 5B,D]. Finally, when PFC inactivation ceased in these simulated control rats for the last test block, retrieval behavior returned to about the same level as the simulated muscimol rats, as seen in the data.

Figure 6 illustrates our explanation for the impairment of memory encoding and retrieval that results from PFC inactivation. In the absence of inhibitory control by top-down contextual signals (see Figure 6, lower right panels), various sorts of context-inappropriate memory traces simultaneously emerge in the subfields of the hippocampus. In our model, the cause of these interfering signals is the indiscriminate activation of cells within the multiple contextual ensembles in DG to varying degrees in response to the current input. These DG traces promote the activation of their corresponding attractor states within the CA3 recurrent network via their previously tuned mossy fiber projections. The presence of multiple potential CA3 attractor states proactively conflicts with the auto-associative learning of the relevant activity pattern within CA3 during encoding. In particular, CA3 could randomly slip into one of the attractor states in each training trial. As a result, the trial-to-trial learning of CA3-to-CA1 Schaffer collaterals will be slower because of the lack of consistency in the CA3 activity pattern for the same discrimination problem across multiple acquisition trials while the PFC is inactivated. Moreover, the tuned Schaffer collaterals from CA3 cells to CA1 cells, and the tuned connections from CA1 cells to ECout cells, representing a previously encoded memory engram (to which CA3 converged)



will also offer some inertia to the formation of the new memory. Thus, the consequent activation of conflicting memory traces in CA3 not only slows the learning of new contextual memories (see **Figure 6**, right panel in second row), but also impairs the cue-based retrieval of previously learned context-specific memories. Similar to encoding under PFC inactivation, during recall as well, there would be competition among multiple potential attractor states within CA3. Further, it is possible that the CA3 recurrent network may not converge to any previously learned memory engram owing to the fragmentary nature of the activated traces within various DG pools. In any case, the previously tuned CA3 → CA1 and CA1 → ECoat projections will likely interfere with the generation of the correct recall pattern in the output layers of LEC (**Figure 6**, lower right).

Our model also simulates the data from Experiment 1B of Peters et al. (2013), in which rats learned each odor

discrimination memory one at a time to criterion before training on another in the set (see **Figure 7**). As in the data, simulated muscimol rats were slower overall in acquiring the memories [trials to criterion difference: $t_{(18)} = -3.85, p = 0.0012$], and less accurate through their blocked acquisition [percent correct difference: $t_{(18)} = 2.52, p = 0.022$]. When the odor discrimination problems were separated into the best and worst halves according to the number of trials to criterion, the model shows a significant interaction between the discrimination difficulty and inactivation condition factors [$F_{(1, 36)} = 21.03, p = 0.0051$; two-way repeated measures ANOVA], as in the data. Peters et al. (2013) interpreted this data as implying that PFC is not critical when acquiring memories one at a time in a single context, but that PFC is essential when many discrimination problems are to be learned and remembered at the same time. In contrast, our simulations suggest that interference from older

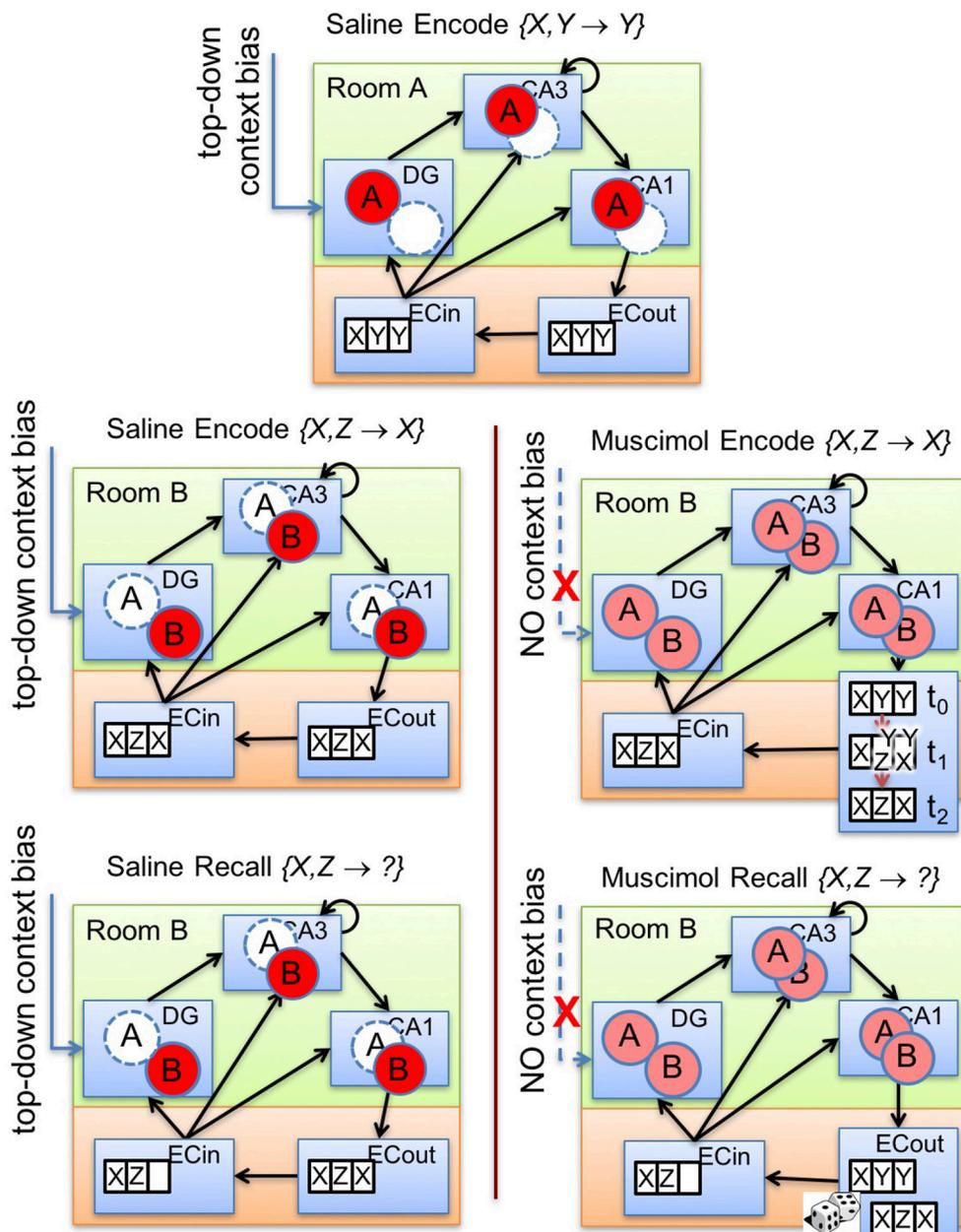
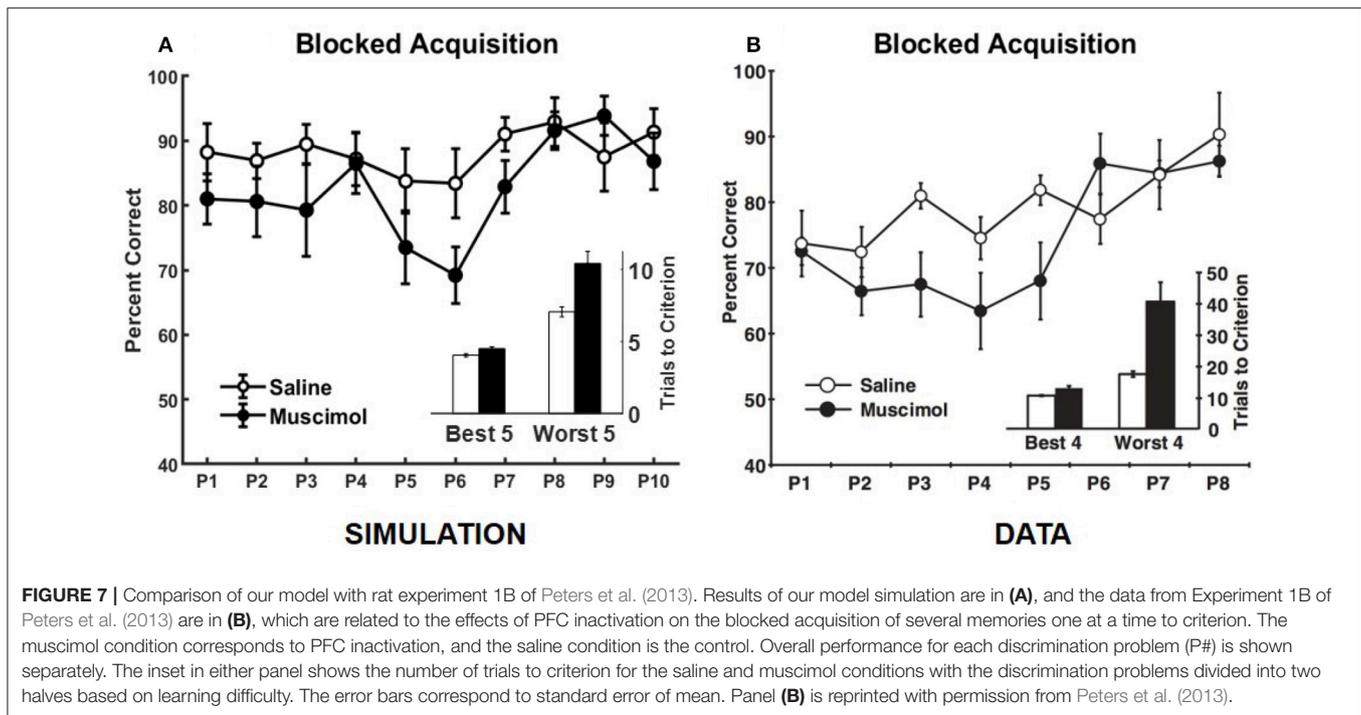


FIGURE 6 | Illustration of our theory. Top-down bias on DG affects memory encoding and recall by controlling the activation of memory traces within the hippocampal subfields (namely, DG, CA3, CA1). The top panel exemplifies prior experience where an associative memory $\{X, Y\} \rightarrow Y$ is encoded in room A with top-down bias from PFC. The other panels depict the encoding (second row) and retrieval (third row) for a conflicting associative memory $\{X, Z\} \rightarrow X$ in room B with (lower left panels) and without (lower right panels) the context bias. Note that normal PFC operation allows only context-appropriate memory traces to become active. With PFC inactivation, the learning of the conflicting association $\{X, Z\} \rightarrow X$ in room B is slowed because of the interfering memory traces triggered by concurrent activation of granule cells in the two DG contextual ensembles in response to the common cue X. If the two memories $\{X, Y\} \rightarrow Y$ and $\{X, Z\} \rightarrow X$ were acquired in contexts A and B, respectively, under normal conditions, PFC inactivation during recall performance $\{X, Z\} \rightarrow ?$ in context B would also lead to concurrent activation of memory traces related to both associations in response to the common cue X in all hippocampal subfields starting with DG. Pattern completion processes in CA3 would then probabilistically converge to either memory engram in different trials, leading to impaired retrieval behavior.

memories is potentially present even for blocked acquisition of single memories, and the strength of interference is not equal for all memories. In addition, the encoding process cannot be equal for all memories. It depends on the particular high-level cortical

patterns that need to be encoded and the precise connectivity structure within the entorhinal-hippocampal system at the time of encoding. In other words, each discrimination problem does not have the same exact probability of being encoded.



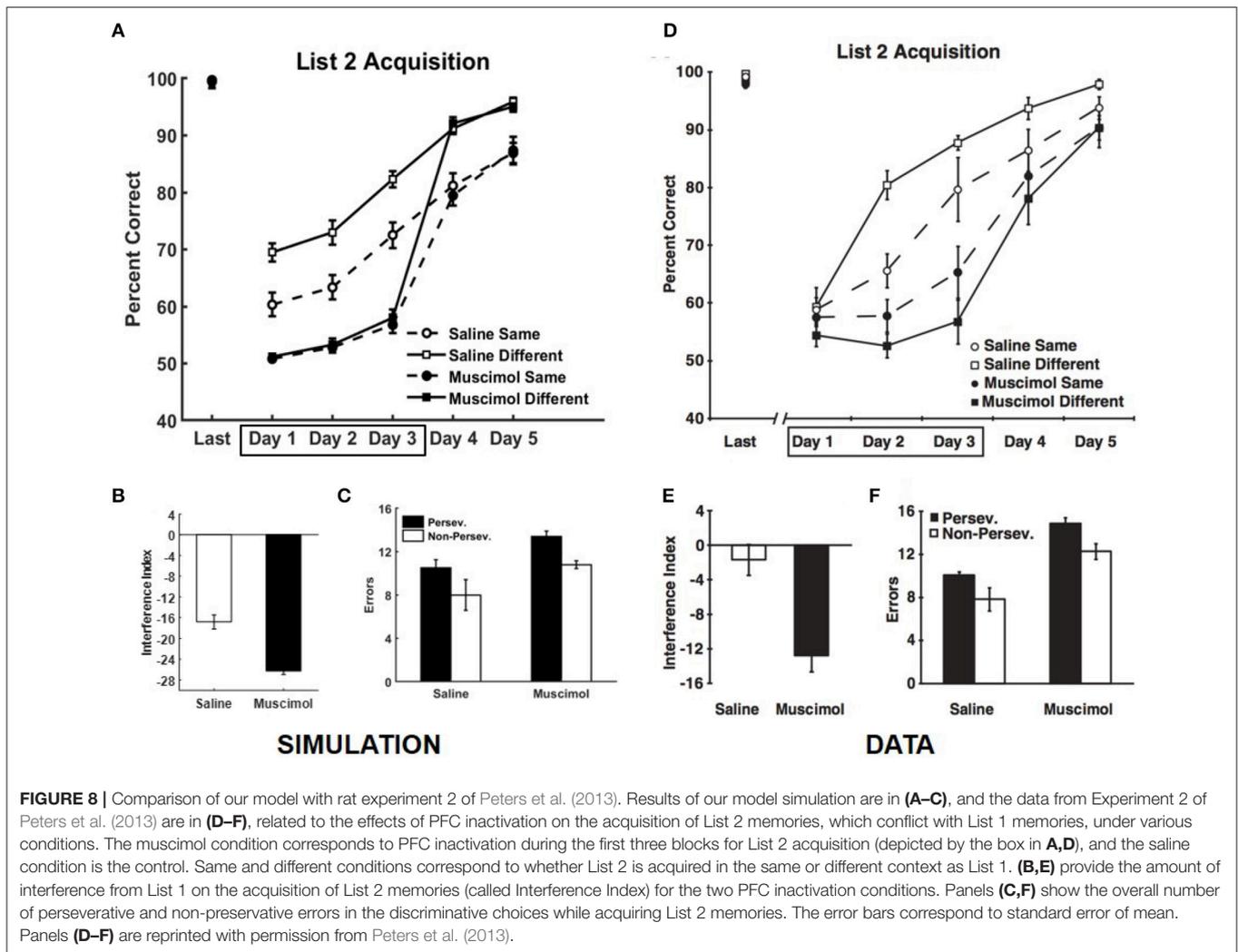
Thus, any untuned discrepancies between the experimental and simulated data, such as the differences in Day 1 performance between the saline and muscimol conditions for the concurrent acquisition of memories (see **Figure 5**) and in the overall number of trials to criterion for the blocked acquisition of memories (see **Figure 7**), are likely due to the particular neural codes chosen for the high-level cortical representation of the odor cues, the particular values chosen for the various parameters, and the particular initial entorhinal-hippocampal connectivity that was randomly sampled for each simulated rat, among others.

4.3. Acquisition of Conflicting Memories

Our simulations match the data from Experiment 2 of Peters et al. (2013), which is related to the effects of PFC inactivation on the dynamics of acquiring conflicting memories in the same and different contexts (see **Figure 8**). We simulate the data from control rats on the well-known effect of slower learning in a context where conflicting memories were previously formed (Barnes and Underwood, 1959). In this experiment, the rats were required to learn a new set of odor discrimination problems either within the same context or a different context, with and without PFC inactivation, where the new set conflicted with the set used in Experiment 1A. Here the memories of the two lists overlapped in one component, so that one odor in each X_1, Y_1 pair was retained and the reward prediction was reversed compared to the first set. For example, for the episode $\{X_1, Y_1\} \rightarrow X_1$, if X_1 were retained then the new event would be $\{X_1, Z_1\} \rightarrow Z_1$; whereas if Y_1 were retained then the new event would be $\{Z_1, Y_1\} \rightarrow Z_1$ (Z_1 is a new pattern, distinct from either X_1 or Y_1). The four different groups of simulated rats (PFC inactivation x context) first learned the list of memories

employed in Experiment 1A (namely, List 1) and showed no difference in retrieval performance following acquisition to criterion [$F_{(3, 36)} = 1.17, p = 0.33$]. The performance data of the simulated rats from the first three training blocks for List 2 was analyzed using a two-way ANOVA with inactivation and context conditions as the between subject factors and training block as the within subject factor. Similar to the data, this analysis revealed a significant main effect of inactivation condition [$F_{(1, 110)} = 303.76, p = 0$] and a significant interaction between inactivation and context conditions [$F_{(1, 110)} = 22.29, p = 0$]; see **Figures 8A,D**. Simulated control rats exhibited a significant main effect of context condition [$F_{(1, 54)} = 32.28, p = 0$], while simulated muscimol rats did not differentiate between the same and different context conditions [$F_{(1, 54)} = 0.62, p = 0.44$] with respect to encoding the new memories from List 2, which interfered with older memories from List 1. Further, the interference from List 1 when learning List 2 memories was higher for simulated muscimol rats than simulated control rats [$t_{(18)} = 5.99, p = 0$; **Figure 8B**], as in the data shown in **Figure 8E**. This was quantified using an interference index metric, which is defined as the difference between the average percent correct for List 2 vs. that for List 1 in their first five training blocks (Peters et al., 2013).

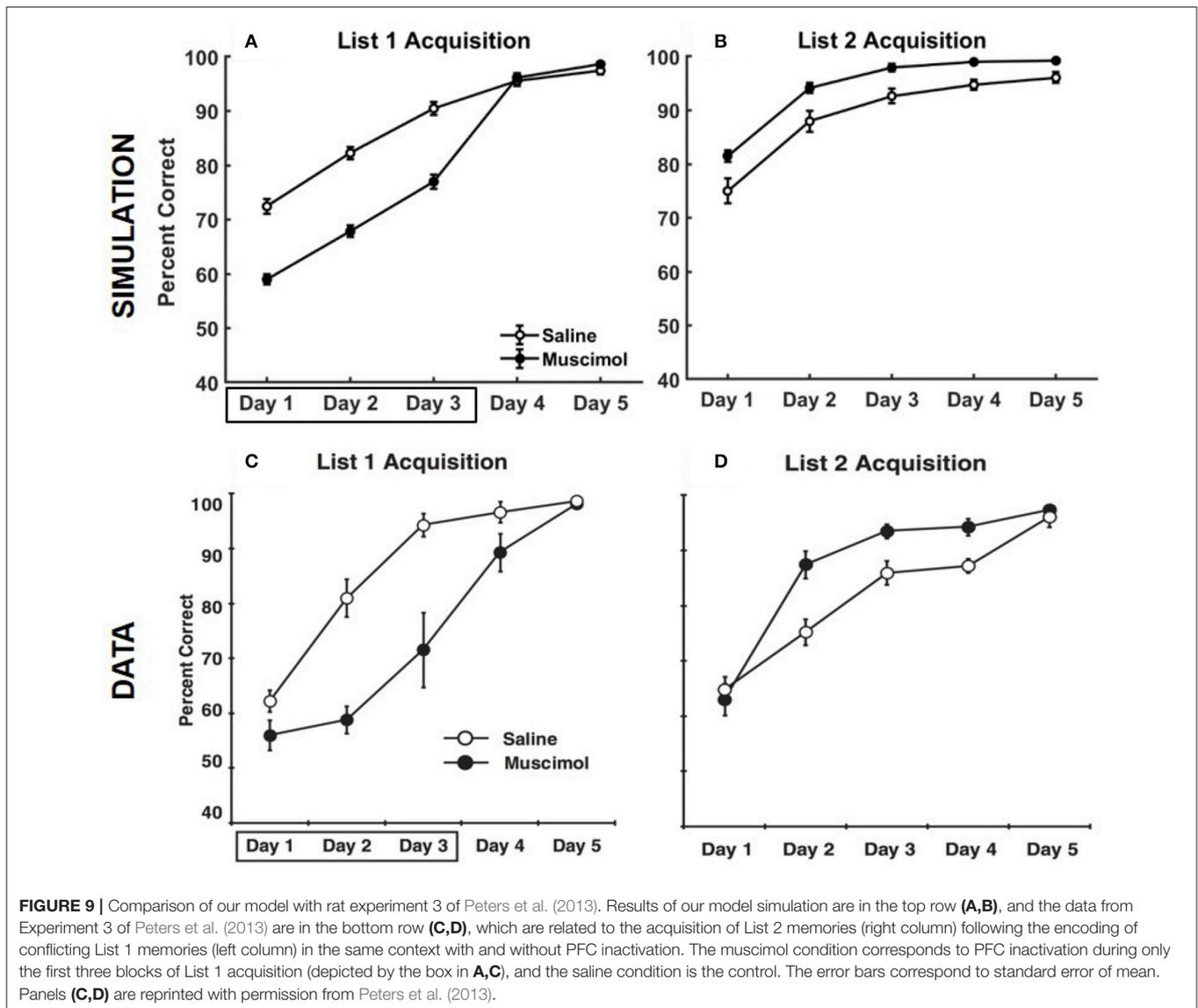
The errors that occur during the learning of the conflicting List 2 memories are of two types: perseverative and non-perseverative. Consider an odor that was employed in both Lists 1 and 2 and was rewarded in List 1, but not in List 2 by design. If the rat makes an error by choosing it during List 2 acquisition, it would be a perseverative error (i.e., choosing the same odor despite the change from List 1 to List 2). Along the same lines, a non-perseverative error is one in which a new odor in a List 2 problem is chosen when it is paired with an unrewarded odor



from List 1. The number of errors made by the simulated rats during the five training blocks of List 2 was analyzed by a two-way repeated measures ANOVA with inactivation condition and error type as factors. As in the data, there were more perseverative errors than non-perseverative errors [significant main effect of error type, $F_{(1, 36)} = 7.79, p = 0.0084$], and while the simulated muscimol rats made more errors than the simulated control rats [significant main effect of inactivation condition, $F_{(1, 36)} = 9.73, p = 0.0036$], the proportion of perseverative and non-perseverative errors was not significantly different between the simulated control and muscimol rats [no interaction between error type and inactivation condition, $F = 0.16, p = 0.70$; two-way ANOVA]; see **Figures 8C,F**.

When learning a List 2 problem in the same context, interference occurs because of the activation of conflicting memory traces (related to the corresponding List 1 problem) in each of the hippocampal subfields in general (and CA3 in particular). The projections from each CA3 engram corresponding to a List 1 associative memory to CA1, as well as auto-associative projections within CA3, need to

overcome their prior tuning before new connections can be established to be able to correctly retrieve the List 2 memory. This impairment, which slows encoding, does not occur if List 2 memories are acquired in a new context; note that the “Saline Different” condition is uniformly better than “Saline Same” in **Figure 8A**, as the data shown in **Figure 8D**. When PFC is inactivated, there is even greater interference because the lack of top-down bias over DG allows innumerable traces within various contextual ensembles of DG cells that respond partially to the current association to be activated (see **Figure 6**). This includes the memory traces corresponding to List 1. Acquisition of List 2 memories was thus drastically affected by PFC inactivation, and also did not benefit from employing a different context. In the model, both perseverative and non-perseverative errors during List 2 learning occur because of interference from prior associations, or lack thereof, to the component of the memory engram signifying the rewarded odor. So, the relative proportion of perseverative and non-preservative errors did not significantly differ between the simulated control and muscimol rats.



4.4. Long-Term Effects of PFC Inactivation

Figure 9 provides the results of our model simulation of Experiment 3 of Peters et al. (2013), showing rats that learned the first set of odor memory associations with a deactivated PFC were surprisingly better at acquiring the second set of interfering memories in the same context under normal PFC operation. Similar to Experiment 1A, the performance data of the simulated rats from the first three training blocks of List 1 was analyzed using a two-way repeated measures ANOVA with inactivation condition and training block as factors. As in the data, simulated muscimol rats were initially less accurate in learning List 1 memories compared to simulated control rats [significant main effect of inactivation condition, $F_{(1, 54)} = 186.28, p = 0$]. Also, there was a significant main effect of training block [$F_{(2, 54)} = 105.97, p = 0$], but no interaction between inactivation condition and training block [$F_{(2, 54)} = 0.1, p = 0.91$]. Though the simulated muscimol rats were

initially impaired, the performance was not different between the saline and muscimol conditions by the fifth training block [$t_{(18)} = -1.41, p = 0.17$]. Next, the performance data of the simulated rats from the five training blocks of List 2 was analyzed using a two-way repeated measures ANOVA with inactivation condition and training block as factors. Note that muscimol rats had their PFC inactivated only during the learning of List 1 for the first three blocks, and not during the learning of List 2. As in the data, simulated muscimol rats were surprisingly more accurate in learning List 2 memories compared to simulated control rats [significant main effect of inactivation condition, $F_{(1, 54)} = 21.18, p = 0$]. Also, there was a significant main effect of training block [$F_{(2, 54)} = 61.74, p = 0$], but no interaction between inactivation condition and training block [$F_{(2, 54)} = 0.09, p = 0.92$]. In our model, List 1 memories for muscimol rats were initially acquired with activities enabled for all contextual ensembles in DG. As a result, CA3 memory engrams gradually

became tuned to DG cells across various contextual ensembles during initial encoding. Although simulated muscimol rats did continue learning for two additional blocks with the top-down contextual bias over DG in place (i.e., with only one appropriate ensemble of DG cells facilitated), we suggest that the distributed projections from across various contextual DG ensembles to CA3 for the List 1 memory engrams were still sufficiently preserved by the time List 2 learning began. During List 2 acquisition in the same context, the activation of the conflicting memory traces in CA3 for simulated muscimol rats was not as strong as that for simulated control rats because the DG support for pattern completion in CA3 was critically shrunk to just one contextual ensemble. In other words, List 2 acquisition underwent relatively weaker interference from previously acquired List 1 memory traces for simulated muscimol rats, explaining the reversal in the encoding dynamics for List 2 between the muscimol and control rats.

5. DISCUSSION

The main contribution of this article is its computational and algorithmic explanation for how the hippocampus could employ top-down contextual signals originating in prefrontal cortex (PFC) to reduce interference when encoding and retrieving conflicting associative memories experienced in different contexts. Specifically, we demonstrated that external signals that modulate cell excitability to select a contextually relevant subset of dentate gyrus (DG) neurons play a key role in replicating the rat data on various negative as well as positive effects of inactivating PFC on associative memory formation and recall (Navawongse and Eichenbaum, 2013; Peters et al., 2013). Pre-training the model to learn conflicting memory associations in contexts other than those used in the experiments is the other key element we found necessary to account for the multifarious aspects of the behavioral data. This pre-training qualitatively simulates prior experiences of the adult lab rats, which can conflict with target memories when the contextual bias is suppressed as a result of PFC inactivation. In this regard, even though Experiment 1 of Peters et al. (2013) did not directly manipulate context (see **Figures 5, 7**), we show that inactivating PFC can lead to additional corollary activation of interfering memory traces from multiple other, previously experienced contexts. Along these lines, Experiment 2 of Peters et al. (2013) and our model simulations (see **Figure 8**) show that under PFC inactivation, the advantage of a new context in reducing interference while learning a conflicting set of memory associations is lost.

As mentioned in the Introduction section, we define “context” as the functional background with respect to which salient objects in the foreground (comprising an episode) are experienced. Context is not simply a collection of spatial locations that encompass an environment; it is a more holistic construct based on rapidly fusing several relevant spatial and non-spatial cues. Its chief function is to aid in disambiguating the perception of stimuli, recall of memories, and choice of actions (Doboli et al., 2000; Lee and Lee, 2013). Several studies have

implicated PFC in rapid extraction of contextual information, which is then leveraged for top-down contextual facilitation of memory-related functions ranging from recognition to recall (Bar, 2004). Moreover, different sub-regions of PFC have been shown to be involved in different aspects of cue-based episodic memory retrieval (Dobbins et al., 2002). This emerging understanding is in concert with the notion that PFC is the source of the external contextual bias to the hippocampus.

Alternative theories posit that contextual information arrives at the hippocampus through the entorhinal cortex (EC) as an explicit input that is associated with other cortical signals coding for salient objects to form an episodic memory (O’Reilly and Rudy, 2001; O’Reilly, 2004). These Complementary Learning Systems models (O’Reilly et al., 2014) perform pattern completion for the AB-AC associative learning task (Barnes and Underwood, 1959) using EC as the gateway for cortical inputs representing list elements and context, and result in the encoding of context and list items as a joint representation within DG. Although the CLS models could, in theory, simulate context-dependent memory retrieval by providing context as one of the recall cues, they would not exhibit contextual clustering of cells within the DG. Very similar list items in different contexts would result in significantly overlapping DG ensembles across contexts, which would be inconsistent with well-known data. In particular, the degree of overlap among sets of DG cells that get activated for various memories within two different contexts is about just 1% (Thompson and Best, 1989; Ramirez et al., 2013). Moreover, the selection of a new ensemble of DG cells can also be triggered by a change in the behavioral task within the same environment (Chawla et al., 2005; Leutgeb et al., 2007; Satvat et al., 2011; Schmidt et al., 2012; Deng et al., 2013). Recent experimental work has examined the role of a local inhibitory circuit in DG between granule cells and hilar perforant path (HIPP) cells that modulates the strength of contextual memories (Raza et al., 2017). In particular, the study found that the mechanism behind the formation of contextual granule cell ensembles in DG involves HIPP cells. To our knowledge, there has been only one computational model (Doboli et al., 2000) that simulates the recruitment and sustenance of contextual ensembles (“latent attractors”) of DG cells, while responding to time-varying inputs from EC. Indeed, as EC inputs change, different cells within the winner DG ensemble become active. In this model, contextual hilar cells within DG are pre-configured to receive excitatory projections from the corresponding granule cells, while they inhibit granule cells configured for the other contexts, and vice versa. Recruitment of an arbitrary ensemble can be achieved by an excitatory perturbation of the appropriate set of DG cells, pointing to control by an external source. This model, however, does not exploit the contextual coding in DG to understand interference-free pattern completion and memory retrieval in an adaptive hippocampus.

It is interesting to note that, in addition to its large size and highly sparse activity compared to other regions of the hippocampus, DG is one of the few brain areas where adult neurogenesis has been reported in mammalian brains including

humans (Kuhn et al., 1996). Moreover, the rate at which new DG granule cells are born in adult mice has been shown to covary with the variety and richness of their experiences (Kempermann et al., 1997, 1998). We speculate that the self-regulated birth of new DG cells continuing through adulthood is what enables animals and humans alike to acquire memories in ever-new contexts through their lifespans without catastrophic forgetting of older contextual memories. The effect of neurogenesis in the hippocampus has been examined in computational models that simulate this experience-dependent recruitment of new granule cells and their extended process of maturation resulting in changes in their excitability (Aimone et al., 2009). This could facilitate the use of temporal information as another aspect of the contextual modulation in the encoding process (Rangel et al., 2014). Despite the high excitability of immature DG granule cells, they do not respond broadly to afferent activity from EC due to their much sparser connectivity than mature cells, supporting their role in providing orthogonal representations for encoding new memories that are similar to those previously experienced (Dieni et al., 2016). Another modeling study has investigated the advantageous role of a heterogeneous population of DG granule cells for encoding an increasingly large number of new inputs (Severa et al., 2017). In addition, the reduction of neurogenesis in DG has been experimentally observed when using focal cranial irradiation to eliminate the advantage of a new context for the acquisition of conflicting odor discrimination memories (Luu et al., 2012). While DG has been previously shown to be critical for both the encoding and recall of contextual memories of emotional experiences in rats (Hernández-Rabaza et al., 2008; Bernier et al., 2017), a recent optogenetic study has demonstrated that successful contextual fear conditioning relies on adult-born DG granule cells arising from neurogenesis (Huckleberry et al., 2018).

Depue (2012) reviewed prior literature to develop a conceptual model of prefrontal-hippocampal interactions in which PFC plays the key role of suppressing task-irrelevant memories. Three specific hypotheses for PFC-mediated inhibitory control of cue-triggered memory retrieval were considered (see their **Figure 1**); namely, (1) “direct inhibition” of specific memory engrams in the hippocampus (Anderson et al., 2004), (2) “reactivation inhibition” of the memory retrieval output of the hippocampus to its neocortical targets including EC (Depue et al., 2007), and (3) “competitive attentional inhibition” by shifting attention to distractor memory engrams in the hippocampus (Wimber et al., 2008). Depue (2012) concluded that prior behavioral and imaging studies related to white bear suppression, directed forgetting, think/no-think tasks provide support only for the direct inhibition and reactivation inhibition hypotheses. While Depue’s model relates more to proactive memory suppression than to context-dependent memory retrieval, our hypothesis of PFC inhibitory control to suppress inappropriate contextual ensembles of DG cells is closer to the direct inhibition hypothesis. Large-scale optical imaging of single cells in each hippocampal subfield and the superficial and deep layers of both LEC and MEC will be required to clarify the precise modus operandi of PFC inhibition on memory circuits.

5.1. Possible Indirect Route for PFC Bias of Hippocampus

No direct mechanism for PFC control of inhibitory bias on DG cells has been identified, but there are some possibilities for an indirect connection. PFC can disynaptically modulate hippocampal cell excitability via hypothalamus and thalamus (Wyss et al., 1979). Engin et al. (2015) showed that tonic inhibition of DG granule cells by GABAergic interneurons, which regulates their excitability, is indeed critical for minimizing interference in the encoding of similar or overlapping memories experienced at different times. Such inhibitory control signals can originate from PFC and arrive at DG by means of hypothalamic supramammillary nucleus (SuMN) (Vertes and McKenna, 2000; Nakanishi et al., 2001). SuMN is connected strongly to DG as well as to several structures that project to the hippocampus, and has been suggested to gate the flow of neural signals within the hippocampus (Vertes, 1992) possibly by modulating cell excitability (Mizumori et al., 1989; Jiang and Khanna, 2006). SuMN has also been shown to key for achieving memory specificity (Xu and Südhof, 2013). Fear conditioning in the absence of SuMN signals to DG would lead to encoding of fear memories across various DG contextual ensembles, resulting in over-generalization to other contexts. While the PFC-SuMN-DG circuit seems pertinent as a mechanism for inhibitory bias on DG, it is unclear how the relatively small number of neurons in the SuMN are capable of biasing the huge number of very specific contextual ensembles that may be activated in the DG.

5.2. Testable Hypotheses

Future experiments are in order to clarify the precise influences of contextual signals on hippocampal processing in general and DG in particular. In this regard, our mechanistic explanation makes several clear predictions. First, in order to match the data from Experiments 1A, 1B, and 2 of Peters et al. (2013), it was crucial that we pre-train the hippocampus with interfering memories in contexts different than those used to train the target memories in the experiments. This would suggest that PFC inactivation should not affect hippocampal function as assessed in these experiments for animals raised in a single environment, compared to animals with a large number of experiences across multiple environments. In other words, inexperienced animals do not have a storehouse of contradictory associations across different contexts that are unleashed when PFC is inactivated. On the other hand, we predict that the facilitation of List 2 acquisition due to prior PFC inactivation during learning of conflicting List 1 memories, seen in Experiment 3 of Peters et al. (2013), would still hold for these inexperienced rats. This is because the interference from List 1 on the acquisition of List 2 memories would be instantiated within the time course of the experiment. **Figures 10, 11** (top row) show quantitative predictions of our model in the absence of pre-training for Experiments 1A and 3 of Peters et al. (2013). As can be seen in **Figure 10** with respect to Experiment 1A of Peters et al. (2013), the difference in memory behavior between the saline and muscimol conditions is eliminated during concurrent acquisition [no effect of inactivation condition, $F_{(1, 54)} = 0.15, p = 0.70$] and greatly reduced during concurrent performance (from 30.45%

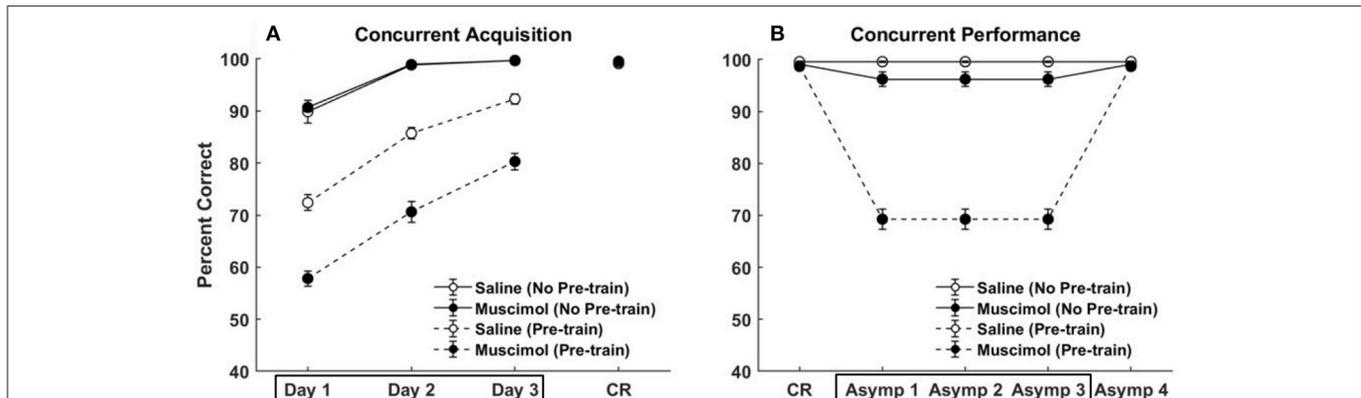


FIGURE 10 | Modeling rat experiment 1A of Peters et al. (2013) without pre-training. Panels (A,B) show our model simulation results in the absence of pre-training any conflicting memories, on the effects of PFC inactivation during concurrent acquisition (A) and concurrent performance (B) of multiple odor discrimination memories. As in Figure 5, the muscimol condition corresponds to PFC inactivation during the first three training blocks (depicted by the box in A), and the saline condition is the control. CR on the x-axis in both panels refers to the training block in which rats reached criterion performance. Control rats with saline injection during acquisition received muscimol infusion during the three test blocks following training to criterion (depicted by the box in B), but not the muscimol rats. Our model simulation results with pre-training, which are shown in Figure 5, are repeated here with dashed lines for comparison. The error bars correspond to standard error of mean.

drop in Figure 5B to 3.36% drop in Figure 10B). In particular, there was no effect of inactivation condition during concurrent acquisition [$F_{(1, 54)} = 0.15, p = 0.70$]. However, as can be seen in Figure 11 (top row) with respect to Experiment 3 of Peters et al. (2013), while the difference in memory behavior between the saline and muscimol conditions is eliminated during List 1 acquisition [no effect of inactivation condition, $F_{(1, 54)} = 2.52, p = 0.12$], the facilitation in List 2 acquisition for muscimol rats remains [significant main effect of inactivation condition, $F_{(1, 54)} = 28.39, p = 0$]. Furthermore, it seems that the lack of previous conflicting experiences also renders the task of acquiring a list of 8 discrimination memories (namely, List 1) from scratch much easier due to the absence of any implicit interference within the hippocampal network (especially CA3).

We also explored the effects of varying the sparseness of activity in DG on the acquisition dynamics of conflicting List 1 and List 2 memories for Experiment 3 of Peters et al. (2013). These simulations were done in the absence of pre-training any conflicting memories to better isolate the role of sparse activities in DG for this data (see Figure 11). As DG activities become more dense, the pattern separability within DG for similar memories would be expected to be weakened leading to slower memory formation. However, we see no apparent differences in List 1 acquisition across the three sparseness levels (compare List 1 results from the 3 rows in Figure 11). While the memory capacity of our hippocampal network is reduced with more dense DG activities, it is likely much higher than the 8 discrimination memories that are being acquired in the case of no pre-training. Interestingly, the facilitation in List 2 acquisition for muscimol rats seems to exhibit a gradual decrease as sparseness levels in DG are reduced [$k = 10\%$: significant main effect of inactivation condition, $F_{(1, 54)} = 17.89, p = 0.0001$; $k = 25\%$: significant main effect of inactivation condition, $F_{(1, 54)} = 6.11, p = 0.0166$]. For the muscimol condition, as the activities in DG become more

dense across all contextual ensembles (with increasing k), the proportional drop in the absolute number of active DG cells from List 1 acquisition to List 2 acquisition would likely not preclude the pattern completion of List 1 memory traces in the recurrent CA3 network (and thereby in CA1 and ECout) to interfere with the formation of List 2 memories.

Second, the number of active DG cells should increase when PFC is inactivated, because numerous contextual ensembles are enabled due to the lack of top-down inhibition, or cell excitability modulation, from PFC. This can be verified using existing optical sensing technologies to measure *c-Fos* (or some other immediate-early gene) expression per unit area in DG (Liu et al., 2012; Ramirez et al., 2013). Third, if the mossy fibers from DG to CA3 are lesioned, then all the contextual effects simulated here from the Navawongse and Eichenbaum (2013) and Peters et al. (2013) studies should be eliminated or greatly reduced. In particular, these animals will exhibit significant impairments in behavioral tasks that require distinguishing similar or overlapping memories acquired in different contexts.

6. CONCLUSION

The new understanding of how PFC exerts top-down control of hippocampal memory circuits can be leveraged to causally enhance memory performance. Prior rat data suggests increased top-down attention to the context in which events occur improves the long-term stability of memories (Kentros et al., 2004; Muzzio et al., 2009a,b). Moreover, a number of human studies demonstrating declarative memory enhancement by transcranial current stimulation (tCS) during encoding have targeted the PFC (Javadi and Walsh, 2012; Javadi and Cheng, 2013). Interestingly, transfer learning, where a schema learned in one context is applied to a different context, has been shown to be inversely related to functional connectivity between

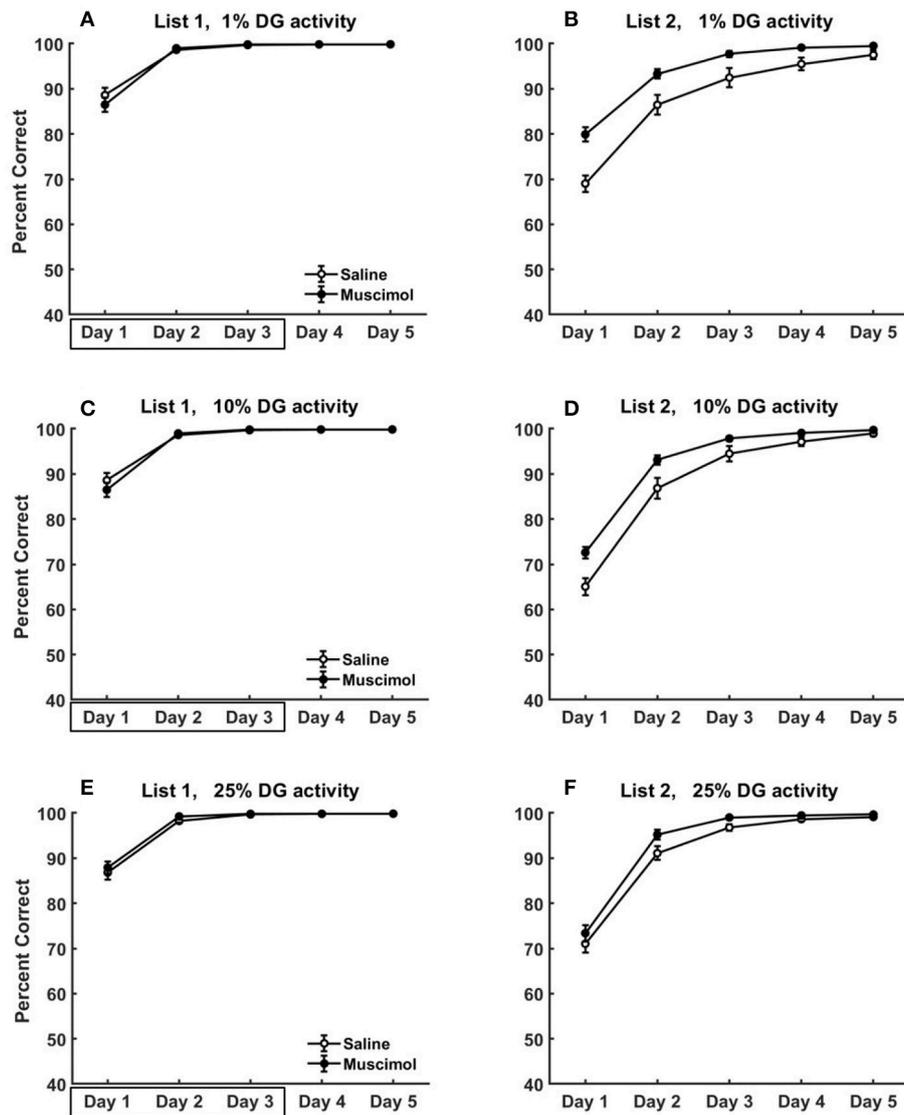


FIGURE 11 | Modeling rat experiment 3 of Peters et al. (2013) by varying the sparseness of activity in DG. The three rows provide our model simulation results in the absence of pre-training any conflicting memories, and at different sparseness levels of activity in DG (top row: $k = 1\%$, middle: $k = 10\%$, bottom: $k = 25\%$), on the acquisition of List 2 memories (**B,D,F**) following the encoding of conflicting List 1 memories (**A,C,E**) in the same context with and without PFC inactivation. Note that as k is increased (see Equation 5) for DG, the activities in DG become more dense (i.e., less sparse). As in **Figure 9**, the muscimol condition corresponds to PFC inactivation during only the first three blocks of List 1 acquisition (depicted by the box in **A,C,E**), and the saline condition is the control. The error bars correspond to standard error of mean.

PFC and hippocampus (van Kesteren et al., 2012; Gerraty et al., 2014). When contextual biasing signals from PFC on DG are weak, memories learned in one context generalize to other contexts. Stronger contextual signals from PFC lead to greater memory specificity during encoding, resulting in more stable memories. In summary, our memory model provides a computational account for the top-down contextual modulation of encoding and recall processes in the hippocampus by postulating contextual recruitment of independent pools of DG cells that critically depends on normal PFC operation.

AUTHOR CONTRIBUTIONS

PP conceived the study. MH and PP performed the simulation experiments. MH, PP, and RB analyzed the data. PP and MH wrote the paper.

FUNDING

This material is based upon work supported by the Intelligence Advanced Research Projects Agency (IARPA) via Department of the Interior (DOI) Contract No. D10PC20021,

and by the Defense Advanced Research Projects Agency (DARPA) via SPAWAR Systems Center Pacific (SSC Pacific) Contract No. N66001-14-C-4066 (Restoring Active Memory: RAM) and via Air Force Research Laboratory (AFRL) Contract No. FA8750-18-C-0103 (Lifelong Learning

Machines: L2M). 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 IARPA, DOI, DARPA, SSC Pacific, AFRL, or the US Government.

REFERENCES

- Aimone, J. B., Wiles, J., and Gage, F. H. (2009). Computational influence of adult neurogenesis on memory encoding. *Neuron* 61, 187–202. doi: 10.1016/j.neuron.2008.11.026
- Aisa, B., Mingus, B., and O'Reilly, R. (2008). The emergent neural modeling system. *Neural Netw.* 21, 1146–1152. doi: 10.1016/j.neunet.2008.06.016
- Amaral, D., and Witter, M. (1995). Hippocampal formation: the rat nervous system. *Neural Netw.* 21, 1146–1152.
- Amaral, D. G., and Witter, M. P. (1989). The three-dimensional organization of the hippocampal formation: a review of anatomical data. *Neuroscience* 31, 571–591. doi: 10.1016/0306-4522(89)90424-7
- Anderson, M. C., Ochsner, K. N., Kuhl, B., Cooper, J., Robertson, E., Gabrieli, S. W., et al. (2004). Neural systems underlying the suppression of unwanted memories. *Science* 303, 232–235. doi: 10.1126/science.1089504
- Bar, M. (2004). Visual objects in context. *Nat. Rev. Neurosci.* 5:617. doi: 10.1038/nrn1476
- Barnes, J. M., and Underwood, B. J. (1959). "Fate" of first-list associations in transfer theory. *J. Exp. Psychol.* 58:97. doi: 10.1037/h0047507
- Bernier, B. E., Lacagnina, A. F., Ayoub, A., Shue, F., Zemelman, B. V., Krasne, F. B., et al. (2017). Dentate gyrus contributes to retrieval as well as encoding: evidence from context fear conditioning, recall, and extinction. *J. Neurosci.* 37, 6359–6371. doi: 10.1523/JNEUROSCI.3029-16.2017
- Brincat, S. L., and Miller, E. K. (2015). Frequency-specific hippocampal-prefrontal interactions during associative learning. *Nat. Neurosci.* 18:576. doi: 10.1038/nn.3954
- Brun, V. H., Otnass, M. K., Molden, S., Steffenach, H. A., Witter, M. P., Moser, M. B., et al. (2002). Place cells and place recognition maintained by direct entorhinal-hippocampal circuitry. *Science* 296, 2243–2246. doi: 10.1126/science.1071089
- Chawla, M. K., Guzowski, J. F., Ramirez-Amaya, V., Lipa, P., Hoffman, K. L., Marriott, L. K., et al. (2005). Sparse, environmentally selective expression of Arc RNA in the upper blade of the rodent fascia dentata by brief spatial experience. *Hippocampus* 15, 579–586. doi: 10.1002/hipo.20091
- Deng, W., Mayford, M., and Gage, F. H. (2013). Selection of distinct populations of dentate granule cells in response to inputs as a mechanism for pattern separation in mice. *eLife* 2:e00312. doi: 10.7554/eLife.00312
- Depue, B. E. (2012). A neuroanatomical model of prefrontal inhibitory modulation of memory retrieval. *Neurosci. Biobehav. Rev.* 36, 1382–1399. doi: 10.1016/j.neubiorev.2012.02.012
- Depue, B. E., Curran, T., and Banich, M. T. (2007). Prefrontal regions orchestrate suppression of emotional memories via a two-phase process. *Science* 317, 215–219. doi: 10.1126/science.1139560
- Dieni, C. V., Panichi, R., Aimone, J. B., Kuo, C. T., Wadiche, J. I., and Overstreet-Wadiche, L. (2016). Low excitatory innervation balances high intrinsic excitability of immature dentate neurons. *Nat. Commun.* 7:11313. doi: 10.1038/ncomms11313
- Dobbins, I. G., Foley, H., Schacter, D. L., and Wagner, A. D. (2002). Executive control during episodic retrieval: multiple prefrontal processes subsolve source memory. *Neuron* 35, 989–996. doi: 10.1016/S0896-6273(02)00858-9
- Doboli, S., Minai, A. A., and Best, P. J. (2000). Latent attractors: a model for context-dependent place representations in the hippocampus. *Neural Comput.* 12, 1009–1043. doi: 10.1162/089976600300015484
- Douchamps, V., Jeewajee, A., Blundell, P., Burgess, N., and Lever, C. (2013). Evidence for encoding versus retrieval scheduling in the hippocampus by theta phase and acetylcholine. *J. Neurosci.* 33, 8689–8704. doi: 10.1523/JNEUROSCI.4483-12.2013
- Engin, E., Zarnowska, E. D., Benke, D., Tsvetkov, E., Sigal, M., Keist, R., et al. (2015). Tonic inhibitory control of dentate gyrus granule cells by $\alpha 5$ -containing GABAA receptors reduces memory interference. *J. Neurosci.* 35, 13698–13712. doi: 10.1523/JNEUROSCI.1370-15.2015y
- Gerraty, R. T., Davidow, J. Y., Wimmer, G. E., Kahn, I., and Shohamy, D. (2014). Transfer of learning relates to intrinsic connectivity between hippocampus, ventromedial prefrontal cortex, and large-scale networks. *J. Neurosci.* 34, 11297–11303. doi: 10.1523/JNEUROSCI.0185-14.2014
- Grossberg, S. (1976). Adaptive pattern classification and universal recoding: I. Parallel development and coding of neural feature detectors. *Biol. Cybern.* 23, 121–134. doi: 10.1007/BF00344744
- Hassabis, D., Kumaran, D., Summerfield, C., and Botvinick, M. (2017). Neuroscience-inspired artificial intelligence. *Neuron* 95, 245–258. doi: 10.1016/j.neuron.2017.06.011
- Hasselmo, M. E., Bodelón, C., and Wyble, B. P. (2002). A proposed function for hippocampal theta rhythm: separate phases of encoding and retrieval enhance reversal of prior learning. *Neural Comput.* 14, 793–817. doi: 10.1162/089976602317318965
- Henze, D. A., Card, J. P., Barrionuevo, G., and Ben-Ari, Y. (1997). Large amplitude miniature excitatory postsynaptic currents in hippocampal CA3 pyramidal neurons are of mossy fiber origin. *J. Neurophysiol.* 77, 1075–1086. doi: 10.1152/jn.1997.77.3.1075
- Hernández-Rabaza, V., Hontecillas-Prieto, L., Velázquez-Sánchez, C., Ferragud, A., Pérez-Villaba, A., Arcusa, A., et al. (2008). The hippocampal dentate gyrus is essential for generating contextual memories of fear and drug-induced reward. *Neurobiol. Learn. Mem.* 90, 553–559. doi: 10.1016/j.nlm.2008.06.008
- Hinton, G. E. (1990). "Connectionist learning procedures," in *Machine Learning, Vol. III*, eds Y. Kodratoff and R. S. Michalski (San Mateo, CA: Morgan Kaufmann Press), 555–610.
- Huckleberry, K. A., Shue, F., Copeland, T., Chitwood, R. A., Yin, W., and Drew, M. R. (2018). Dorsal and ventral hippocampal adult-born neurons contribute to context fear memory. *Neuropsychopharmacology* 43, 2487–2496. doi: 10.1038/s41386-018-0109-6
- Javadi, A. H., and Cheng, P. (2013). Transcranial direct current stimulation (tDCS) enhances reconsolidation of long-term memory. *Brain Stimul.* 6, 668–674. doi: 10.1016/j.brs.2012.10.007
- Javadi, A. H., and Walsh, V. (2012). Transcranial direct current stimulation (tDCS) of the left dorsolateral prefrontal cortex modulates declarative memory. *Brain Stimul.* 5, 231–241. doi: 10.1016/j.brs.2011.06.007
- Jiang, F., and Khanna, S. (2006). Microinjection of carbachol in the supramammillary region suppresses CA1 pyramidal cell synaptic excitability. *Hippocampus* 16, 891–905. doi: 10.1002/hipo.20219
- Jones, M. W., and Wilson, M. A. (2005). Theta rhythms coordinate hippocampal-prefrontal interactions in a spatial memory task. *PLoS Biol.* 3:e402. doi: 10.1371/journal.pbio.0030402
- Jung, M. W., and McNaughton, B. L. (1993). Spatial selectivity of unit activity in the hippocampal granular layer. *Hippocampus* 3, 165–182. doi: 10.1002/hipo.450030209
- Kempermann, G., Kuhn, H. G., and Gage, F. H. (1997). More hippocampal neurons in adult mice living in an enriched environment. *Nature* 386:493. doi: 10.1038/386493a0
- Kempermann, G., Kuhn, H. G., and Gage, F. H. (1998). Experience-induced neurogenesis in the senescent dentate gyrus. *J. Neurosci.* 18, 3206–3212. doi: 10.1523/JNEUROSCI.18-09-03206.1998
- Kentros, C. G., Agnihotri, N. T., Streater, S., Hawkins, R. D., and Kandel, E. R. (2004). Increased attention to spatial context increases both place field stability and spatial memory. *Neuron* 42, 283–295. doi: 10.1016/S0896-6273(04)00192-8
- Ketz, N., Morkonda, S. G., and O'Reilly, R. C. (2013). Theta coordinated error-driven learning in the hippocampus. *PLoS Computat. Biol.* 9:e1003067. doi: 10.1371/journal.pcbi.1003067

- Kim, J., Delcasso, S., and Lee, I. (2011). Neural correlates of object-in-place learning in hippocampus and prefrontal cortex. *J. Neurosci.* 31, 16991–17006. doi: 10.1523/JNEUROSCI.2859-11.2011
- Kubie, J. L., and Muller, R. U. (1991). Multiple representations in the hippocampus. *Hippocampus* 1, 240–242. doi: 10.1002/hipo.450010305
- Kuhn, H. G., Dickinson-Anson, H., and Gage, F. H. (1996). Neurogenesis in the dentate gyrus of the adult rat: age-related decrease of neuronal progenitor proliferation. *J. Neurosci.* 16, 2027–2033. doi: 10.1523/JNEUROSCI.16-06-02027.1996
- Kunec, S., Hasselmo, M. E., and Kopell, N. (2005). Encoding and retrieval in the CA3 region of the hippocampus: a model of theta-phase separation. *J. Neurophysiol.* 94, 70–82. doi: 10.1152/jn.00731.2004
- Lee, I., and Lee, S. H. (2013). Putting an object in context and acting on it: neural mechanisms of goal-directed response to contextual object. *Rev. Neurosci.* 24, 27–49. doi: 10.1515/revneuro-2012-0073
- Leutgeb, J. K., Leutgeb, S., Moser, M. B., and Moser, E. I. (2007). Pattern separation in the dentate gyrus and CA3 of the hippocampus. *Science* 315, 961–966. doi: 10.1126/science.1135801
- Leutgeb, S., Leutgeb, J. K., Barnes, C. A., Moser, E. I., McNaughton, B. L., and Moser, M. B. (2005). Independent codes for spatial and episodic memory in hippocampal neuronal ensembles. *Science* 309, 619–623. doi: 10.1126/science.1114037
- Leutgeb, S., Leutgeb, J. K., Treves, A., Moser, M. B., and Moser, E. I. (2004). Distinct ensemble codes in hippocampal areas CA3 and CA1. *Science* 305, 1295–1298. doi: 10.1126/science.1100265
- Liao, Q., Leibo, J. Z., and Poggio, T. A. (2016). “How important is weight symmetry in backpropagation?” in AAAI (Phoenix, AZ), 1837–1844.
- Lillicrap, T. P., Cownden, D., Tweed, D. B., and Akerman, C. J. (2016). Random synaptic feedback weights support error backpropagation for deep learning. *Nat. Commun.* 7:13276. doi: 10.1038/ncomms13276
- Liu, X., Ramirez, S., Pang, P. T., Puryear, C. B., Govindarajan, A., Deisseroth, K., et al. (2012). Optogenetic stimulation of a hippocampal engram activates fear memory recall. *Nature* 484:381. doi: 10.1038/nature11028
- Luu, P., Sill, O. C., Gao, L., Becker, S., Wojtowicz, J. M., and Smith, D. M. (2012). The role of adult hippocampal neurogenesis in reducing interference. *Behav. Neurosci.* 126:381. doi: 10.1037/a0028252
- Markus, E. J., Qin, Y. L., Leonard, B., Skaggs, W. E., McNaughton, B. L., and Barnes, C. A. (1995). Interactions between location and task affect the spatial and directional firing of hippocampal neurons. *J. Neurosci.* 15, 7079–7094. doi: 10.1523/JNEUROSCI.15-11-07079.1995
- Marr, D. (1971). Simple memory: a theory for archicortex. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* 262:23. doi: 10.1098/rstb.1971.0078
- McClelland, J. L., McNaughton, B. L., and O'Reilly, R. C. (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.* 102:419. doi: 10.1037/0033-295X.102.3.419
- McNaughton, B. L., and Morris, R. G. (1987). Hippocampal synaptic enhancement and information storage within a distributed memory system. *Trends Neurosci.* 10, 408–415. doi: 10.1016/0166-2236(87)90011-7
- Miller, E. K., and Cohen, J. D. (2001). An integrative theory of prefrontal cortex function. *Ann. Rev. Neurosci.* 24, 167–202. doi: 10.1146/annurev.neuro.24.1.167
- Mizumori, S., Mcnaughton, B. L., and Barnes, C. A. (1989). A comparison of supramammillary and medial septal influences on hippocampal field potentials and single-unit activity. *J. Neurophysiol.* 61, 15–31. doi: 10.1152/jn.1989.61.1.15
- Muzzio, I. A., Kentros, C., and Kandel, E. (2009a). What is remembered? Role of attention on the encoding and retrieval of hippocampal representations. *J. Physiol.* 587, 2837–2854. doi: 10.1113/jphysiol.2009.172445
- Muzzio, I. A., Levita, L., Kulkarni, J., Monaco, J., Kentros, C., Stead, M., et al. (2009b). Attention enhances the retrieval and stability of visuospatial and olfactory representations in the dorsal hippocampus. *PLoS Biol.* 7:e1000140. doi: 10.1371/journal.pbio.1000140
- Naber, P. A., Lopes da Silva, F. H., and Witter, M. P. (2001). Reciprocal connections between the entorhinal cortex and hippocampal fields CA1 and the subiculum are in register with the projections from CA1 to the subiculum. *Hippocampus* 11, 99–104. doi: 10.1002/hipo.1028
- Nakanishi, K., Saito, H., and Abe, K. (2001). The supramammillary nucleus contributes to associative EPSP-spike potentiation in the rat dentate gyrus *in vivo*. *Euro. J. Neurosci.* 13, 793–800. doi: 10.1046/j.1460-9568.2001.01446.x
- Nakazawa, K., Quirk, M. C., Chitwood, R. A., Watanabe, M., Yeeckel, M. F., Sun, L. D., et al. (2002). Requirement for hippocampal CA3 NMDA receptors in associative memory recall. *Science* 297, 211–218. doi: 10.1126/science.1071795
- Navawongse, R., and Eichenbaum, H. (2013). Distinct pathways for rule-based retrieval and spatial mapping of memory representations in hippocampal neurons. *J. Neurosci.* 33, 1002–1013. doi: 10.1523/JNEUROSCI.3891-12.2013
- O'Reilly, R. C. (2004). “The division of labor between the neocortex and hippocampus,” in *Connectionist Models in Cognitive Psychology*, ed G. Houghton (New York, NY: Psychology Press), 143.
- O'Reilly, R. C., Bhattacharyya, R., Howard, M. D., and Ketz, N. (2014). Complementary learning systems. *Cogn. Sci.* 38, 1229–1248. doi: 10.1111/j.1551-6709.2011.01214.x
- O'Reilly, R. C., and Munakata, Y. (2000). *Computational Explorations in Cognitive Neuroscience: Understanding the Mind by Simulating the Brain*. Cambridge, MA: MIT Press.
- O'Reilly, R. C., and Rudy, J. W. (2001). Conjunctive representations in learning and memory: principles of cortical and hippocampal function. *Psychol. Rev.* 108, 311–345. doi: 10.1037/0033-295X.108.2.311
- Peters, G. J., David, C. N., Marcus, M. D., and Smith, D. M. (2013). The medial prefrontal cortex is critical for memory retrieval and resolving interference. *Learn. Mem.* 20, 201–209. doi: 10.1101/lm.029249.112
- Peyrache, A., Khamassi, M., Benchenane, K., Wiener, S. I., and Battaglia, F. P. (2009). Replay of rule-learning related neural patterns in the prefrontal cortex during sleep. *Nat. Neurosci.* 12:919. doi: 10.1038/nn.2337
- Poo, M., Pignatelli, M., Ryan, T. J., Tonegawa, S., Bonhoeffer, T., Martin, K. C., et al. (2016). What is memory? The present state of the engram. *BMC Biol.* 14:40. doi: 10.1186/s12915-016-0261-6
- Preston, A. R., and Eichenbaum, H. (2013). Interplay of hippocampus and prefrontal cortex in memory. *Curr. Biol.* 23, R764–R773. doi: 10.1016/j.cub.2013.05.041
- Ramirez, S., Liu, X., Lin, P. A., Suh, J., Pignatelli, M., Redondo, R. L., et al. (2013). Creating a false memory in the hippocampus. *Science* 341, 387–391. doi: 10.1126/science.1239073
- Rangel, L. M., Alexander, A. S., Aimone, J. B., Wiles, J., Gage, F. H., Chiba, A. A., et al. (2014). Temporally selective contextual encoding in the dentate gyrus of the hippocampus. *Nat. Commun.* 5:3181. doi: 10.1038/ncomms4181
- Raza, S. A., Albrecht, A., Çalişkan, G., Müller, B., Demiray, Y. E., Ludewig, S., et al. (2017). HIPP neurons in the dentate gyrus mediate the cholinergic modulation of background context memory salience. *Nat. Commun.* 8:189. doi: 10.1038/s41467-017-00205-3
- Rumelhart, D. E., and Zipser, D. (1986). “Chapter 5: Feature discovery by competitive learning,” in *Parallel Distributed Processing: Explorations in the Microstructure of Cognition, Vol. 1*, eds J. A. Feldman, P. J. Hayes, and D. E. Rumelhart (Cambridge, MA: MIT Press), 151–193. USA.
- Satvat, E., Schmidt, B., Argraves, M., Marrone, D. F., and Markus, E. J. (2011). Changes in task demands alter the pattern of zif268 expression in the dentate gyrus. *J. Neurosci.* 31, 7163–7167. doi: 10.1523/JNEUROSCI.0094-11.2011
- Scellier, B., and Bengio, Y. (2017). Equilibrium propagation: Bridging the gap between energy-based models and backpropagation. *Front. Comput. Neurosci.* 11:24. doi: 10.3389/fncom.2017.00024
- Schmidt, B., Marrone, D. F., and Markus, E. J. (2012). Disambiguating the similar: the dentate gyrus and pattern separation. *Behav. Brain Res.* 226, 56–65. doi: 10.1016/j.bbr.2011.08.039
- Severa, W., Parekh, O., James, C. D., and Aimone, J. B. (2017). A combinatorial model for dentate gyrus sparse coding. *Neural Comput.* 29, 94–117. doi: 10.1162/NECO_a_00905
- Siapas, A. G., and Wilson, M. A. (1998). Coordinated interactions between hippocampal ripples and cortical spindles during slow-wave sleep. *Neuron* 21, 1123–1128. doi: 10.1016/S0896-6273(00)80629-7
- Squire, L. R., and Alvarez, P. (1995). Retrograde amnesia and memory consolidation: a neurobiological perspective. *Curr. Opin. Neurobiol.* 5, 169–177. doi: 10.1016/0959-4388(95)80023-9
- Thompson, L. T., and Best, P. J. (1989). Place cells and silent cells in the hippocampus of freely-behaving rats. *J. Neurosci.* 9, 2382–2390. doi: 10.1523/JNEUROSCI.09-07-02382.1989

- Treves, A., and Rolls, E. T. (1994). Computational analysis of the role of the hippocampus in memory. *Hippocampus* 4, 374–391. doi: 10.1002/hipo.450040319
- van Kesteren, M. T., Ruiters, D. J., Fernández, G., and Henson, R. N. (2012). How schema and novelty augment memory formation. *Trends Neurosci.* 35, 211–219. doi: 10.1016/j.tins.2012.02.001
- Vertes, R. P. (1992). PHA-L analysis of projections from the supramammillary nucleus in the rat. *J. Comparat. Neurol.* 326, 595–622. doi: 10.1002/cne.903260408
- Vertes, R. P., and McKenna, J. T. (2000). Collateral projections from the supramammillary nucleus to the medial septum and hippocampus. *Synapse* 38, 281–293. doi: 10.1002/1098-2396(20001201)38:3<281::AID-SYN7>3.0.CO;2-6
- Whittington, J. C. R., and Bogacz, R. (2017). An approximation of the error backpropagation algorithm in a predictive coding network with local hebbian synaptic plasticity. *Neural Comput.* 29, 1229–1262. doi: 10.1162/NECO_a_00949
- Wimber, M., Bäuml, K. H., Bergström, Z., Markopoulos, G., Heinze, H. J., and Richardson-Klavehn, A. (2008). Neural markers of inhibition in human memory retrieval. *J. Neurosci.* 28, 13419–13427. doi: 10.1523/JNEUROSCI.1916-08.2008
- Wyss, J. M., Swanson, L. W., and Cowan, W. M. (1979). A study of subcortical afferents to the hippocampal formation in the rat. *Neuroscience* 4, 463–476. doi: 10.1016/0306-4522(79)90124-6
- Xu, W., and Südhof, T. C. (2013). A neural circuit for memory specificity and generalization. *Science* 339, 1290–1295. doi: 10.1126/science.1229534

Conflict of Interest Statement: The authors were employed by HRL Laboratories, LLC, and have a pending patent application on a contextually biased memory system.

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2018 Pilly, Howard and Bhattacharyya. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.