# The Computational and Neural Bases of Context-Dependent Learning

## James B. Heald,[1] Daniel M. Wolpert,[1,2,*] and Máté Lengyel[2,3,*]

[1]Department of Neuroscience and Zuckerman Mind Brain Behavior Institute, Columbia University, New York, NY, USA; email: jamesbheald@gmail.com, wolpert@columbia.edu

[2]Computational and Biological Learning Lab, Department of Engineering, University of Cambridge, Cambridge, United Kingdom; email: m.lengyel@eng.cam.ac.uk

[3]Center for Cognitive Computation, Department of Cognitive Science, Central European University, Budapest, Hungary

*These authors contributed equally to this article

## Keywords

memory, learning, context-dependent learning, continual learning, Bayesian inference, neural computation, prefrontal cortex, hippocampus, thalamus, motor cortex

## Abstract

Flexible behavior requires the creation, updating, and expression of memories to depend on context. While the neural underpinnings of each of these processes have been intensively studied, recent advances in computational modeling revealed a key challenge in context-dependent learning that had been largely ignored previously: Under naturalistic conditions, context is typically uncertain, necessitating contextual inference. We review a theoretical approach to formalizing context-dependent learning in the face of contextual uncertainty and the core computations it requires. We show how this approach begins to organize a large body of disparate experimental observations, from multiple levels of brain organization (including circuits, systems, and behavior) and multiple brain regions (most prominently the prefrontal cortex, the hippocampus, and motor cortices), into a coherent framework. We argue that contextual inference may also be key to understanding continual learning in the brain. This theory-driven perspective places contextual inference as a core component of learning.

## Contents

## 1. INTRODUCTION

Flexible behavior relies on the ability to lay down a variety of memories, including for episodic events, motor skills, and decision-making. Research on learning and memory has, for the most part, studied how individual memories are laid down, recalled, and modified, one at a time. These studies use carefully controlled laboratory experiments, in which salient events and unambiguous stimuli cue when a memory needs to be stored, and when and which memory needs to be recalled, or modified. For example, studies of animal learning typically use easily discriminable stimuli for conditioning or study spatial memory for the location of a single platform in a water maze. However, in the real world, we process a continuous stream of sensorimotor experience and need to lay down multiple distinct memories. This leads to the problem of how to manage our growing repertoire of memories, when to create a new memory, and how to express and update existing memories such that old memories are not (necessarily) forgotten as new memories are acquired.

There have been recent developments that place the idea of context central to the management of such repertoires. Rooted in so-called mixture models that have long been pursued in machine learning (Jacobs et al. 1991, Singh 1991), a number of theories have argued that the brain organizes memories according to discrete internally constructed contexts (Wolpert & Kawato 1998; Wolpert et al. 1998; Haruno et al. 2001; Gershman et al. 2010, 2014; Collins & Koechlin 2012; Collins & Frank 2013; Heald et al. 2021). As such, context serves to tag or index memories for subsequent recall and updating. The effects of context on memory have been demonstrated in a diverse set of cognitive domains, including classical conditioning (Redish et al. 2007, Gershman et al. 2010), episodic memory (Howard & Kahana 2002, Zacks et al. 2007, Gershman et al. 2014), economic decision-making (Collins & Koechlin 2012, Collins & Frank 2013), spatial memory (Gulli et al. 2020, Julian & Doeller 2021, Plitt & Giocomo 2021), and motor learning (Heald et al. 2018, 2021; Oh & Schweighofer 2019). These results suggest that successfully managing a repertoire of

memories hinges on our ability to determine which contexts are appropriate at different points in time.

Although context is unambiguously defined in most laboratory experiments (at least from the experimenter's point of view), in the wild (and in general, from the animal's perspective), it is anything but (Heald et al. 2023). Thus, determining, or inferring, the currently appropriate context is a core computational challenge in context-dependent learning. We start by reviewing a theoretical framework for formalizing contextual inference and how it controls memory expression, updating, and creation. We then use this framework to organize our understanding of the neural bases of context-dependent learning. Finally, we review recent advances in training artificial neural networks (ANNs) on continual learning tasks, that is, tasks that require managing multiple, potentially interfering memories. We discuss how such models may inform the study of the neural bases of context-dependent learning and, conversely, how a better understanding of the way the brain solves contextual inference may provide useful clues for constructing more efficient artificial continual learning systems.

## 2. A COMPUTATIONAL FRAMEWORK FOR CONTEXT-DEPENDENT LEARNING

In order to study the neural bases of context-dependent learning, we introduce a normative approach that starts from a generative model, which summarizes the causal and statistical relationships between relevant quantities in the environment, such as contexts, sensory cues, and actions—including the ways they may change over time. Some of these quantities, such as sensory cues and actions, are directly observed by the brain, while others, such as contexts and reward contingencies, are hidden, as they cannot be observed directly. While there have been a number of different proposals for the precise form of the generative model (Wolpert & Kawato 1998; Wolpert et al. 1998; Haruno et al. 2001; Gershman et al. 2010, 2014; Collins & Koechlin 2012; Collins & Frank 2013; Sanders et al. 2020; Heald et al. 2021), here we focus on one particular, recent example, the contextual inference (COIN) model of motor learning (Heald et al. 2021), because it unifies a number of aspects of previous models, and it has been particularly successful in accounting for behavioral data. Although, in its original form, the COIN model was developed to understand the principles of motor adaptation, its foundational concepts are domain general.

Once the generative model is defined, learning amounts to the process of inferring the hidden variables of the environment based on the observed variables and, in particular, inferring those hidden variables that are assumed to change only slowly, or not at all, over time.[1] Below, we describe this computational framework in formal detail, but the uninitiated reader is welcome to skip forward to Section 3, where we begin by summarizing the main properties and predictions of this framework more informally.

### 2.1. Generative Model

The generative model of the environment that is relevant for contextual learning is a variant of the standard Markov decision process or, more generally, partially observable Markov decision process that forms the basis of reinforcement learning (Sutton & Barto 2018) (see also below) and its applications to neuroscience (Dayan & Abbott 2005). We illustrate the generative model

**Continual learning:** the ability to learn multiple tasks sequentially without forgetting old tasks

**Normative approach:** a normative approach is based on defining a goal (e.g., maximizing reward) and finding an optimal solution to it under constraints (e.g., available computation time, noise)

**Markov decision process:** a generative model in which the current state and action fully determine, albeit in general only probabilistically, the next state

**Partially observable Markov decision process:** a Markov decision process in which states are hidden and need to be inferred (e.g., from sensory cues, actions, and sensory feedback)

**Reinforcement learning:** the study of how artificial or biological agents learn from experience to adapt their behavior so as to improve their performance as measured by cumulative rewards

---

[1]Inferring hidden variables that vary on fast timescales (e.g., an animal's location in the environment, or the configuration of one's arm) is also a highly relevant and challenging task for the brain, but one that is usually not considered to belong to the realm of learning. In the following, to focus on learning, we ignore the inference of most fast-changing hidden variables.

(**Figure 1**) with respect to a mouse in a maze seeking food. In this framework, the (biological or artificial) agent has a state at time $t$, $s_t$ (the mouse's location in the maze), and can generate an action, $a_t$ (the mouse turning left or right). The mouse then receives sensory feedback, $r_t$ (the absence or presence of a food reward), and experiences a state transition (the mouse moves to a new location).
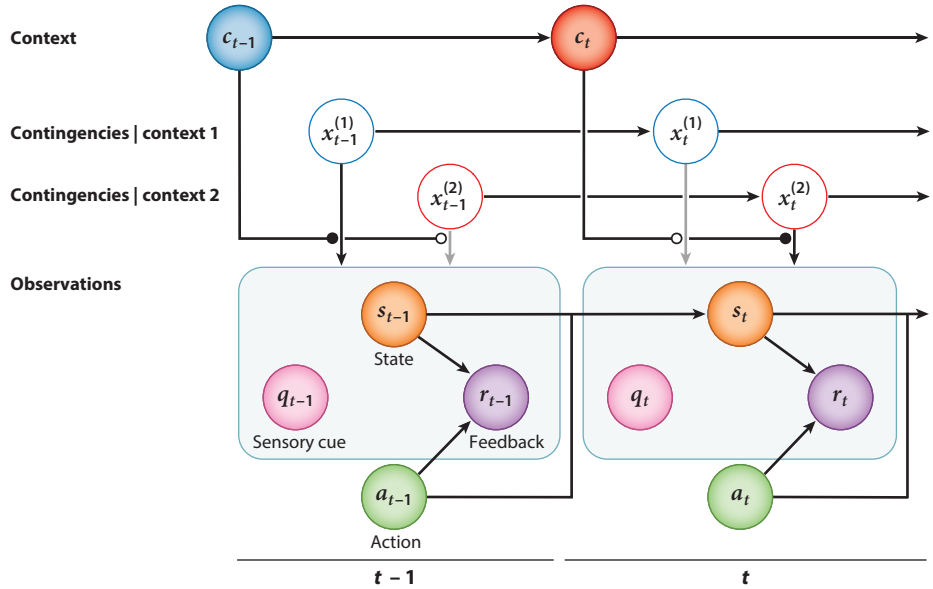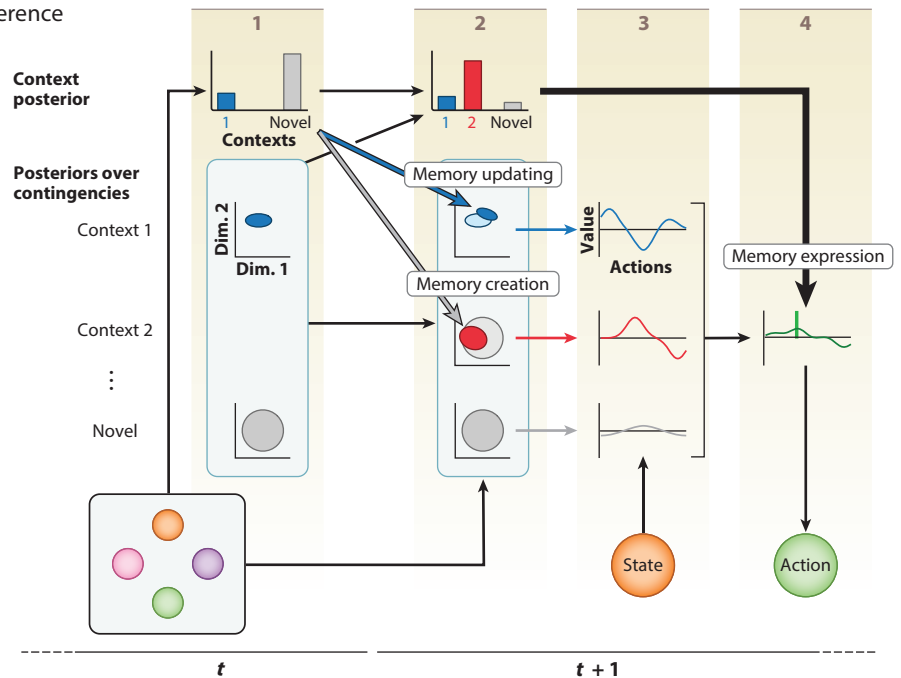


(*Caption appears on following page*)

**Figure 1** (*Figure appears on preceding page*)

A normative model of context-dependent learning. (*a*) Generative model, showing two consecutive time steps, $t - 1$ and $t$. Context, $c_t$, evolves over time (color represents the identity of the active context). Each context $j$ is associated with a set of contingencies, $x_t^{(j)}$, that may also evolve in time, independent of the contingencies of the other contexts. (Only two contexts are shown for simplicity; in general, the number of contexts and sets of contingencies can be unbounded.) The contingencies of the currently active context (*filled* versus *empty circles*, respectively, gating *black* versus *gray arrows*) determine what sensory cue is received, $q_t$ (*pink*), and how the agent's action, $a_t$ (*green*), changes its state, $s_t$ (*orange*), and leads to sensory feedback, $r_t$ (*purple*). In general, the contingencies of the active context can also affect the next context transition (not shown for clarity), and states may not be directly observable. Panel *a* adapted from Heald et al. (2023). (*b*) Inference, showing two consecutive time steps. The agent infers contexts (*top*, inferred posterior distributions shown as histograms) and context-specific contingencies [*middle*, inferred posterior distributions in a multidimensional space of contingencies illustrated with covariance ellipses, dimensions (dim.) 1 and 2 shown] based on observed states, sensory cues, sensory feedback, and its own actions (*bottom*, colors as in panel *a*). Contexts are color coded in top and middle as in panel *a*. Note that the probability that a yet-unseen, novel context with default contingencies is encountered is also always computed (*gray*). In columns 1 and 2, the context posterior at time $t$ (column 1, *top*) is computed by fusing prior expectations (propagated context posterior from time $t - 1$, not shown) with information about the current observations (column 1, *bottom*). This posterior determines the degree to which information from the current observations is used to update the inferred context-specific contingencies at time $t + 1$ (column 2, *middle*, memory updating, *pale blue* distribution updated to *bright blue* distribution) and whether a new memory is instantiated (column 2, *middle*, memory creation, the *red* posterior over contingencies updated from the *gray* posterior over contingencies of a novel context). In column 3, for each represented context, the expected values of actions (curves in context-specific colors) are computed based on the inferred contingencies of that context. In column 4, the final expected values of actions (*green curve*) are the weighted sum of their context-specific values, with the weighting determined by the posterior probabilities of the corresponding contexts (memory expression). The action with the highest final expected value is chosen (*vertical green line*; potentially with some decision or motor noise, not shown). Note that memory creation, expression and updating can all occur on the same time step, but they are shown here on separate time steps to avoid visual clutter.

Therefore, according to this generative model, (the distribution of) $r_t$ and the (distribution of the) next state, $s_{t+1}$, depend on the current state, $s_t$, together with the agent's action, $a_t$, in a manner defined by a set of primary contingencies (e.g., the current configuration of walls in the maze and the location of food inside it). In addition, there may exist sensory cues, $q_t$ (the general appearance of the room containing the maze), that are unrelated to states, actions, and feedback and also depend on the primary contingencies.[2] The primary contingencies may themselves change over time (some passageways become obstructed or new ones may open), controlled by some other (secondary) contingencies [e.g., the volatility of the environment (Piray & Daw 2020)].

Critically, central to context-dependent learning is the assumption that there may exist several such sets of (primary and secondary) contingencies in parallel, $x_t^{(j)}$ ($j = 1, 2, \ldots$), and the identity of the currently active one (i.e., the one that gets to control sensory cues, state transitions, and feedback) is determined by the current context, $c_t$ (e.g., the maze identity) (**Figure 1a**). Note that it is because there are multiple contexts that sensory cues become relevant: They can be informative as to the context (different mazes may be associated with different-looking rooms). However, as opposed to states, once the context is known, sensory cues are no longer informative as to the relation between actions and feedback.

Finally, context can also change over time (depending on context-specific transition probabilities, which can be regarded as further secondary contingencies), and when it does, it is thought to

**State:** the features of the environment (and agent) that, together with the agent's action, determine sensory feedback and the next state

**Sensory feedback:** sensory input that either has intrinsic value or is related to task performance (e.g., rewards, punishments, movement accuracy)

**Primary contingencies:** parameters that control the statistics of how state transitions, sensory feedback, and sensory cues depend on states and actions

---

[2]Although sensory inputs do not come to the brain already labeled as sensory feedback or sensory cues, studies in humans suggest that the brain spontaneously categorizes stimuli into these two classes (Collins & Frank 2016).

instantaneously switch which set of contingencies is active (**Figure 1***a*; gating connections from context nodes to the active versus nonactive context-specific contingencies, expressing their control, or lack of control, of observations). One important aspect by which previous and more recent models of contextual learning differ is the extent to which they allow context transition probabilities to be structured, that is, whether and how much transition probabilities depend on (or conversely, generalize across) "from" contexts, "to" contexts, both, or neither (for a review, see Heald et al. 2023; for a more general comparison beyond transition probabilities, see Heald et al. 2021, section 7 of the supplementary information and extended data table 1).

Note that, theoretically, the distinction between contexts and states need not be as clear cut as we depict it here. States (rather than just contexts) may also be hidden, sensory cues may also depend on states (rather than just on contexts), and transitions between contexts (rather than just those between states) may also depend on actions and states. In this more general framework, contexts may simply correspond to higher-level, more slowly changing components of a high-dimensional, hierarchical hidden state representation. For example, although in typical experiments contexts may be naturally operationalized by different experimental rooms and states as positions within a maze situated in a room, the mouse may actually integrate all these into a single multiresolution map of its environment. Nevertheless, it may be computationally advantageous to distinguish between high-level contexts [often referred to as tasks (Xie et al. 2021)] and low-level states. Moreover, there is both behavioral and neural evidence suggesting that (the internal representations of) contexts may have a special status (see below), and so for the purposes of this review, we maintain a qualitative distinction between contexts and states.

## 2.2. Contextual Learning as Recursive Inference

The ultimate goal of a reinforcement learning agent, and arguably also of the brain, is to choose actions, $a_t$, adaptively so as to maximize (cumulative) rewards (that are included in our notion of sensory feedback, $r_t$). Although there are many algorithmic strategies that can be used for this (Sutton & Barto 2018), the most flexible class of algorithms that makes the best use of limited experience with the environment (i.e., most data efficient) is based on a generative model of the environment (i.e., it is model based). Thus, the task for the brain in context-dependent learning can be formalized as the inference of the hidden variables of the generative model described in Section 2.1, based on the observed sensory inputs and motor outputs (actions). To reduce the computational complexity of learning, and especially of choosing actions (i.e., expressing memories, albeit at the cost of reduced data efficiency and behavioral flexibility), the brain may also combine model-based inference of context with model-free strategies for learning within individual contexts (Singh 1991). Although we consider the fully model-based setting below, our focus will be on the inference of context, which thus remains relevant even in such a hybrid learning system.

### 2.2.1. The targets of inference: environmental contingencies and context.
The key hidden variables to be inferred are the contingencies specific to each context that has been encountered so far, $x_t^{(j)}$, and the identity of the current context, $c_t$ (see the top half of **Figure 1***a*). In our example with the mouse foraging for food, it needs to infer the layout (and other features such as the volatility; see below) of each maze it has encountered (the contingencies) and the identity of the maze it is currently in (the context). The observations that can be used for this are the state, $s_t$; the sensory feedback, $r_t$; the sensory cue, $q_t$; and the agent's own action, $a_t$ (see the bottom half of **Figure 1***a*).[3] While treating the contingencies of the environment as hidden variables has been a

---

[3] In general, states are themselves hidden variables, and as such they need to be inferred by the brain, but in order to focus on learning we treat them as directly observable here (see also Section 2.1).

common feature of formal theories of learning, treating the current context as an additional hidden variable is a relatively recent innovation (Gershman et al. 2010, 2014; Collins & Koechlin 2012; Heald et al. 2021) that has far-reaching consequences, as we see below. This choice is motivated by a simple fact we already highlighted in Section 1: From an animal's perspective, context is only defined ambiguously in most situations, especially in ecologically relevant ones.

**2.2.2. The result of inference: the posterior distribution.**  Optimally, the result of inference should be a posterior probability distribution defining the joint probability of all possible combinations of the current values of all hidden variables given the history of all previous observations, $P(c_t, x_t^{(j=1,\dots)} | s_{0:t}, a_{0:t}, r_{0:t}, q_{0:t})$. This distribution can be computed recursively, incorporating the most recent observations into the posterior computed in the previous time step. The joint posterior over all hidden variables at any point in time can be conceptually separated into two components: the posterior distribution over the current context (see the top row with histograms in **Figure 1b**), and the posterior distributions over the contingencies of each possible context (see the middle three rows with covariance ellipses in **Figure 1b**). In particular, it is sufficient to represent posterior context probabilities and posteriors over contingencies for only those contexts inferred to have been encountered so far. Besides these, the novel context probability, that is, the total probability of all yet-unseen contexts, also always needs to be represented (see the top row in **Figure 1b**), together with an additional set of default contingencies for all such contexts (see the lowest covariance ellipses in the middle of **Figure 1b**).

Critically, all the agent's experience pertinent to a given context is captured by the corresponding posterior over contingencies. In other words, in this framework, the posterior over contingencies specific to a given context constitutes the memory of that context. Alternatively, to reduce computational and representational complexity, instead of inferring a full posterior distribution, most theories posit that only single estimates of contingencies are learned (but see Section 4.3). Yet other, altogether model-free strategies may not try to infer or estimate the contingencies of the environment at all and instead learn other kinds of representations (such as value functions or policies). Whatever the precise nature of the context-specific representations being learned, we henceforth refer to them as memories. Importantly, the principles of contextual learning we describe in the next section apply to nearly all kinds of memories, regardless of their representational content.

In contrast to inferring the values of hidden variables based only on past and present observations (**Figure 1b**), in some situations, it may be desirable to retrospectively refine inferences made in the past based on subsequent observations, such as when recalling an episodic memory (Gershman et al. 2014). In Bayesian inference, this type of retrospective inference is known as smoothing and is a computationally challenging operation that requires storing the entire history of observations in memory. How the brain might approximate the fully Bayesian solution to smoothing is unknown, though consolidation-like processes that replay previous experiences (Pfeiffer 2020) are likely involved.

# 3. COMPUTATIONAL PRINCIPLES FOR MANAGING MULTIPLE MEMORIES

The unique challenge of contextual learning is that it requires the agent to manage multiple memories, each corresponding to a different context. In particular, it needs to ensure that at all times it expresses and updates existing memories and keeps creating new memories as appropriate. The normative theory developed in Section 2 reveals the computational principles governing how these processes should optimally be organized. In particular, it implies that memory expression, updating, and creation need to be graded, or probabilistic, rather than all-or-none and deterministic, as

in standard conceptualizations of these processes. It also specifies how the current context needs to be continually inferred and how this contextual inference needs to control each of these graded processes. In this section, we describe these computational principles in more detail. These principles provide a set of important constraints on the putative neural mechanisms of contextual learning, reviewed in Section 4. To aid intuition about the computational principles underlying the managing of multiple memories, we continue to use the mouse in a maze example. We also briefly refer to relevant work demonstrating these effects behaviorally in specific experimental paradigms (for a recent more comprehensive review on behavioral effects, see Heald et al. 2023).

## 3.1. Contextual Inference

In order for memory expression, creation, and updating to be optimal, these operations should be performed in a graded or probabilistic manner. This in turn requires the full distribution over contexts, expressing the probability that each known context or a yet-unknown novel context is currently active, and hence context uncertainty, to be represented at each point in time. This contrasts with alternative approaches that ignore context uncertainty by only estimating the most probable context. The distribution over contexts should not be independent at each point in time (a priori) but should take into account the previous context distribution and the context transition probabilities (or more generally, the dependence of the current context on the history of past contexts).

The generative model described above (Section 2.1) is, perhaps, the simplest for contextual inference as it assumes that exactly one discrete context (a hidden cause) is active at each point in time. There are three major extensions to such a model in terms of compositional, continuous, and hierarchical representations of contexts, which can each contribute to more powerful generalization. In compositional models, multiple hidden causes can be active at each point in time, and hence a context is formalized as a unique combination of hidden causes. For example, when navigating a maze to find a reward, it can be beneficial to represent the reward function (goal location) and state transition function (mapping from states and actions to the next states) as two separate hidden causes so that they can be combined in novel ways to solve new tasks (Franklin & Frank 2020). Continuous representations of context allow the contingencies associated with different contexts to vary along a continuum. For example, when manipulating objects, it may be beneficial to consider different objects within a class (e.g., different cups) to lie on a manifold of contexts, defined by continuously varying contingencies [such as their weight and visual size (Braun et al. 2009)]. Finally, contexts may also be hierarchically organized, for example, because different mazes in the same room may have different configurations but still share many of their sensory cues, while mazes in different rooms will even differ in their sensory cues. In the following, we focus on how the posterior distribution over contexts controls memory expression, memory updating, and memory creation.

## 3.2. Memory Expression

The ultimate goal of having memories is to express them in behavior. The simplest form of how contextual inference can control this is when the current context is known to the learner with absolute certainty, that is, when the posterior distribution over contexts is all or none. In this case, behavior is driven by a single memory: the one corresponding to the current context. The normative strategy for choosing an action in a given context is to compute the value of every available action in terms of the cumulative reward the agent expects to receive in the future after choosing that action now—depending, of course, on its current state (see the arrow from state at the bottom to context-specific value functions in the third column of **Figure 1b**) and its

context-specific memories (see the arrows from the posteriors over contingencies to the value functions in the middle of **Figure 1*b***). Then, the action with the highest value can be chosen.

As there is often uncertainty about the current context, in principle, the memories for all contexts need to be mixed (Haruno et al. 2001, Gershman et al. 2017, Heald et al. 2021). In this case, once the context-specific action values have been computed (see above), according to the rules of Bayesian decision theory, the final expected value of each action is computed by averaging its value across contexts,[4] weighted by the posterior context probabilities, and the action with the highest expected value can be chosen (see the arrow pointing from the context posterior of the second column to the action-value function in the fourth column of **Figure 1*b***). Thus, a key feature of the model is that contextual inference controls the expression of memories in the actions the agent chooses. In particular, the contents of several memories (those associated with contexts that have a nonzero posterior probability) need to be mixed for optimal behavior. This is notably different from classical accounts of memory recall that focus on how a single memory is retrieved at a time (Hopfield 1982).

The graded mixing of memories provides an alternative account to even such a basic process as memory extinction. Classical (Rescorla & Wagner 1972) and even more modern (Smith et al. 2006) accounts treat extinction as the overwriting (or erasure) of some previously established memory trace and, as such, as being related to memory updating (see also below). In contrast, when viewed through the lens of contextual learning, extinction may instead be due to the gradual reexpression of the memory for the original context preceding the beginning of acquisition (Gershman et al. 2010), as the newly acquired memory (associated with the acquisition context) has no reason to be expressed. This naturally explains ubiquitous postextinction phenomena that are widely regarded as difficult to reconcile with a simple overwriting-based account, such as the slow spontaneous recovery of the extinguished memory with the passage of time (Rescorla 2004), its fast evoked recovery in response to context-specific sensory cues [as in so-called ABA renewal (Bouton & King 1983)] or sensory feedback [reinstatement (Rescorla & Heth 1975) or evoked recovery (Heald et al. 2021)], or its rapid reacquisition [compared to the speed of its original acquisition (Napier et al. 1992)].

## 3.3. Memory Updating

On each time step, the context posterior and the memories are updated by fusing information from the past with information arriving in the current time step (see the transition from the first to second column, respectively corresponding to time step $t$ and $t + 1$, in **Figure 1*b***). For example, imagine that a passageway that the mouse found to be open on previous trials now appears to be closed. In principle, this experience in the current time step will require updating its existing memories, representing information from the past (that the passageway used to be open). The extent of this updating fundamentally depends on the balance between the reliability (or relevance) of present versus past information. In principle, the higher the reliability of current information relative to that of past information, the more memories should be updated in light of this new experience.

The key question then becomes, What determines the relative reliabilities of these two sources of information? The reliability of existing memories (i.e., the relevance of past information they store for the present) depends on the context-specific secondary contingencies describing how

---

[4]In general, in each context, the overall value of each action needs to incorporate its contributions to both exploitation, by steering the organism toward more rewarding states, and exploration, by resolving uncertainty about context (and other hidden variables), thus leading to higher rewards in the long run. Moreover, if actions can also affect context transitions (Section 2.1), this will also need to be incorporated into action values.

(and how much) the primary contingencies are expected to change over time. If, based on previous experience, the mouse has inferred that the maze configuration in a room tends to remain the same for long periods of time, the new information about the passageway being closed may be chalked up to happenchance and as such largely ignored. However, if previous experience tells the mouse that the maze in a room is often reconfigured by some malevolent power (i.e., the experimenter), then this same information will be used to update the mouse's cognitive map of the maze. This predicts higher effective learning rates in more volatile environments—an effect that has been confirmed experimentally (Burge et al. 2008). Note that, however, this prediction changes when the volatility of the environment is at the level of contexts rather than the level of contingencies (i.e., between- rather than within-context volatility), for example, if the mouse is often transported to an altogether different maze. In this case, rather than changes in the rate of adaptation being driven by memory updating (proper learning), they may be due to changes in memory expression [apparent learning, so called because from behavior alone it may appear as though proper learning has taken place (Heald et al. 2021, 2023)].

The reliability/relevance of currently incoming information is determined by two factors. First, the better the quality of evidence it supplies, the more relevant it is for updating memories. For example, if the mouse only saw the passageway to be closed from a distance, this observation may be attributed to poor visibility, and as such ignored again. If, however, the mouse went all the way to the end of the passageway to confirm it was closed, this strong evidence is worth registering in memory. Second, specific to contextual learning, a given memory should only be updated inasmuch as the current information comes from the same context as that associated with that memory. More precisely, contextual inference should gate and modulate the updating of memories, such that the extent to which memories are updated is proportional to the posterior probabilities of the contexts to which they belong. Again, in contrast to classical notions of memory updating (e.g., via reconsolidation), this implies that when there is uncertainty about the current context (so that several context probabilities are nonzero), multiple memories need to be updated in parallel. That is, if the mouse is unsure which of two mazes it is currently navigating, the information about the closed passageway needs to be incorporated into both memories. Consistent with this idea, studies of single-trial motor learning have shown that under experimentally controlled contextual uncertainty, multiple memories are updated in proportion to their posterior probabilities (Heald et al. 2021).

### 3.4. Memory Creation

The notion of memory creation is naturally captured by Bayesian nonparametric models such as the so-called Chinese restaurant process [and its hierarchical extension, the Chinese restaurant franchise (Teh et al. 2006)]. These models assume that the number of contexts in the environment are unknown and unbounded (i.e., can be infinite), such that when a new context is inferred, a new memory is created. These models have been used to account for a wide range of context-dependent phenomena in a variety of domains, including classical conditioning (Gershman et al. 2010), episodic memory (Gershman et al. 2014), economic decision-making (Collins & Koechlin 2012), and motor learning (Heald et al. 2021; for a more comprehensive overview of Bayesian nonparametric models in context-dependent learning, see Heald et al. 2023).

According to these models, in each time step, with some probability, a hitherto unseen context is inferred to have been encountered, and a new memory for this context is added to the repertoire of existing memories (see **Figure 1b**). The probability for this is given by the probability of a novel context in the previous time step (see the novel context probability in the first column in **Figure 1b**). The memory associated with this newly encountered context is an updated version of the default posterior over contingencies of a novel, yet-unseen context (see the creation of the

posterior and its updating in the second column of **Figure 1*b***). As all contexts are assumed to have existed indefinitely before they are first encountered, the contingencies for a novel context have a stationary state distribution that is never updated (see the covariance ellipse, which remains unchanged, in the second column of **Figure 1*b***). For the mouse, this means that if no familiar mazes are consistent with a set of observations (so that the probabilities of existing contexts are all low, and consequently—as probabilities add to one—the probability of the novel context is high), there is probably a new maze in place, and it is time to create a new memory for it. Thus, contextual inference also controls memory creation.

Memory creation in the model is unlikely when small changes occur in the environment, as errors will be small and current memories will be able to explain sensory cues and feedback. In contrast, abrupt changes in the environment increase the probability of a new memory being created. Consistent with this, deadaptation in motor control is slower after the removal of a gradually, versus abruptly, introduced perturbation. The gradual introduction leads to the adaptation of an existing memory so that, unlike the abrupt case, deadaptation cannot be quickly achieved by switching the context back to the original memory. In contrast, the abrupt perturbation leads to a new memory being created, allowing preservation of the original memory for the unperturbed situation and subsequently rapid switching back to the unperturbed memory [a rapid process (Heald et al. 2021)]. Analogous phenomena have also been described in a variety of other domains of learning (Heald et al. 2023).

# 4. NEURAL BASES OF CONTEXTUAL LEARNING

The computational principles of contextual learning reviewed above (Section 3) provide a useful framework for organizing the large body of data on its neural bases. At present, there is no unified circuit-level understanding of contextual inference and learning as described above. In this section, we highlight studies implicating specific circuits and mechanisms underlying different aspects of contextual inference and learning. As contextual inference is critical for controlling all aspects of contextual learning, we begin by reviewing experimental data about its neural underpinning. We then review data relevant for each of the main operations that contextual inference controls—memory expression, updating, and creation—focusing on their aspects that are relevant when multiple context-dependent memories need to be managed.

## 4.1. Contextual Inference

Although many brain areas have been implicated in some aspect of context-dependent learning, here we focus on three areas that have been most extensively studied in paradigms tapping into contextual inference: the prefrontal cortex (PFC), hippocampus, and thalamus.

### 4.1.1. Prefrontal cortex.
The PFC is thought to play a key role in representing contextual information (Botvinick 2008, Donoso et al. 2014, Soltani & Koechlin 2022). In our terminology, PFC may be involved in both computing context probabilities and updating them based on context transitions and sensory cues. This is broadly consistent with multiple areas within PFC encoding abstract task rules (Wallis et al. 2001, Saez et al. 2015) and medial PFC being necessary for switching between different rules based on sensory cues (Marton et al. 2018).

Implementation-level models of PFC corroborate its role in contextual inference. Collins & Frank (2013) developed a neurobiologically explicit network model consisting of multiple modules corresponding to distinct areas within the corticobasal ganglia loop. When the network was trained on an instrumental learning task with multiple hidden contexts (by pure reinforcement learning), the module corresponding to PFC developed activations consistent with contextual

inference. This network model also provided an approximate implementation of a Bayesian non-parametric model of contextual learning, similar to the one described in Section 2, which itself explained a wealth of behavioral data from humans performing the same task.

If PFC is involved in computing context probabilities, then lesioning it, or interfering with its function, should impair behaviors that rely on the maintenance and updating of context probabilities. This is broadly consistent with a large body of neuropsychological literature describing how patients with PFC lesions have difficulties maintaining behavioral flexibility between contexts (Shallice & Cipolotti 2018). Going beyond simple behavioral impairments, perturbing (by transcranial magnetic stimulation) the dorsolateral PFC led to the directed forgetting of episodic memories—a process that is thought to be based on the suppression or substitution of the representation of the context relevant to those memories (Anderson & Hulbert 2021).

In animal learning experiments with rats, lesioning the medial orbitofrontal cortex (OFC), a major subdivision of PFC, resulted in a pattern of decision-making impairments (diminished Pavlovian-instrumental transfer and outcome devaluation) that was mostly well accounted for by the putative role of OFC in contextual inference [except for a lack of effects on reinstatement (Bradfield et al. 2015)]. In addition, as we saw above, extinction and subsequent savings on relearning are behavioral effects that bear the signatures of contextual inference. In particular, they reflect the increasing expression of the memory of the pre- or postacquisition context, respectively, driven by the growing probability that each of these contexts is active. Indeed, lesions of OFC slow down extinction (or the relearning of a previous set of contingencies) in classical conditioning (Butter 1969). This may be because, rather than reexpressing a previous memory (by apparent learning, a potentially rapid process), OFC-lesioned animals need to update their current memory (by proper learning, a relatively slow process) (Wilson et al. 2014).

Spontaneous recovery has been suggested to be another behavioral signature of memory expression driven by contextual inference (Heald et al. 2021) (see also Section 3.2). Here, introducing a working memory task, which is known to tax PFC resources, led to a boosting of spontaneous recovery (Keisler & Shadmehr 2010). This seemingly paradoxical effect (interference makes a behavioral effect greater rather than smaller) has been explained by a model in which the working memory task specifically affects the active maintenance of current context probabilities (Heald et al. 2021). According to this explanation, when context probabilities are not actively maintained and updated, they revert to a set of default values that are stored in long-term memory (instead of working memory) and that favor the acquisition context, as it has been experienced for a long time.

Functional imaging studies have shown that OFC represents the abstract state of the task (see, e.g., Wilson et al. 2014), especially when it is hidden and must be inferred (Schuck et al. 2016, Nassar et al. 2019, Schuck & Niv 2019). One prominent demonstration of this effect used a paradigm in which participants had to judge the age (in our terminology, the state; old or young) of one of two simultaneously displayed categories (house or face) on each trial (Schuck et al. 2016, Schuck & Niv 2019). The relevant category (in our terminology, the context) was not indicated directly to participants but depended on the relevant category of the previous trial, as well as on the age on the previous two trials. To fully account for these dependencies, the authors defined the task state as a conjunction of the current and previous category and age. Interestingly, three out of these four variables (previous and current category and previous age) were represented in OFC, but not the current age. The authors argued that this was because the current age was directly observable, whereas the other variables were hidden. However, Bayesian models of perception suggest that stimulus features such as age are also hidden and must be inferred. Instead, we suggest that the previous age and previous category (together with the age from two trials ago, which was untested) are sufficient statistics of the posterior distribution over the current context,

while the current age is not. Thus, if OFC is involved in contextual inference, we would expect the combination of these variables to be decodable in OFC—just as has been found (Schuck et al. 2016). This interpretation makes the prediction that the age from two trials ago should also be decodable, which remains to be tested.

More direct evidence for the role of PFC in contextual inference comes from studies that combine functional imaging with Bayesian multiple-context models. Such studies suggest that PFC does not represent a single context at a time (e.g., the most likely active one) but represents multiple contexts and the probability that each of them is currently active. For example, in an economic decision-making task in which state-action-feedback contingencies could switch based on a hidden context, anterior regions of PFC, and ventromedial PFC in particular, were shown to encode the current reliability of different context-specific behavioral strategies (Donoso et al. 2014) (**Figure 2a**). These reliability signals are formally analogous to the posterior context probabilities (often termed responsibilities) of Bayesian multiple-context models (Heald et al. 2021).

The aforementioned studies demonstrated the representation of inferred contexts in PFC even in the absence of explicit information about the existence of different contexts and when contextual inference needed to combine information about context transitions as well as action-outcome contingencies. However, they did not specifically test whether full posterior distributions are represented over contexts, as would be required theoretically. More direct evidence for this comes from an experiment in which participants had to infer the current context (which sector of a safari park they are in) after viewing a sample of animals from that context. OFC responses were best explained as representing the (log-transformed) posterior distribution over contexts given those stimuli, outperforming alternative representations, such as the most probable context or the uncertainty over contexts (Chan et al. 2016). However, the paradigm used in this study was specifically designed so that full-context posteriors needed to be represented to perform well while—compared to some of the previous studies—only relatively impoverished context-specific information was required to be learned (the distribution of animals in each context, but no actions or structured context transitions). Therefore, it remains an open question whether these results generalize to richer, more challenging contextual inference settings.

Direct neural recordings from monkey OFC (and other prefrontal cortical areas) confirmed a neural representation of context in the PFC, even on trials when context was hidden and as such had to be inferred (Saez et al. 2015, Bernardi et al. 2020). Nevertheless, similar context-encoding representations were also found in other brain areas, including the hippocampus and amygdala. These studies indicate that context is encoded in an abstract way in prefrontal (and hippocampal) areas, orthogonal to other task-relevant variables [e.g., stimulus identity (Bernardi et al. 2020)]. Such abstraction allows flexible generalization to novel combinations of context and other variables. Interestingly, when assessed using a simple condition-specific linear decoder, only activities in the amygdala, but not those in OFC or other cortical areas, were predictive of behavioral errors (Saez et al. 2015). However, the level of abstraction of neural representations, assessed using a decoder that was required to generalize across conditions, was predictive of behavioral errors in prefrontal (and other cortical) areas (Bernardi et al. 2020). More importantly, although the tasks used in these studies often required the animals to infer context from past observations, without providing unambiguous sensory cues on a given trial, only the representation of the experimentally defined veridical context was tested. Thus, whether and how probabilistic inferences are reflected in neural activities in these tasks remains unknown.

At the level of neural circuits, the evidence for the causal role of PFC in contextual inference is somewhat mixed, at least in the relatively simple behavioral paradigms used so far. Optogenetic suppression of PFC (anterior cingulate) projections to primary visual cortex in an odd-ball paradigm (in which deviant stimuli ostensibly lead to the inference of a novel context) has shown
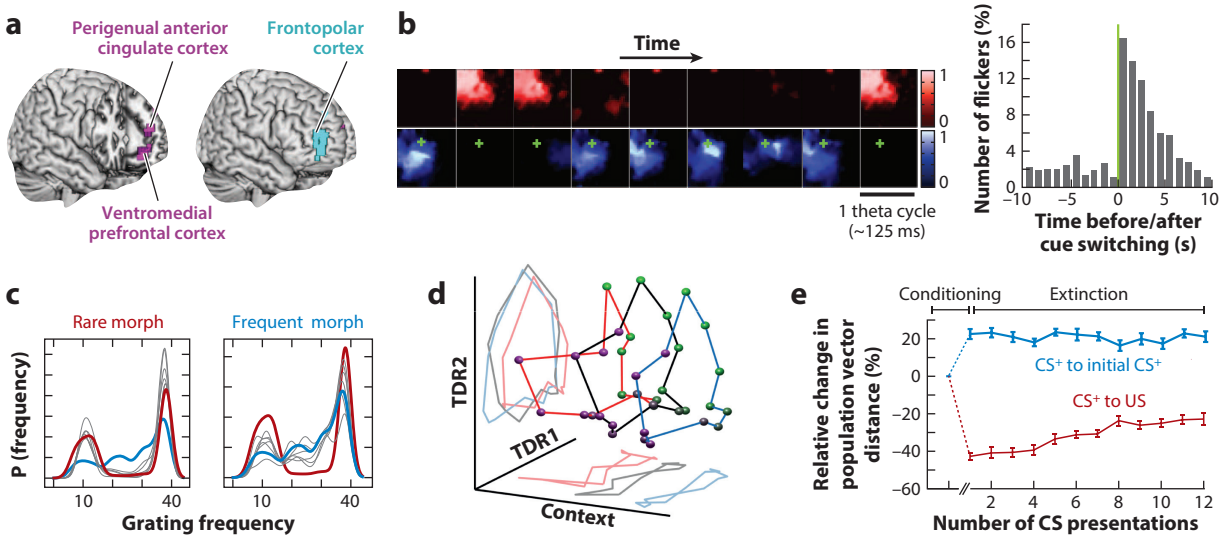
**Figure 2**

Neural bases of contextual learning. (*a*) Prefrontal cortex encodes context probabilities. Participants learned strategies that involved choosing an action (correct button press) in response to a state observation (number displayed on a screen). The state-action contingencies occasionally switched between three possibilities, necessitating different context-specific strategies. Brain areas in which (functional MRI) activations correlated with the probability (strategy reliability, derived from a behavioral model) of the context associated with either the participants' current strategy (*magenta*) or with alternative strategies (*cyan*). Panel *a* adapted with permission from Donoso et al. (2014). (*b*) Hippocampal activity samples from context-specific maps under context uncertainty. Hippocampal activity was recorded as rats navigated one of two spatial environments. Occasionally, animals would be teleported between the environments by instantaneously switching sensory cues. (*Left*) Time series (columns show successive theta cycles) of correlations between current neural population activity and reference activity associated with each location in environment A (*top*, *red*) and environment B (*bottom*, *blue*). A teleportation from environment A to B took place shortly before this series. The correlation shows evidence of spontaneous flickering between the hippocampal place-field maps associated with each environment despite no change in the environment during this period (*green cross* shows current location). (*Right*) Histogram shows a temporary increase in the number of flickers after a change in the environment. Panel *b* adapted with permission from Jezek et al. (2011). (*c*) Hippocampal representation of a prior distribution over contexts. Hippocampal activity was recorded as mice navigated through virtual reality environments that could morph between two extremes by continuously varying the frequency of a sinusoidal grating on the walls, thus giving rise to a continuum of contexts. Mice experienced one of two different distributions of contexts, characterized by different distributions of the morphs (rare versus frequent morph distributions shown as *red* versus *blue thick lines*, respectively). Thin lines show the prior reconstructed from the hippocampal activity of individual mice exposed to the rare (*left*) and frequent (*right*) distributions. The prior approximated the true distribution for each condition. Panel *c* adapted with permission from Plitt & Giocomo (2021). (*d*) Context-dependent motor learning. Nonhuman primates reached to targets under two possible force perturbations applied to the hand [forces applied by a robotic interface that act in clockwise (CW) or counterclockwise directions (CCW)] or unperturbed. Neural state space shows preparatory activity (after dimensionality reduction) in dorsal premotor and primary motor cortex for movements to the same 12 targets (*points*) under the three possible contexts: before learning (*black*), CW perturbation (*red*), and CCW perturbations (*blue*). Projections of the neural state space are shown on the floor and wall. Panel *d* plotted from data in Sun et al. (2022). (*e*) Neural representation of extinction in the mouse amygdala during fear conditioning. Change in a distance measure between the neural representation of the conditioned stimulus (CS) during training (conditioning and extinction) and the prelearning representation of the unconditioned stimulus (US) (*red*) or CS (*blue*). Negative values mean the representations become more similar. Panel *e* adapted with permission from Grewe et al. (2017).

that contextually appropriate responses (larger responses to deviants) depend on an intact PFC input (Hamm et al. 2021). Another study used an olfactory delayed match-to-sample paradigm (Wu et al. 2020) that can also be seen as engaging contextual inference, such that the sample stimulus acts as a sensory cue indicating the context, which in turn determines the mapping from the test stimulus to the correct response. In this study, although context could be decoded from OFC, other relevant cortical areas also carried such signals, including a primary sensory cortical area (the

piriform cortex) and a premotor area [the anterolateral motor cortex (ALM)]. More importantly, optogenetic inactivation of OFC did not impair context-dependent behavior in this case (effects on neural responses were not studied), while that of ALM did.

### 4.1.2. Hippocampus.
Information stored in the hippocampus has long been considered to serve as indices to patterns of neocortical activity produced by particular episodes (Káli & Dayan 2004, Teyler & Rudy 2007). This is broadly consistent with a role in contextual inference, inasmuch as context probabilities are essentially indexing which memories need to be expressed and updated.

More specifically, several lines of evidence point to the hippocampus as a key brain region that represents information about context and, in particular, spatially defined contexts (which arguably have special ecological relevance for most animals). Indeed, the hippocampus has been implicated in maintaining a cognitive map of the environment (O'Keefe & Nadel 1978). Moreover, like the OFC (see Section 4.1.1), it has been suggested to encode cognitive maps in the broad sense, capturing the relationships between sensory cues, states, actions, and feedback (Wikenheiser & Schoenbaum 2016, Behrens et al. 2018)—much like the relationships governed by the context variable in our model (**Figure 1***a*).

Classical work on fear conditioning demonstrated that pairing a foot shock with an auditory tone in a particular chamber (i.e., spatial context) leads to occasional fear responses (behavioral freezing) in the same chamber even in the absence of the tone and, conversely, to the conditioned auditory tone even when in a different chamber. Critically, lesioning the hippocampus selectively dampens fear to the context (but not to the tone) previously associated with the shock, albeit this role seems to be time limited (Kim & Fanselow 1992). In some cases, inactivating the hippocampus may also dampen fear renewal when the tone is presented outside of the extinction context (Holt & Maren 1999).

The neural representation of spatial context is widely thought to rely on the activity of place cells. As place cells have place fields (the region of physical space that needs to be occupied by the animal in order for a given cell to become active), their joint population activity provides a neural representation of the animal's location (see, e.g., the heat map in the leftmost bottom frame of **Figure 2***b*). Critically, due to remapping (a change in relative place field locations with a change in environment), the activity of place cells also encodes the environment, or the spatial context (see the leftmost top versus bottom frame of the heat map in **Figure 2***b*). While remapping is most often induced by placing animals in a different environment, and thus changing the combination of sensory cues available, remapping has also been observed when the animal's behavioral policy needed to be different in the same environment (Markus et al. 1995). This is in line with context determining not only sensory cues but also the mapping of states and actions to feedback and thus, ultimately, the policy that is best for acquiring rewards (**Figure 1***a*). Indeed, a formal mathematical model of contextual inference very similar to the one we propose here (**Figure 1***b*) (but without a concept of context transitions) has been able to account for a number of different (and sometimes puzzling) features of remapping, such as its dependence on sensory cues (including their variability), training and testing protocol, and its variability across animals (Sanders et al. 2020).

Contextual inference requires that, at any time, the brain represents a probability distribution over contexts, rather than just a single (e.g., the most probable) context (see Section 2.2). One proposal for how hippocampal responses might encode probability distributions is based on probabilistic sampling (Savin et al. 2014). This predicts that context representations in the hippocampus should dynamically alternate over time even under constant experimental conditions. Evidence for this proposal comes from experiments showing that the context encoded by hippocampal activity can flicker between representing two different spatial environments, such

**Probabilistic sampling:** instantaneous neural activities encode a single value of a hidden variable, such that over time they sample different values with frequencies proportional to the corresponding posterior probabilities

that in each cycle of the hippocampal theta oscillation (lasting for ∼125 ms) a single context is represented coherently across the population, but the represented context occasionally changes across subsequent theta cycles spontaneously, without changes in sensory cues (Jezek et al. 2011) (see the left side of **Figure 2b**). Moreover, also in line with sampling-based inference, the frequency of flickers was shown to increase following a switch in sensory cues (see the right side of **Figure 2b**), which—under a suitable model of context transitions—are ostensibly periods of time with increased contextual uncertainty.

Contextual inference, as a form of Bayesian inference, involves the combination of current sensory information with prior expectations to make inferences about the current context. In particular, prior expectations should reflect the probability distribution over contexts experienced during training. In accordance with this, recent recordings of large neural populations showed that hippocampal ensembles encode context in an experience-dependent manner (Plitt & Giocomo 2021). More specifically, the particular prior distribution of contexts (defined over a spectrum of continuously morphed environments) used for training could be decoded from hippocampal activity (**Figure 2c**).

The hippocampus may also contribute to contextual inference more indirectly via its effects on OFC. During periods of rest, a functional MRI study found that the human hippocampus replayed sequences of previously experienced task states (Schuck & Niv 2019) that embody our notion of context (see Section 4.1.1). Moreover, the more faithfully the statistics of these sequences reflected the transition structure of task states, as defined by the task, the more accurately task states could be decoded from OFC during task performance. In turn, decoding accuracy in OFC was positively correlated with task performance. As long as there is a causal relationship underlying these correlations, these results suggest that hippocampal replay may play a role in constructing the task-relevant representations in OFC that underlie contextual inference.

**4.1.3. Thalamus.** The PFC has reciprocal connections with thalamus, which in turn receives strong inputs from the basal ganglia, a brain structure that has been implicated in context-dependent action selection (Hikosaka et al. 2000, Collins & Frank 2013, Klaus et al. 2019). By integrating inputs from many different cortical areas, thalamic circuits are well situated to compute contextual signals that can rapidly reconfigure cortical representations. Indeed, it has been suggested that the thalamus acts as an optimal Bayesian observer of context, computing likelihoods that update posterior context probabilities in cortex (Rikhye et al. 2018). Consistent with this view, inhibition of mediodorsal thalamus in mice leads to perseveration of actions appropriate for a previous context following a change in context, such as a reversal of action-outcome contingencies (Parnaudeau et al. 2013, 2015).

## 4.2. Memory Expression

An emerging view of how memories are recalled by networks of recurrently connected neurons is based on a dynamical systems perspective. This perspective was anticipated by classical theories (Hopfield 1982), and it has recently gained experimental support, particularly in the field of motor control, that is, in the context of the expression of motor memories (Vyas et al. 2020). According to this perspective, the evolution of the state of a neural population is best understood in terms of the dynamical rules that govern how the next neural state is caused by the current neural state and inputs. In general, inputs can be either transient, such as those that set the initial (preparatory) state of motor cortex before movement offset, or tonic, persisting while the neural trajectory unfolds, or even longer, across many consecutive movements. Transient inputs are ideally suited to convey information about the current environmental state (e.g., the target of reaching) for which an action needs to be chosen, while tonic inputs are ideally suited to convey information about the current

context. Indeed, population dynamics in motor cortex during reaching have been well described by a mostly autonomous (i.e., not input-driven) dynamical system, in which different reaching directions correspond to different preparatory neural states to which the system is driven by appropriately tuned transient inputs (Kao et al. 2021b). Thus, the memory of a context, determining the mapping of environmental states to actions, is embedded in the dynamics of the circuit that determine how a neural trajectory evolves from an initial state to generate a sequence of motor commands to reach a target.

Due to the nonlinearity of neural circuit dynamics, different subspaces of the neural state space can support different local dynamics. Thus, these subspaces can form the dynamical substrates of independent modules capable of implementing context-specific memories (Mante et al. 2013, Remington et al. 2018a, Wang et al. 2018, Flesch et al. 2022). A particular context-dependent module can then be addressed by driving the neural state to the appropriate subspace by tonic (or slowly changing) context-dependent inputs (Mante et al. 2013, Remington et al. 2018a, Wang et al. 2018), providing a mechanism by which context-specific memories can be differentially expressed. Furthermore, the continuous nature of the state space means that context-specific memories can be interpolated between, for example, by driving the neural state to a preparatory state in between those for different contexts, allowing graded context probabilities to influence memory expression under context uncertainty. Alternatively, rather than selecting between different subspaces, context-dependent inputs might change neural gains (Stroud et al. 2018) or engage different thalamo-cortical loops via the basal ganglia (Logiaco et al. 2021), thus controlling the effective connectivity of the network and the resulting neural dynamics even within a single subspace. This mechanism might be particularly suitable for interpolating between different memories.

An elegant demonstration of these principles comes from recent studies of contextual motor learning in nonhuman primates. In these experiments, macaques performed different reaching movements while one of two opposing external force fields (corresponding to two different contexts) was applied to their arm. Recordings from motor cortex revealed that the subspace spanned by different preparatory neural states was context dependent (Sun et al. 2022) (in **Figure 2d**, the same colored points on each curve correspond to the same 12 reach directions, and the curve colors to movement under different force fields). For example, learning to reach through two opposing force fields resulted in two parallel subspaces that were each parallel to the subspace spanned by the before-learning preparatory states and were separated along a dimension that was orthogonal to this subspace (context dimension). The common dimensions shared by these context-dependent subspaces may provide a mechanism by which learned behaviors can generalize across contexts.

Similar principles have been shown to underlie flexible motor timing (Remington et al. 2018a). In a time-interval reproduction task (Remington et al. 2018b), a random time interval between two visual flashes was presented that monkeys needed to reproduce up to some context-dependent multiplicative factor (1 or 1.5), which changed over blocks and was indicated by color cues. Recordings from the dorsomedial frontal cortex revealed neural trajectories that were separated according to initial condition (determined by the presented time interval) and an orthogonal context-dependent axis, with trajectories in different parts of the state space evolving at different speeds to reproduce different timing intervals. Analogous results were also found in a variant of the same task, in which contexts were associated with different prior distributions over the possible time intervals that needed to be reproduced rather than different multiplicative factors with which they needed to be reproduced (Sohn et al. 2019). Using a context-dependent cross-modal decision task, a recent study found a similar representational geometry of context even in visual cortex (Hajnal et al. 2021).

## 4.3. Memory Updating

Within each context, memory updating should take into account uncertainty about the contingencies being inferred. In line with this, recent evidence suggests that synapses adjust their learning rates based on uncertainty, framing synaptic plasticity as Bayesian inference (Aitchison et al. 2021). Across contexts, behavioral studies have shown that memory updating is modulated by the posterior of contextual inference (Heald et al. 2021). A neural mechanism for such modulation, at the level of synapses, would require an additional component to simple two-factor Hebbian mechanisms. One class of mechanisms that has been proposed is three-factor mechanisms in which synaptic modulation depends on presynaptic and postsynaptic activity as well as the influence of neuromodulators that can convey scalar information about novelty or reward (Frémaux & Gerstner 2015). For example, tonically active cholinergic neurons of the striatum have been suggested to gate the learning of striatum-dependent behaviors, providing a mechanism by which memories can be protected from unlearning during extinction (Crossley et al. 2013). However, conveying information about context probabilities requires more than just representing a single scalar. How neuromodulatory influences could achieve this remains an open question.

## 4.4. Memory Creation

Memory creation involves two processes: novelty detection, that is, determining whether sensory experience is explained by current representations or is novel, and memory instantiation. Novelty has been proposed to be encoded in neuromodulatory signals. Brain areas such as the ventral tegmental area (Lisman & Grace 2005, McNamara et al. 2014) and locus coeruleus (Takeuchi et al. 2016) signal novelty by releasing dopamine into the hippocampus. In this way, newly encoded hippocampal representations, which are initially labile, are stabilized (McNamara et al. 2014, Takeuchi et al. 2016), providing a mechanism by which new long-term memories can be created. Contextual novelty signals also arrive in the hippocampus from the hypothalamus (Chen et al. 2020) and could, in principle, be used to trigger the creation of a new map of space (Sanders et al. 2020). In addition to dopamine, norepinephrine (the main neurotransmitter produced by the locus coeruleus) has been suggested to signal unexpected uncertainty, when unsignaled context switches produce strongly unexpected observations (Yu & Dayan 2005). Critically, through its effects on synaptic plasticity, norepinephrine has also been implicated in the formation of memories (Palacios-Filardo & Mellor 2019), making it an ideal substrate for inducing memory creation.

The cerebellum may play a key role in memory creation in conditioning paradigms. Neuroimaging data show that a memory formed during fear conditioning remains within the cerebellum following extinction (Batsikadze et al. 2022), suggesting that distinct context-dependent memories have been formed for acquisition and extinction in the cerebellum. Conversely, in cerebellar patients performing an ABA renewal eye-blink conditioning paradigm, extinction (during B) is slowed and renewal (during the second A) is absent (Steiner et al. 2019), consistent with the idea that a single memory has been updated throughout learning, independent of context, and that an intact cerebellum is needed to create new context-specific memories for both the acquisition of and extinction of eye-blink conditioning.

At the circuit level, evidence of the instantiation of a new memory has been observed in the amygdala during fear conditioning and extinction in rodents. This was achieved by tracking a measure of representational distance between the neural population responses to the conditioning stimulus (CS) and population responses to the CS prior to conditioning (**Figure 2e**) and the unconditioned stimulus (US) (Grewe et al. 2017). CS responses became more similar to the US response during conditioning. During extinction, the distance to US response increased, paralleling the reduction in freezing. However, the representation did not return to baseline, as indicated by

the constant distance of responses from the initial CS representation. This is consistent with the idea that the neural circuit has been fundamentally changed by the incorporation of a new memory and hence cannot return to baseline. That is, extinction is not simply the reversal of learning. This result parallels that seen in monkey motor cortex, whereby following deadaptation to a previously learned force field, the neural activity did not return to baseline but instead retained a memory of the force field (Sun et al. 2022).

Further evidence that extinction results in two context-specific memories in the amygdala (conditioning and extinction memories) comes from studies demonstrating that unique neuronal ensembles emerge following extinction. That is, fear neurons and extinction neurons coexist and preferentially fire in response to the same CS when the animal freezes (renewal context) or does not (extinction context), respectively (Orsini et al. 2013).

In cortical areas that use different subspaces of the neural state space to implement context-specific computations (see Section 4.2), memory instantiation may correspond to expanding the neural repertoire into a previously unused (i.e., outside-manifold) subspace (Oby et al. 2019). Moreover, the creation of new memories may also be facilitated by increasing the number of neurons in the circuit by neurogenesis, for example, in the dentate gyrus (Becker 2005, Gershman et al. 2017).

## 5. CONTINUAL LEARNING

Contextual inference might also be key to understanding how the brain solves a fundamental problem that has long haunted learning in ANNs. Although ANNs can solve challenging computational tasks when learned individually (LeCun et al. 2015), albeit typically with extensive training to achieve high performance, crucially, they have difficulty learning multiple tasks, especially when those tasks need to be learned sequentially (continual learning, also known as lifelong learning). As task-optimized ANNs have also been found to predict neural activity in a number of brain areas, including the visual cortex (Yamins et al. 2014, Echeveste et al. 2020) and PFC (Mante et al. 2013, Stroud et al. 2021), understanding how they might be able to perform continual learning may provide useful clues as to how the brain achieves the same.

There are two fundamental factors that make multitask and continual learning difficult in ANNs (Rolnick et al. 2019): interference and catastrophic forgetting. Interference arises when multiple tasks are incompatible, that is, they have different input-output mappings. This happens regardless of whether the tasks are trained in an interleaved or blocked fashion (Kessler et al. 2021). In contrast, catastrophic forgetting can happen even when the tasks are not incompatible. In this case, the tasks are all well learned by an ANN when they are presented in an interleaved manner (standard multitask learning), but when the tasks are experienced in a blocked fashion (continual learning), substantial forgetting is observed (French 1999). Such catastrophic forgetting occurs because the ANN weights are reoptimized for performance on each new task without reference to performance on previous tasks.

A variety of mechanisms have been explored to prevent these problems and promote continual learning in ANNs by using knowledge about context. A powerful solution is to create separate networks (Aljundi et al. 2017), or subnetworks (Rusu et al. 2016, Li & Hoiem 2017), for each task (context) and use an input representing the current context (task identity) to index the appropriate (sub)network that needs to be called upon. To avoid having to grow the network with time, which may limit generalization across tasks and require biologically implausible mechanisms, fixed network architectures with context-dependent gating have also been studied (Masse et al. 2018, Podlaski et al. 2020, Flesch et al. 2022). In these networks, a fraction of the units is used for each task (while the rest are gated by setting their activities to 0). Alternatively, contextual inputs have also been used to induce learned rotations of the input space (Zeng et al. 2019). These rotations

**Interference:** degradation in task performance when two or more tasks are learned that are incompatible (i.e., require different outputs for the same inputs)

**Catastrophic forgetting:** severe degradation of performance on previously learned tasks due to one or more subsequent tasks having been learned

allow the same inputs to map to different outputs for different tasks within the same network, thereby reducing interference. It remains an open question which, if any, of these mechanisms may be relevant to how context controls continual learning in biological circuits.

However, all these approaches are still task aware: They require access to ground-truth information about task identity during training [and often at test time, too (van de Ven et al. 2022)]. This information is often explicitly given in the form of task labels (Rusu et al. 2016, Aljundi et al. 2017, Li & Hoiem 2017, Masse et al. 2018, Zeng et al. 2019), or sometimes it is more implicit, for example, when only the boundaries between tasks are sign posted, but the identities of individual tasks are not provided (Kirkpatrick et al. 2017, Lopez-Paz & Ranzato 2017, Rebuffi et al. 2017, Shin et al. 2017, Zenke et al. 2017). Furthermore, to ensure that each task is learned equally well, several approaches also require networks to be provided with an equal amount of training for each task (Lopez-Paz & Ranzato 2017, Rebuffi et al. 2017, Shin et al. 2017, Kao et al. 2021a). Arguably, this is an unreasonable requirement for a lifelong learning algorithm, as, in general, an equal amount of experience per task cannot be guaranteed in the wild and should not be expected.

In contrast, biological neural networks do not seem to suffer from the same problems. Over the course of a lifetime, humans are able to learn a wide variety of tasks from a continuous stream of sensorimotor experience (Pisupati & Niv 2022). In fact, humans can even benefit from blocked training relative to interleaved training (Flesch et al. 2018)—the exact opposite of how the performance of ANNs depends on the training regimen. This effect could be reproduced in an ANN that received as contextual input a moving average of explicit information about task identity (Flesch et al. 2022). The sluggishness of this moving average was similar to what can be expected from contextual inference under a generative model of context transitions that prefers self-transitions (Heald et al. 2021). This suggests that contextual inference may play a role in biological continual learning, such that different tasks are assigned to different contexts. In this way, context can be more reliably inferred when it is predictable, such as in a blocked design, and as a consequence, proper learning can also be more efficiently focused on the appropriate context-specific representations. In line with this (Lengyel 2022), interference is much reduced in roving paradigms in perceptual learning when using training protocols in which trials corresponding to each task (e.g., defined by the reference orientation in a fine orientation-discrimination paradigm) are either blocked (Dosher et al. 2020) or organized into predicable sequences (Zhang et al. 2008) compared to when they are randomly interleaved (Banai et al. 2010).

Importantly, humans are able to perform continual learning even without ever having been given explicit information about task labels, that is, in a task-agnostic/task-free manner (cf. task-aware continual learning in ANNs). By leveraging contextual (task) inference, continual learning in ANNs can also begin to tackle the ecologically relevant and challenging task-agnostic setting. Some recent progress has been made in this direction using Bayesian nonparametric models (Nagabandi et al. 2018, Jerfel et al. 2019, Xu et al. 2020) that are broadly similar to the models that have been suggested to underlie contextual inference in humans and other animals (see Section 3.4). Other approaches have performed contextual inference without inverting a well-defined generative model. For example, in the dynamic mixture of experts model, reinforcement learning was used to train a gating network that received input from the environment and output a decision regarding which expert to use for the current task (Tsuda et al. 2020). This model was also able to grow in capacity (like Bayesian nonparametric models), adding a new expert (also trained via reinforcement learning) whenever existing experts were unable to solve a task. Some of these approaches (Nagabandi et al. 2018, Jerfel et al. 2019) take context uncertainty into account when performing memory updating, modifying the parameters (in our terminology, contingencies) of all context-specific models in parallel based on the posterior probabilities of their corresponding contexts. However, none of these approaches (Nagabandi et al. 2018, Jerfel et al. 2019, Tsuda et al.

2020, Xu et al. 2020) take context uncertainty into account for expressing memories, as they only express a single context-specific model/expert (the one associated with the most probable context) at any moment.

## 6. CHALLENGES AND OPPORTUNITIES

We have reviewed the computations required for contextual continual learning and elements of their putative neural underpinnings. However, there are still many gaps in our understanding of the neural implementation of contextual inference and memory creation, expression, and updating. For a full understanding across different domains of contextual learning, including for episodic events, motor skills, and decision-making, there are five key computational elements that will need to be mapped onto neural circuit mechanisms.

First, neural circuits need to be able to maintain multiple memories simultaneously. There are mathematical theories for how synaptic and cellular forms of plasticity might achieve this, but comparatively little direct experimental evidence supporting these theories, especially in continual learning settings when multiple memories are explicitly tested. Second, contexts need to be continually inferred, incorporating not only current sensory information but also prior experience with the probabilities of contexts and their transitions. Current experimental evidence for this is sporadic and mostly circumstantial. Outstanding questions include the neural representation of posterior distributions over contexts and the division of labor—and interplay—between PFC, hippocampus, and thalamus in contextual inference. Formal models of contextual inference, similar in spirit to the one we presented here, making quantitative predictions about posterior context distributions under specific conditions, should be useful for discovering and resolving the neural correlates of these posteriors. Third, the neural mechanisms of memory expression need to be context dependent. Fourth, the updating of memories also needs to be context dependent. Fifth, there need to be distinct mechanisms for the creation of altogether new memories. While there are experimental demonstrations for each of these, it remains to be shown that they can go beyond all-or-none operations and instead be graded or probabilistic—specifically as controlled by contextual inferences. Such graded forms of memory processing will likely rely on a large array of mechanisms that are traditionally studied in isolation, such as neural gain control, neuromodulation, synaptic plasticity, and large-scale interactions between cortical and subcortical regions. Understanding the neural bases of context-dependent continual learning will require studying these disparate mechanisms through the unifying lens of appropriate computational frameworks.

## LITERATURE CITED

Aitchison L, Jegminat J, Menendez JA, Pfister JP, Pouget A, Latham PE. 2021. Synaptic plasticity as Bayesian inference. *Nat. Neurosci.* 24(4):565–71

Aljundi R, Chakravarty P, Tuytelaars T. 2017. Expert gate: Lifelong learning with a network of experts. In *Proceedings of the 30th IEEE Conference on Computer Vision and Pattern Recognition*, pp. 3366–75. Piscataway, NJ: IEEE

Anderson MC, Hulbert JC. 2021. Active forgetting: adaptation of memory by prefrontal control. *Annu. Rev. Psychol.* 72:1–36

Banai K, Ortiz JA, Oppenheimer JD, Wright BA. 2010. Learning two things at once: differential constraints on the acquisition and consolidation of perceptual learning. *Neuroscience* 165(2):436–44

Batsikadze G, Diekmann N, Ernst TM, Klein M, Maderwald S, et al. 2022. The cerebellum contributes to context-effects during fear extinction learning: a 7T fMRI study. *NeuroImage* 253:119080

Becker S. 2005. A computational principle for hippocampal learning and neurogenesis. *Hippocampus* 15(6):722–38

Behrens TEJ, Muller TH, Whittington JC, Mark S, Baram AB, et al. 2018. What is a cognitive map? Organizing knowledge for flexible behavior. *Neuron* 100:490–509

Bernardi S, Benna MK, Rigotti M, Munuera J, Fusi S, Salzman CD. 2020. The geometry of abstraction in the hippocampus and prefrontal cortex. *Cell* 183(4):954–67.e21

Botvinick MM. 2008. Hierarchical models of behavior and prefrontal function. *Trends Cogn. Sci.* 12(5):201–8

Bouton ME, King DA. 1983. Contextual control of the extinction of conditioned fear: tests for the associative value of the context. *J. Exp. Psychol. Anim. Behav. Process.* 9(3):248–65

Bradfield LA, Dezfouli A, van Holstein M, Chieng B, Balleine BW. 2015. Medial orbitofrontal cortex mediates outcome retrieval in partially observable task situations. *Neuron* 88(6):1268–80

Braun DA, Aertsen A, Wolpert DM, Mehring C. 2009. Motor task variation induces structural learning. *Curr. Biol.* 19(4):352–57

Burge J, Ernst MO, Banks MS. 2008. The statistical determinants of adaptation rate in human reaching. *J. Vis.* 8(4):20.1–19

Butter CM. 1969. Perseveration in extinction and in discrimination reversal tasks following selective frontal ablations in *Macaca mulatta*. *Physiol. Behav.* 4(2):163–71

Chan SCY, Niv Y, Norman KA. 2016. A probability distribution over latent causes, in the orbitofrontal cortex. *J. Neurosci.* 36(30):7817–28

Chen S, He L, Huang AJ, Boehringer R, Robert V, et al. 2020. A hypothalamic novelty signal modulates hippocampal memory. *Nature* 586(7828):270–74

Collins A, Koechlin E. 2012. Reasoning, learning, and creativity: frontal lobe function and human decision-making. *PLOS Biol.* 10(3):e1001293

Collins AG, Frank MJ. 2013. Cognitive control over learning: creating, clustering, and generalizing task-set structure. *Psychol. Rev.* 120(1):190–229

Collins AGE, Frank MJ. 2016. Neural signature of hierarchically structured expectations predicts clustering and transfer of rule sets in reinforcement learning. *Cognition* 152:160–69

Crossley MJ, Ashby FG, Maddox WT. 2013. Erasing the engram: the unlearning of procedural skills. *J. Exp. Psychol. Gen.* 142(3):710–41

Dayan P, Abbott LF. 2005. *Theoretical Neuroscience: Computational and Mathematical Modeling of Neural Systems*. Cambridge, MA: MIT Press

Donoso M, Collins AGE, Koechlin E. 2014. Foundations of human reasoning in the prefrontal cortex. *Science* 344(6191):1481–86

Dosher BA, Liu J, Chu W, Lu ZL. 2020. Roving: the causes of interference and re-enabled learning in multi-task visual training. *J. Vis.* 20(6):9

Echeveste R, Aitchison L, Hennequin G, Lengyel M. 2020. Cortical-like dynamics in recurrent circuits optimized for sampling-based probabilistic inference. *Nat. Neurosci.* 23(9):1138–49

Flesch T, Balaguer J, Dekker R, Nili H, Summerfield C. 2018. Comparing continual task learning in minds and machines. *PNAS* 115(44):E10313–22

Flesch T, Nagy DG, Saxe A, Summerfield C. 2022. Modelling continual learning in humans with Hebbian context gating and exponentially decaying task signals. arXiv:2203.11560 [q-bio.NC]

Franklin NT, Frank MJ. 2020. Generalizing to generalize: Humans flexibly switch between compositional and conjunctive structures during reinforcement learning. *PLOS Comput. Biol.* 16(4):e1007720

Frémaux N, Gerstner W. 2015. Neuromodulated spike-timing-dependent plasticity, and theory of three-factor learning rules. *Front. Neural Circuits* 9:85

French RM. 1999. Catastrophic forgetting in connectionist networks. *Trends Cogn. Sci.* 3(4):128–35

Gershman SJ, Blei DM, Niv Y. 2010. Context, learning, and extinction. *Psychol. Rev.* 117(1):197–209

Gershman SJ, Monfils MH, Norman KA, Niv Y. 2017. The computational nature of memory modification. *eLife* 6:e23763

Gershman SJ, Radulescu A, Norman KA, Niv Y. 2014. Statistical computations underlying the dynamics of memory updating. *PLOS Comput. Biol.* 10(11):e1003939

Grewe BF, Gründemann J, Kitch LJ, Lecoq JA, Parker JG, et al. 2017. Neural ensemble dynamics underlying a long-term associative memory. *Nature* 543(7647):670–75

Gulli RA, Duong LR, Corrigan BW, Doucet G, Williams S, et al. 2020. Context-dependent representations of objects and space in the primate hippocampus during virtual navigation. *Nat. Neurosci.* 23(1):103–12

Hajnal MA, Tran D, Einstein M, Martelo MV, Safaryan K, et al. 2021. Continuous multiplexed population representations of task context in the mouse primary visual cortex. bioRxiv 2021.04.20.440666. **https://doi.org/10.1101/2021.04.20.440666**

Hamm JP, Shymkiv Y, Han S, Yang W, Yuste R. 2021. Cortical ensembles selective for context. *PNAS* 118(14):e2026179118

Haruno M, Wolpert DM, Kawato M. 2001. MOSAIC model for sensorimotor learning and control. *Neural Comput.* 13(10):2201–20

Heald JB, Ingram JN, Flanagan JR, Wolpert DM. 2018. Multiple motor memories are learned to control different points on a tool. *Nat. Hum. Behav.* 2(4):300–11

Heald JB, Lengyel M, Wolpert DM. 2021. Contextual inference underlies the learning of sensorimotor repertoires. *Nature* 600(7889):489–93

Heald JB, Lengylel M, Wolpert DM. 2023. Contextual inference in learning and memory. *Trends Cogn. Sci.* 27(1):43–64

Hikosaka O, Takikawa Y, Kawagoe R. 2000. Role of the basal ganglia in the control of purposive saccadic eye movements. *Physiol. Rev.* 80(3):953–78

Holt W, Maren S. 1999. Muscimol inactivation of the dorsal hippocampus impairs contextual retrieval of fear memory. *J. Neurosci.* 19(20):9054–62

Hopfield JJ. 1982. Neural networks and physical systems with emergent collective computational abilities. *PNAS* 79(8):2554–58

Howard MW, Kahana MJ. 2002. A distributed representation of temporal context. *J. Math. Psychol.* 46(3):269–99

Jacobs R, Jordan M, Nowlan S, Hinton G. 1991. Adaptive mixtures of local experts. *Neural Comput.* 3(1):79–87

Jerfel G, Grant E, Griffiths T, Heller KA. 2019. Reconciling meta-learning and continual learning with online mixtures of tasks. *Adv. Neural Inf. Process. Syst.* 32:9119–30

Jezek K, Henriksen EJ, Treves A, Moser EI, Moser MB. 2011. Theta-paced flickering between place-cell maps in the hippocampus. *Nature* 478(7368):246–49

Julian JB, Doeller CF. 2021. Remapping and realignment in the human hippocampal formation predict context-dependent spatial behavior. *Nat. Neurosci.* 24(6):863–72

Káli S, Dayan P. 2004. Off-line replay maintains declarative memories in a model of hippocampal-neocortical interactions. *Nat. Neurosci.* 7(3):286–94

Kao TC, Jensen KT, Bernacchia A, Hennequin G. 2021a. Natural continual learning: success is a journey, not (just) a destination. arXiv:2106.08085 [cs.LG]

Kao TC, Sadabadi MS, Hennequin G. 2021b. Optimal anticipatory control as a theory of motor preparation: a thalamo-cortical circuit model. *Neuron* 109(9):1567–1581.e12

Keisler A, Shadmehr R. 2010. A shared resource between declarative memory and motor memory. *J. Neurosci.* 30(44):14817–23

Kessler S, Parker-Holder J, Ball P, Zohren S, Roberts SJ. 2021. Same state, different task: continual reinforcement learning without interference. arXiv:2106.02940 [cs.LG]

Kim JJ, Fanselow MS. 1992. Modality-specific retrograde amnesia of fear. *Science* 256(5057):675–77

Kirkpatrick J, Pascanu R, Rabinowitz N, Veness J, Desjardins G, et al. 2017. Overcoming catastrophic forgetting in neural networks. *PNAS* 114(13):3521–26

Klaus A, da Silva JA, Costa RM. 2019. What, if, and when to move: basal ganglia circuits and self-paced action initiation. *Annu. Rev. Neurosci.* 42:459–83

LeCun Y, Bengio Y, Hinton G. 2015. Deep learning. *Nature* 521(7553):436–44

Lengyel G. 2022. *A common probabilistic framework explaining learning and generalization in perceptual and statistical learning*. PhD thesis, Central European Univ., Vienna

Li Z, Hoiem D. 2017. Learning without forgetting. *IEEE Trans. Pattern Anal. Mach. Intell.* 40(12):2935–47

Lisman JE, Grace AA. 2005. The hippocampal-VTA loop: controlling the entry of information into long-term memory. *Neuron* 46(5):703–13

Logiaco L, Abbott LF, Escola S. 2021. Thalamic control of cortical dynamics in a model of flexible motor sequencing. *Cell Rep.* 35(9):109090

Lopez-Paz D, Ranzato M. 2017. Gradient episodic memory for continual learning. *Adv. Neural Inf. Process. Syst.* 30:6467–76

Mante V, Sussillo D, Shenoy KV, Newsome WT. 2013. Context-dependent computation by recurrent dynamics in prefrontal cortex. *Nature* 503(7474):78–84

Markus EJ, Qin YL, Leonard B, Skaggs WE, McNaughton BL, Barnes CA. 1995. Interactions between location and task affect the spatial and directional firing of hippocampal neurons. *J. Neurosci.* 15(11):7079–94

Marton TF, Seifikar H, Luongo FJ, Lee AT, Sohal VS. 2018. Roles of prefrontal cortex and mediodorsal thalamus in task engagement and behavioral flexibility. *J. Neurosci.* 38(10):2569–78

Masse NY, Grant GD, Freedman DJ. 2018. Alleviating catastrophic forgetting using context-dependent gating and synaptic stabilization. *PNAS* 115(44):E10467–75

McNamara CG, Tejero-Cantero Á, Trouche S, Campo-Urriza N, Dupret D. 2014. Dopaminergic neurons promote hippocampal reactivation and spatial memory persistence. *Nat. Neurosci.* 17(12):1658–60

Nagabandi A, Finn C, Levine S. 2018. Deep online learning via meta-learning: continual adaptation for model-based RL. arXiv:1812.07671 [cs.LG]

Napier RM, Macrae M, Kehoe EJ. 1992. Rapid reacquisition in conditioning of the rabbit's nictitating membrane response. *J. Exp. Psychol. Anim. Behav. Process.* 18(2):182–92

Nassar MR, McGuire JT, Ritz H, Kable JW. 2019. Dissociable forms of uncertainty-driven representational change across the human brain. *J. Neurosci.* 39(9):1688–98

Oby ER, Golub MD, Hennig JA, Degenhart AD, Tyler-Kabara EC, et al. 2019. New neural activity patterns emerge with long-term learning. *PNAS* 116(30):15210–15

Oh Y, Schweighofer N. 2019. Minimizing precision-weighted sensory prediction errors via memory formation and switching in motor adaptation. *J. Neurosci.* 39(46):9237–50

O'Keefe J, Nadel L. 1978. *The Hippocampus as a Cognitive Map*. Oxford, UK: Clarendon Press

Orsini CA, Yan C, Maren S. 2013. Ensemble coding of context-dependent fear memory in the amygdala. *Front. Behav. Neurosci.* 7:199

Palacios-Filardo J, Mellor J. 2019. Neuromodulation of hippocampal long-term synaptic plasticity. *Curr. Opin. Neurobiol.* 54:37–43

Parnaudeau S, O'Neill PK, Bolkan SS, Ward RD, Abbas AI, et al. 2013. Inhibition of mediodorsal thalamus disrupts thalamofrontal connectivity and cognition. *Neuron* 77(6):1151–62

Parnaudeau S, Taylor K, Bolkan SS, Ward RD, Balsam PD, Kellendonk C. 2015. Mediodorsal thalamus hypofunction impairs flexible goal-directed behavior. *Biol. Psych.* 77(5):445–53

Pfeiffer BE. 2020. The content of hippocampal "replay". *Hippocampus* 30(1):6–18

Piray P, Daw ND. 2020. A simple model for learning in volatile environments. *PLOS Comput. Biol.* 16(7):e1007963

Pisupati S, Niv Y. 2022. The challenges of lifelong learning in biological and artificial systems. *Trends Cogn. Sci.* 26(12):1051–53

Plitt MH, Giocomo LM. 2021. Experience-dependent contextual codes in the hippocampus. *Nat. Neurosci.* 24(5):705–14

Podlaski WF, Agnes EJ, Vogels TP. 2020. Context-modular memory networks support high-capacity, flexible, and robust associative memories. bioRxiv 2020.01.08.898528. **https://doi.org/10.1101/2020. 01.08.898528**

Rebuffi SA, Kolesnikov A, Sperl G, Lampert CH. 2017. iCaRL: incremental classifier and representation learning. In *Proceedings of the 2017 IEEE Conference on Computer Vision and Pattern Recognition*, pp. 2001–10. Piscataway, NJ: IEEE

Redish AD, Jensen S, Johnson A, Kurth-Nelson Z. 2007. Reconciling reinforcement learning models with behavioral extinction and renewal: implications for addiction, relapse, and problem gambling. *Psychol. Rev.* 114(3):784–805

Remington ED, Egger SW, Narain D, Wang J, Jazayeri M. 2018a. A dynamical systems perspective on flexible motor timing. *Trends Cogn. Sci.* 22(10):938–52

Remington ED, Narain D, Hosseini EA, Jazayeri M. 2018b. Flexible sensorimotor computations through rapid reconfiguration of cortical dynamics. *Neuron* 98(5):1005–19.e5

Rescorla RA. 2004. Spontaneous recovery. *Learn. Mem.* 11(5):501–9

Rescorla RA, Heth CD. 1975. Reinstatement of fear to an extinguished conditioned stimulus. *J. Exp. Psychol. Anim. Behav. Process.* 1(1):88–96

Rescorla RA, Wagner AR. 1972. A theory of Pavlovian conditioning: variations in the effectiveness of reinforcement and nonreinforcement. In *Classical Conditioning II: Current Theory and Research*, ed. A Black, W Prokasy, pp. 64–99. New York: Appleton-Century-Crofts

Rikhye RV, Wimmer RD, Halassa MM. 2018. Toward an integrative theory of thalamic function. *Annu. Rev. Neurosci.* 41:163–83

Rolnick D, Ahuja A, Schwarz J, Lillicrap T, Wayne G. 2019. Experience replay for continual learning. *Adv. Neural Inf. Process. Syst.* 32:348–58

Rusu AA, Rabinowitz NC, Desjardins G, Soyer H, Kirkpatrick J, et al. 2016. Progressive neural networks. arXiv:1606.04671 [cs.LG]

Saez A, Rigotti M, Ostojic S, Fusi S, Salzman CD. 2015. Abstract context representations in primate amygdala and prefrontal cortex. *Neuron* 87(4):869–81

Sanders H, Wilson MA, Gershman SJ. 2020. Hippocampal remapping as hidden state inference. *eLife* 9:e51140

Savin C, Dayan P, Lengyel M. 2014. Optimal recall from bounded metaplastic synapses: predicting functional adaptations in hippocampal area CA3. *PLOS Comput. Biol.* 10(2):e1003489

Schuck NW, Cai MB, Wilson RC, Niv Y. 2016. Human orbitofrontal cortex represents a cognitive map of state space. *Neuron* 91(6):1402–12

Schuck NW, Niv Y. 2019. Sequential replay of nonspatial task states in the human hippocampus. *Science* 364(6447):eaaw5181

Shallice T, Cipolotti L. 2018. The prefrontal cortex and neurological impairments of active thought. *Annu. Rev. Psychol.* 69:157–80

Shin H, Lee JK, Kim J, Kim J. 2017. Continual learning with deep generative replay. arXiv:1705.08690 [cs.AI]

Singh S. 1991. The efficient learning of multiple task sequences. *Adv. Neural Inf. Process. Syst.* 4:251–58

Smith MA, Ghazizadeh A, Shadmehr R. 2006. Interacting adaptive processes with different timescales underlie short-term motor learning. *PLOS Biol.* 4(6):e179

Sohn H, Narain D, Meirhaeghe N, Jazayeri M. 2019. Bayesian computation through cortical latent dynamics. *Neuron* 103(5):934–47.e5

Soltani A, Koechlin E. 2022. Computational models of adaptive behavior and prefrontal cortex. *Neuropsychopharmacology* 47(1):58–71

Steiner KM, Gisbertz Y, Chang DI, Koch B, Uslar E, et al. 2019. Extinction and renewal of conditioned eyeblink responses in focal cerebellar disease. *Cerebellum* 18(2):166–77

Stroud JP, Porter MA, Hennequin G, Vogels TP. 2018. Motor primitives in space and time via targeted gain modulation in cortical networks. *Nat. Neurosci.* 21(12):1774–83

Stroud JP, Watanabe K, Suzuki T, Stokes MG, Lengyel M. 2021. Optimal information loading into working memory in prefrontal cortex. bioRxiv 2021.11.16.468360. **https://doi.org/10.1101/2021. 11.16.468360**

Sun X, O'Shea DJ, Golub MD, Trautmann EM, Vyas S, et al. 2022. Cortical preparatory activity indexes learned motor memories. *Nature* 602(7896):274–79

Sutton RS, Barto AG. 2018. *Reinforcement Learning: An Introduction*. Cambridge, MA: MIT Press. 2nd ed.

Takeuchi T, Duszkiewicz AJ, Sonneborn A, Spooner PA, Yamasaki M, et al. 2016. Locus coeruleus and dopaminergic consolidation of everyday memory. *Nature* 537(7620):357–62

Teh YW, Jordan MI, Beal MJ, Blei DM. 2006. Hierarchical Dirichlet processes. *J. Am. Stat. Assoc.* 101(476):1566–81

Teyler TJ, Rudy JW. 2007. The hippocampal indexing theory and episodic memory: updating the index. *Hippocampus* 17(12):1158–69

Tsuda B, Tye KM, Siegelmann HT, Sejnowski TJ. 2020. A modeling framework for adaptive lifelong learning with transfer and savings through gating in the prefrontal cortex. *PNAS* 117(47):29872–82

van de Ven GM, Tuytelaars T, Tolias AS. 2022. Three types of incremental learning. *Nat. Mach. Intell.* 4:1185–97

Vyas S, Golub MD, Sussillo D, Shenoy KV. 2020. Computation through neural population dynamics. *Annu. Rev. Neurosci.* 43:249–75

Wallis JD, Anderson KC, Miller EK. 2001. Single neurons in prefrontal cortex encode abstract rules. *Nature* 411(6840):953–56

Wang J, Narain D, Hosseini EA, Jazayeri M. 2018. Flexible timing by temporal scaling of cortical responses. *Nat. Neurosci.* 21(1):102–10

Wikenheiser AM, Schoenbaum G. 2016. Over the river, through the woods: cognitive maps in the hippocampus and orbitofrontal cortex. *Nat. Rev. Neurosci.* 17(8):513–23

Wilson RC, Takahashi YK, Schoenbaum G, Niv Y. 2014. Orbitofrontal cortex as a cognitive map of task space. *Neuron* 81(2):267–79

Wolpert DM, Kawato M. 1998. Multiple paired forward and inverse models for motor control. *Neural Netw.* 11(7–8):1317–29

Wolpert DM, Miall RC, Kawato M. 1998. Internal models in the cerebellum. *Trends Cogn. Sci.* 2(9):338–47

Wu Z, Litwin-Kumar A, Shamash P, Taylor A, Axel R, Shadlen MN. 2020. Context-dependent decision making in a premotor circuit. *Neuron* 106(2):316–28.e6

Xie A, Harrison J, Finn C. 2021. Deep reinforcement learning amidst continual structured non-stationarity. *PMLR* 139:11393–403

Xu M, Ding W, Zhu J, Liu Z, Chen B, Zhao D. 2020. Task-agnostic online reinforcement learning with an infinite mixture of Gaussian processes. *Adv. Neural Inf. Process. Syst.* 33:6429–40

Yamins DL, Hong H, Cadieu CF, Solomon EA, Seibert D, DiCarlo JJ. 2014. Performance-optimized hierarchical models predict neural responses in higher visual cortex. *PNAS* 111(23):8619–24

Yu AJ, Dayan P. 2005. Uncertainty, neuromodulation, and attention. *Neuron* 46(4):681–92

Zacks JM, Speer NK, Swallow KM, Braver TS, Reynolds JR. 2007. Event perception: a mind-brain perspective. *Psychol. Bull.* 133(2):273–93

Zeng G, Chen Y, Cui B, Yu S. 2019. Continual learning of context-dependent processing in neural networks. *Nat. Mach. Intell.* 1(8):364–72

Zenke F, Poole B, Ganguli S. 2017. Continual learning through synaptic intelligence. *PMLR* 70:3987–95

Zhang JY, Kuai SG, Xiao LQ, Klein SA, Levi DM, Yu C. 2008. Stimulus coding rules for perceptual learning. *PLOS Biol.* 6(8):e197