nature reviews neuroscience

https://doi.org/10.1038/s41583-024-00893-z

Perspective

Schemas, reinforcement learning and the medial prefrontal cortex

Oded Bein [1,2] ✉ **& Yael Niv** [1,3]

Abstract

Schemas are rich and complex knowledge structures about the typical unfolding of events in a context; for example, a schema of a dinner at a restaurant. In this Perspective, we suggest that reinforcement learning (RL), a computational theory of learning the structure of the world and relevant goal-oriented behaviour, underlies schema learning. We synthesize literature about schemas and RL to offer that three RL principles might govern the learning of schemas: learning via prediction errors, constructing hierarchical knowledge using hierarchical RL, and dimensionality reduction through learning a simplified and abstract representation of the world. We then suggest that the orbitomedial prefrontal cortex is involved in both schemas and RL due to its involvement in dimensionality reduction and in guiding memory reactivation through interactions with posterior brain regions. Last, we hypothesize that the amount of dimensionality reduction might underlie gradients of involvement along the ventral–dorsal and posterior–anterior axes of the orbitomedial prefrontal cortex. More specific and detailed representations might engage the ventral and posterior parts, whereas abstraction might shift representations towards the dorsal and anterior parts of the medial prefrontal cortex.

Sections

---

[1] Princeton Neuroscience Institute, Princeton University, Princeton, NJ, USA. [2] Weill Cornell Institute of Geriatric Psychiatry, Department of Psychiatry, Weill Cornell Medicine, New York, NY, USA. [3] Psychology Department, Princeton University, Princeton, NJ, USA. ✉e-mail: odb4002@med.cornell.edu

Introduction

Imagine entering a restaurant. You immediately know the likely sequence of occurrences and the relevant set of behaviours. You will be seated at a table and given a menu. After placing your order, you will hopefully receive a delicious meal and maybe a glass of fine wine. This will be followed by paying the bill and leaving the restaurant. The general knowledge of what typically occurs in an event and in what order, as well as the appropriate behaviour, is referred to as its 'schema'[1–4]. Although schemas are widely investigated in psychology and, more recently, in neuroscience, they also remain notoriously elusive and ill-defined[2,5]. Importantly, in schema theory, a satisfying computational account of how schemas are learned through experience, guide goal-oriented behaviour, and influence perception, attention, learning and memory is lacking.

Reinforcement learning (RL) offers a computational theory of how humans and animals learn goal-oriented behaviours through experience[6]. Schemas are thought to represent information about the environment that is useful for such behaviours (Box 1). In this Perspective, we synthesize research from the seemingly disparate fields of schemas and RL to propose that RL, and complementary algorithms such as dimensionality reduction and latent cause inference, provide a quantitative framework for schema theory. We begin with a brief description of schemas and RL mechanisms to show how these are related. We then focus on three core computational principles that could underlie schemas: learning a summary of the environment through prediction errors, grouping of states through hierarchical RL and latent cause inference, and dimensionality reduction through learning of abstract state representations (Fig. 1). We then build on evidence emphasizing the importance of the medial prefrontal cortex (mPFC) and the medial orbitofrontal cortex (mOFC) to both RL[7–11] and schemas[5,12–16] to hypothesize that these regions mediate dimensionality reduction and guide memory retrieval through communicating with posterior brain regions. We conclude by postulating that graded recruitment along the ventral–dorsal and anterior–posterior axes of the mPFC might reflect the amount of dimensionality reduction required in a current situation.

Conceptual mapping of schema to reinforcement learning theory

In this section, we first briefly introduce the concept of schemas, noting that there is no satisfying computational account for understanding how schemas are learned, instantiated and deployed. Then, we briefly introduce RL as a potential account of schemas. The rest of the paper will then elaborate on this mapping between the mechanisms of RL and the phenomenology of schemas, as well as their neural underpinnings.

Schemas

Schemas are learned knowledge structures that organize knowledge of what typically occurs in a context[2,4,17–19], including associative knowledge of relationships and co-occurrences between the components of recurring events (for example, 'menu' and 'food' as components of a meal at a restaurant). Schemas are learned through the extraction of commonalities across multiple experiences, also termed episodes[2], and as such, schemas are devoid of specific episode details. In this Perspective, we predominantly discuss schemas that are extended in time (similar to the notion of 'scripts'[3] or 'event schemas'[20]) and thus include knowledge of the temporal structure of an event. The process of retrieving schemas from memory is termed schema instantiation. Once instantiated, schemas can be deployed to guide behaviour as they include knowledge of context-appropriate actions (for example, the knowledge that upon receiving a menu, one should read it and place an order). Last, schemas can be thought of as hierarchically organized 'modules' that can be recombined. Both the schema of a restaurant and that of having dinner at home can include a module of sitting at the table and eating, and the schema of an airport can include a restaurant as a module. Thus, schemas can be a part of other schemas, as well as include other schemas.

Despite decades of research on the influence of schemas on cognition[1,4,5,19], it is not completely clear how schemas are learned and instantiated and how they influence perception, action, learning and memory[16,21,22]. Computational models of semantic networks, concepts and category learning[18,23–29] characterize some aspects of schema learning, such as how general knowledge about the co-occurrence of entities is extracted, but do not seem to capture fully the scope and richness of schemas. Particularly relevant is the fact that schemas are learned and instantiated through experiences that are multidimensional, dynamic in time and involve goal-oriented actions. For these, we turn to the framework of RL.

Reinforcement learning

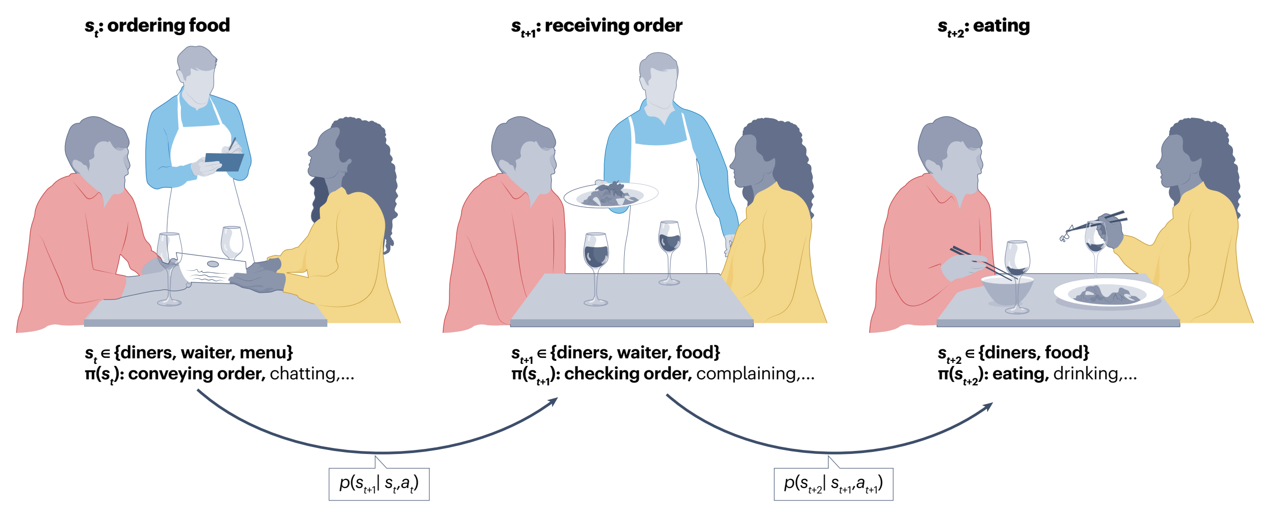
RL provides a set of algorithms for goal-oriented learning and behaviour, in which the goal is typically conceptualized as maximizing reward while minimizing costs or punishments[6]. Through trial and error over multiple instances of a task, an agent learns the sequence of actions most suitable for achieving maximal reward in an environment. In RL theory, tasks are divided into a series of discrete timepoints or contexts, termed 'states'. For instance, a visit to a restaurant can be divided into the states of standing at the entrance, sitting down at the table or having a menu in hand. Each state has an associated action policy – the probability of taking each action available at that state. A state can also be associated with a value, which denotes the expected sum of (possibly discounted) future rewards when in that state and assuming a specific action policy (values can also be learned for each action taken in each state, in which case they are termed state-action values). Tasks can be divided into states at different levels of coarseness and, similarly, action policies can be defined as single actions or high-level action groupings (for example, 'adding salt' groups reaching for the salt, grabbing it and sprinkling salt on the food. See Hierarchical RL and latent cause inference may contribute to learning and instantiating schema hierarchies).

In addition to learning an optimal action policy – what actions lead to maximizing reward in each state – in a sequential task that extends over time, the agent can learn a model of the environment, that is, the probability of transitioning between different states contingent on different actions[30–32], and the probability of encountering (good or bad) outcomes at each state. In RL, learning occurs when an agent experiences a prediction error: a situation in which the actual outcome is different from the expected one[33–35]. Prediction errors include both reward prediction errors, which refer to obtaining more or less reward than expected, and state prediction errors, which refer to transitioning to a state that is different than expected. Updating expectations according to the prediction error aligns expectations with actual outcomes and reduces future prediction errors. In this way, through experience, the agent can learn a world model, which includes representations of states, state-transition probabilities (G) and the distribution of rewards in each state, and can mentally simulate actions within this learned world model to determine which action is best in what situation (termed 'model-based RL' or 'goal-directed decision making'[30]). Alternatively, in 'model-free RL', the agent can learn an

Box 1 | Partial mapping of schemas to standard reinforcement learning components

Schema components mapped onto reinforcement learning (RL) nomenclature (states, action policy and state-transition probability; see the figure) using three consecutive states (s, denoted by the subscripts *t*, t+1 and *t*+2) in the schema of a restaurant. Essential components (∈{…}) of the situation (such as menu or food) differentiate each state from other states (some, but not necessarily all such components are denoted in the figure). Dimensionality reduction and state abstraction can be observed in the schema. For example, the ordering food state (s_t) includes a menu, but does not specify whether this is physical, online or even memorized, as this can change across episodes (dimensionality reduction). Episodic details such as the specific menu items in s_t and the colour of the shirts and the use of chopsticks in the following receiving order (s_{t+1}) and eating (s_{t+2}) states are not included in the schema (state abstraction). In this way, each state is a generic one, and its learned action policy — a probability distribution across possible actions in a state (π(s)) — will pertain to episodes in many similar situations. The three states in this schema of a restaurant have two actions for each state, ordered by their probability from high to low (see the figure, shades of grey). The probability of transitioning from one state to another (state-transition probability: $p(s'|s,a)$) depends on the previous state (s) and the action taken (*a*), and is often not deterministic. For example, if an order is not conveyed in s_t, the probability of transitioning to s_{t+1} and receiving food is markedly lower, and even if food is ordered, a sudden power outage that closes the kitchen can result in the order not being received and no state transition from s_t to s_{t+1}. As schemas are hierarchical, each state can also further divide into a sequence of substates and respective policies.

Despite this close mapping, we are not claiming that RL can account for all aspects of schemas, and instead hypothesize that additional processes such as dimensionality reduction and latent cause inference mediate schema learning (Fig. 1). Indeed, central features of RL, such as reward and value, might not be an essential part of schemas, but that might depend on the level of abstraction or how well learned a schema is. For example, if the food is bad, that might influence the choice of dish or restaurant in the next restaurant experience, but it is unlikely to change the general representation of states, state transitions and action policies in the schema of a restaurant. Of course, if this bad experience repeats itself, one might choose to stop going to restaurants, which could eventually lead to loss of schema knowledge (especially if restaurant culture changes in the meantime, unbeknownst to those converted into home diners). Thus, rewarding experiences might be important to motivate restaurant visits, but might not be essential to learning the components of the restaurant schema. Consistent with this, empirical studies reveal schema representations in the absence of explicit reward[9,12,173].



action policy directly without learning a world model, from trial and error using reward prediction errors.

Core computational principles that could underlie schemas

The above descriptions suggest how schemas might be mapped to a representation of a task, including the world model and the action policy (Box 1). We now turn to computational principles that could support this mapping (Fig. 1). We start by asking whether schemas are learned through prediction errors. Then, we ask whether the hierarchical nature of schemas might be learned and instantiated via hierarchical RL algorithms and complementary algorithms such as latent cause inference. We conclude this section by asking how dimensionality reduction might mediate schema learning and instantiation.

Are schemas learned through prediction errors?

As RL algorithms use prediction error-driven learning, the first question we ask is whether schemas are also learned and updated via prediction errors (Fig. 1a). The alternative hypothesis is that a summary of the typical and repeating structure of the world is learned by tracking the frequency of occurrences (‘unsupervised learning’[26]). In this frequency hypothesis, learning does not require a prediction and an update

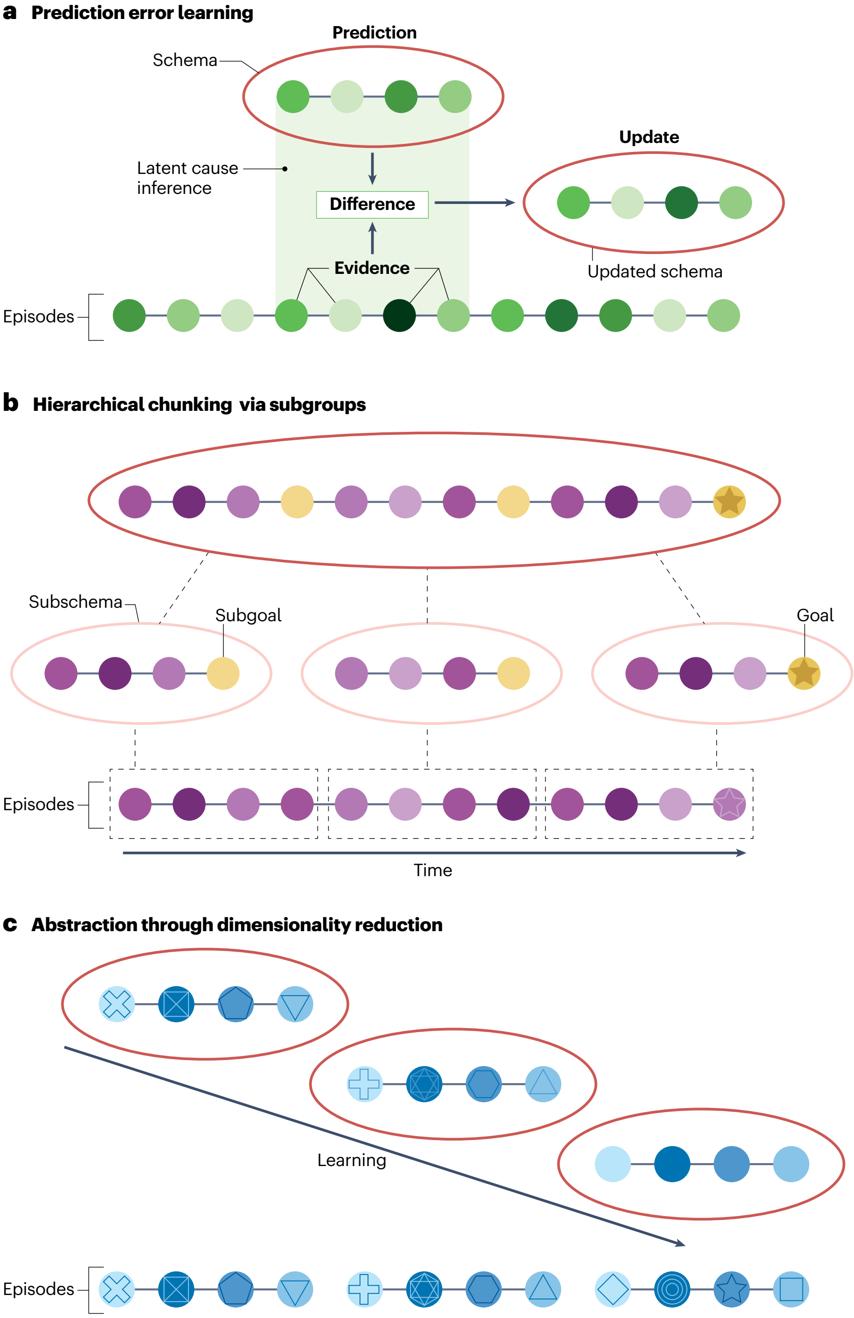


Fig. 1 | Three reinforcement learning principles contribute to schema learning. In the schemas and the episodes, circles represent states or timepoints, with different shades representing different state features (see Box 1). **a**, Prediction errors, namely the difference between a schema-based prediction and the evidence from a specific episode, drive schema updating. Both the evidence and the schema are selected through latent cause inference from episodes in the stream of one's experience, and thus schema updating eventually converges to the typical unfolding of events across episodes. **b**, The hierarchical structure of schemas is learned via identifying subgoals (yellow) that chunk subschemas. Subgoals are states that are identified as breaking points, which are used to train policies of subschemas. How subgoals are identified is an open question, potentially relying on transitions between sequences of frequently occurring states or the discovery of optimal breaking points. We additionally propose salient changes as putative subgoals, offering a mechanism that can identify subgoals rapidly. **c**, Dimensionality reduction, implemented via schema-guided attention, mediates the elimination of episodic details that differ across similar episodes (symbols) during schema learning. This dimensionality reduction results in the inclusion of goal-relevant information in the schema, as well as other repeating information that is not necessarily goal relevant (blue shades).

following a prediction error; instead, learning occurs by averaging of experience, regardless of predictions and their potential violation.

The discovery of 'blocking'[36,37] led animal-learning theorists to shift from assuming that the frequency of co-occurrence (contiguity) is sufficient for associative learning, considering that prediction errors drive learning. In blocking, a neutral stimulus (such as light) previously associated with a motivationally relevant outcome (such as an electric shock or food) prevents a co-occurring neutral stimulus (such as a tone) from also becoming associated with the same outcome. As the first stimulus fully predicts the outcome, no prediction error is possible when the outcome occurs, and thus associative learning about the newly added stimulus is 'blocked'[34,35]. In humans, a wealth of research shows that reward prediction errors not only drive learning[33,38,39], but also facilitate long-term memory[40–43].

As schemas include state-transition probabilities (Box 1), to establish that prediction errors drive schema learning, one can test whether state prediction errors lead to updating of the schema (Fig. 1a) and to changes in behaviour. Recent work in rodents showed blocking of learning of simple stimulus–stimulus associations, thereby establishing that learning of 'neutral' associations required state prediction errors[44,45]. Computational models that learn via state prediction errors explained human and animal choice data in studies that involved frequent changes (reversals) of state-transition probabilities[46–50]. Studies also showed that participants trained on state transitions exhibit enhanced memory for items that violated these transitions[51–55] and reduced memory for items that cued the (surprisingly) not-transitioned-to future state[56,57]. This is consistent with updating a model of the world through state prediction errors, whereby the violating information is encoded and the incorrect prediction is weakened. Whether schemas are learned only via prediction errors or through a combination of frequency of occurrences and prediction errors is a question for future research (see Box 2 for a related discussion on the initial learning of schemas).

These studies, which focused on simplified tasks that trained participants on few associations and over a few trials or sessions, provide evidence that initial learning of schemas might be driven by state prediction errors. By contrast, consolidated and well-learned semantic knowledge is thought to be stable and less amenable to change, largely supported by cortical structures (whereas newly acquired knowledge is supported by the hippocampus; Box 2), and to be more abstract and include fewer specific episodic details[58–62]. Thus, it is not clear that these previous findings from simplified tasks generalize to the updating of complex and well-learned schemas, as work in humans showed that complex semantic knowledge can both impair and enhance learning and memory of new associations[63–71].

In the complexity of everyday life, cues and outcomes are not as clearly defined as in many of the previously mentioned studies, but rather dynamically evolve in time and span multiple temporal scales[72–74]. Indeed, viewers of continuous sport games remembered events within the games that elicited prediction errors better than events that did not[75,76]. In another study that directly targeted the updating of memories, Sinclair and colleagues[77] used rich movie-clip stimuli to elicit predictions of action outcomes learned over a lifetime of everyday experience (for example, a batter hitting a home run during a baseball game). They then violated these action–outcome predictions by stopping the movie clips before the expected outcome and moving on to the next (potentially semantically related) clip. In a subsequent memory test of the movie clips, participants demonstrated memory intrusions, recalling details from the semantically related movie clips as if they were in clips that were stopped prematurely[77,78]. These intrusions might reflect memory update of the movie-clip memories that was enhanced by the violations of expectations.

These studies[51–55,75,77–79], consistent with learning and updating of schemas, tested for memory of unique (one-trial) episodes but not for memory of schema that is semantic in nature. Historically, semantic versus episodic memories have been thought to rely on distinct systems[80,81] (despite early recognition of the interactive nature of these systems[1,4]). However, recent views emphasize overlapping and highly interactive memory representations[82,83]. In our view, as episodes are used to update schemas (via prediction error-based learning), better memory of episodes that elicit a prediction error suggests that these episodes are prioritized in learning and are more likely to influence the updating of schemas. Likewise, poor memory of episodes in which a stimulus or state elicited erroneous predictions might reflect the downweighing of these episodes in representations of schemas. Nonetheless, the specific mechanism by which better memory for unique episodes reflects or influences learning and updating of schemas remains to be elucidated.

In summary, emerging literature suggests that schemas might indeed be learned and updated via prediction errors, similar to learning in RL.

Hierarchical reinforcement learning and latent cause inference may contribute to learning and instantiating schema hierarchies

Schemas are hierarchically organized: each schema can be composed of subschemas and might be a subschema of another, larger schema. Hierarchical RL algorithms[84–91] might provide a blueprint for how such a schema hierarchy is acquired (Fig. 1b). Learning via RL algorithms can become prohibitively slow in complex environments, but hierarchically grouping states and actions into larger units can provide a mechanism to alleviate this scaling problem. In such 'temporal abstraction'[85,92], hierarchical RL algorithms divide a temporally extended task into subunits, called 'subtasks'. Each subtask is defined by a set of possible start states, a subtask-specific action policy and a set of termination states – also called 'subgoals' – in which the subtask will cede control back to the overarching action policy (Fig. 1b, middle and top)[88,90–94]. Subtasks can be defined by a submodel, which includes the states, state-transition probabilities and rewards in the subtask. For example, 'adding salt' could be a subtask that has its start state initiated upon tasting bland food, continues with an action policy that includes reaching for the salt shaker, grasping it and shaking it over the food, and then terminates because its subgoal to salt the food has been reached. Subtasks can be used across disparate tasks[88,95,96] (for example, the 'adding salt' subtask could be used in both 'dining at a restaurant' and 'eating at home' tasks). The term 'subgoal' distinguishes the termination state of the 'adding salt' subtask (food is salted) from the overall goal of the 'dining at a restaurant' and 'eating at home' tasks (having a full stomach; Fig. 1b). In some hierarchical RL algorithms, reaching a subgoal leads to a pseudo-reward signal[88,97,98]. Pseudo-rewards allow standard RL algorithms to learn the optimal action policy for the subtask that maximizes pseudo-rewards, in the same way as these algorithms learn reward-maximizing policies in non-hierarchical settings.

Box 2 | Rapid learning in the hippocampus shapes new schemas

The idea that event boundaries can become subgoals could mean that first instances of events — in which subgoals are created — might be highly influential in shaping models of the structure of the world. This contrasts with the idea that the structure of the environment is extracted solely through incremental and relatively slow learning[102,273]. In reinforcement learning (RL) models, the initial values bias learning and can be hard to overcome[6]. To avoid this, in many RL algorithms, the rate of updating — the learning rate — is high at the beginning of a task and decreases with time. Indeed, Shteingart et al.[274] showed that the first trial experience greatly biases choices in future trials, and that assuming full updating on the first trial (a learning rate of 1) best explained this behaviour. Other studies showed relatively quick learning of regularities[165,207,261,275] and generalization based on such regularities[119] within tens of trials. The hippocampus, known to be involved in rapid learning[61,273,276,277], also mediates learning of the structure of the environment[207,271,278–288]. The hippocampus also plays a central role in event segmentation[112,289–291], which might contribute to the initial learning of schemas (see Hierarchical RL and latent cause inference may contribute to learning and instantiating schema hierarchies). Thus, converging behavioural and neural evidence suggests that rapid initial learning shapes schema learning.

Box 3 | Schemas are broader than cognitive maps

Similarly to schemas, a cognitive map organizes aspects of an experience and can be used to flexibly guide behaviour[147,292–294]. However, we suggest that schemas can include additional types of information and are broader than cognitive maps[1,18,24,26,280]. In most traditional conceptualizations of cognitive maps, information is represented through some notion of distance, which can be physical or mental[292,293] (indeed, although cognitive maps have been studied extensively in spatial navigation, recent research has extended the notion of cognitive maps to non-spatial maps[161,162,165,166,281,286,295,296]). Such distance relationships are, by definition, symmetrical, whereas relationships in schemas do not have to be symmetrical. Others view cognitive maps as more akin to state-transition probabilities in reinforcement learning (RL) (the probability of transitioning from one state to another; Box 1), potentially generalized over several tasks[154,292], which do not have to be symmetrical. Some of these proposals also include in the cognitive map the expected reward in each state[154]. In any case, cognitive maps and state-transition functions do not include the action policy (the probability of executing an action per state; Box 1). By contrast, action policies and options (hierarchical policies) are central to schemas, thus schemas are broader than cognitive maps.

Another sense by which schemas are broader than cognitive maps is that schemas may also include semantic relationships that are difficult to reduce to a distance measure or a state-transition probability (for example, that restaurants include menus). Such relationships have been represented in computational models in various ways, including symbolic representations[24,273,297–299]. How such semantic relationships are encoded in the brain and how they could be integrated with RL to account for schema learning and instantiation are open questions.

An important question in hierarchical RL is how to select subgoals. In terms of schemas, the analogous question is how to segment continuous experience into discrete event schemas[20]. Hierarchical RL offers more than one algorithm[88,90,91,94]. Some algorithms rely on exploring an environment while keeping track of sequences of states and actions that co-occur frequently (statistical learning[99–103]) and use states preceding a transition to another sequence of frequently co-occurring states as subgoals[85,87,104]. Other algorithms use Bayesian inference to discover optimal hierarchal organizations of tasks into subtasks given the structure of the environment[91,94] and the cost of planning[90].

The above algorithms for subgoal discovery all rely on repeated experience to construct a hierarchical model of the world. We propose that saliency can also trigger the creation of a subgoal. Salient stimuli create an intrinsic reward signal and engage motivation-related neural systems, much like rewards[105–111]. Research on event segmentation that focuses on how ongoing and continuous experience is chunked into discrete events[20,112,113] has shown that salient changes, termed 'event boundaries', cause humans to segment their experiences in memory. For example, events that span an event boundary are remembered as happening farther apart in time from each other, and their temporal order is often remembered worse than that of events not separated by a boundary[114–116]. This suggests that event boundaries, like subgoals, structure experience into discrete, segmented units. Indeed, reward prediction errors have been shown to structure memories[117], consistent with our idea that such salience-induced prediction errors create subgoals. A mechanism that relies on salient changes to create subgoals does not require repetition as even in the first instance a change of context, perceptual details or internal state can trigger event segmentation[112,113,115,118]. This discrete event representation could then form a base that future episodes might join, eventually culminating to an event schema. This proposal resonates with recent behavioural work suggesting that schemas can be created rapidly[119] (Box 2). Such rapid extraction of the structure of the environment can facilitate goal-oriented learning and behaviour in new situations[95,120,121], with later learning refining that initial structure extracted[122,123].

Latent cause inference might be the computational process by which salient changes both initiate a new schema and instantiate existing (sub)schemas. Latent cause inference is a computational theory of how observations that are similar to each other are grouped into clusters ('latent causes')[22,95,124,125]. The latent cause underlying the current observations can be inferred using Bayesian inference by combining prior beliefs about the probability of various latent causes (for example, the latent cause responsible for recent observations is most likely to underlie the next observation) with evidence from current observations. Thus, observations that are sufficiently different from existing latent causes prompt the creation of a new latent cause[126,127]. Recent theoretical work has begun to explore how salient changes, such as event boundaries, trigger the inference of a new latent cause[22,113] or instantiate a relevant event schema[22]. Latent cause inference can facilitate the grouping of states into subtasks in hierarchical RL, such that each latent cause is treated as a separate subtask. In this way, for each latent cause, the agent can learn a model (and submodels, as latent causes can be hierarchically constructed) and an action policy, as in the hierarchical RL models above.

Dimensionality reduction through selective attention might mediate schema learning

Schemas summarize information across multiple multidimensional episodes. Learning this summary could be conceptualized as two processes: a form of dimensionality reduction and state abstraction (Fig. 1c). Regarding the former, representing the unique features of each experience probably requires high-dimensional representations, and schemas can be thought of as a summary representation that contains information from multiple experiences while eliminating episode-unique dimensions. Note that unlike other forms of dimensionality reduction that aim to preserve information, here schemas ideally remove idiosyncratic episodic information to allow generalization. To learn such reduced representations, one option is that schema learning involves simply averaging across features in all dimensions across episodes, such that dimensions that repeat persist, whereas features that change average out. Alternatively, schema learning might involve goal-sensitive dimensionality reduction, whereby dimensions that include repeating goal-relevant features (such as meal prices on a menu) are prioritized, whereas unique episodic features in goal-irrelevant dimensions (such as the colour of the server's shirt) are downweighted[128] (Box 1).

In RL, an optimal representation of a state focuses on only goal-relevant information in the environment[6,129,130]. The process by which an agent learns what dimensions of the environment are important to a given task has been termed 'representation learning'[125], and often involves dimensionality reduction. The idea is that through experience, an agent learns what dimensions of the environment are relevant to their goals and therefore should be attended to, as well as what dimensions are irrelevant and thus should be ignored. Indeed, learning the relevant (reward-predicting) dimensions of a state guides attention to these dimensions, which in turn prioritizes learning predictions associated with the reward-predicting dimensions[125,131–137]. These studies suggest that goal relevance and selective attention might mediate dimensionality reduction during schema learning.

However, repetition of features might result in learning of goal-irrelevant dimensions as well. Indeed, people are faster to identify a stimulus that appears in a location where, in other task trials, regularities existed in a stream of symbols[138–140], suggesting that they are attending to that location despite it being goal irrelevant. Similarly, across various behavioural tasks, processing of item pairs that are semantically congruent and encountered repeatedly in daily life (such as restaurant and menu) is typically enhanced (reduced reaction times and increased accuracy) compared with incongruent pairs that are rarely encountered (such as spinach and train), even if congruency is task irrelevant[141] and item pairs are presented only briefly[142,143]. Task-irrelevant congruence also enhances long-term memory[144], but perhaps not for unique episodic details[145]. In addition, studies of statistical learning have shown that predicting a goal-irrelevant but repeating dimension comes at the expense of later memory of unique episodic details[56,57,146]. Together, these findings suggest that attentional mechanisms might prioritize learning of repeating information, goal relevant or irrelevant, potentially at the expense of downweighting unique episodic details, as we propose is key for schema learning. This might be adaptive because it allows flexible behaviour when the world changes[2,132,147] (similar to cognitive maps; Box 3). For example, learning that in restaurants the cashier is typically next to the bar — even if mostly irrelevant because payment is typically made at the table with a waiter — can be useful if one is ever in a rush to leave and must pay at the cashier.

Whereas dimensionality reduction can be seen as determining what dimensions will be included in a schema, state abstraction refers to learning features within a dimension that generalize across similar episodes[148,149] (Box 1). For example, in the price dimension, a schema of a restaurant might represent that appetizers are generally cheaper than entrees, but abstract away specific prices that could change over time and depend on the restaurant. Various computations have been hypothesized to govern state abstraction and this is a topic of ongoing investigation (for example, see refs. 150,151). We hypothesize that prediction error learning could be one mechanism for learning the abstracted features themselves (see ref. 152 and Fig. 1c). We note that both dimensionality reduction and state abstraction can be viewed through a similar lens, depending on the specifics of the representation and the implementation.

Medial PFC involvement in schemas and reinforcement learning

We now turn to mapping the ideas laid out above to potential neural substrates. Substantial evidence supports that the orbitofrontal cortex (OFC) and mPFC are involved in both RL and schema-related processes; however, the functions these cortical regions have are intensely debated (Box 4 and also see refs. 5,8,10,16,59,153–156). Here, we focus on the medial part of the OFC (mOFC), ventromedial PFC (vmPFC) and the mid-mPFC (the area dorsal to the mOFC–vmPFC on the medial wall, but ventral to the most dorsal part of the mPFC), using ‘mPFC’ to collectively refer to these areas. We first summarize evidence that the mPFC represents both schemas and RL states. We then offer that low-dimensional representations in the mPFC that can activate detailed memories in posterior brain regions might underlie these representations[16]. Last, we postulate that the amount of dimensionality reduction in the mPFC determines the involvement of subparts along its ventral–lateral and anterior–posterior axis.

mPFC representations of both schemas and reinforcement learning states

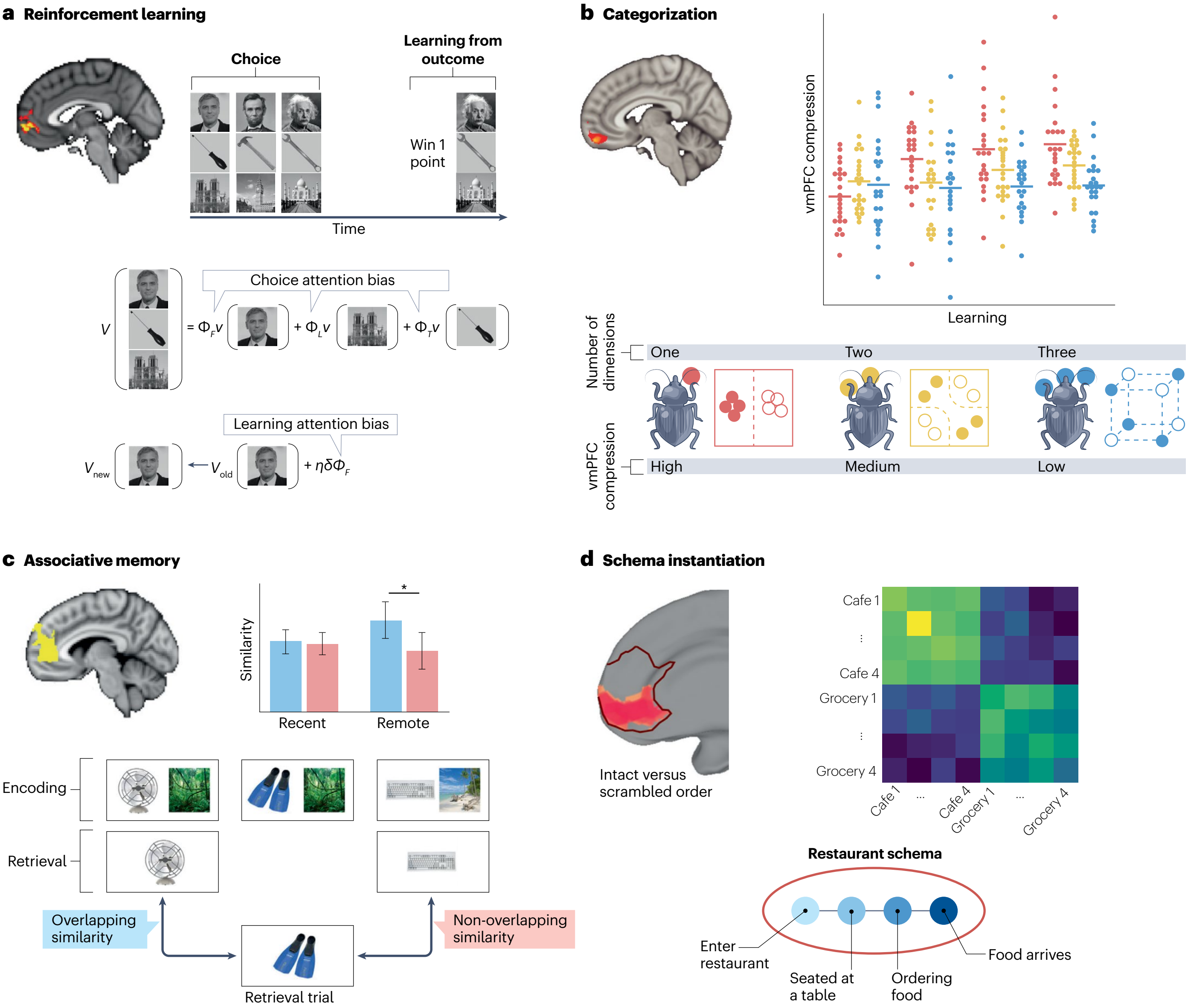
A prominent theory suggests that the mPFC and OFC represent a map of task states in RL[8,10,157–160] (but see Box 4 for alternative theories). Indeed, recent work found mOFC–vmPFC representations

Box 4 | The controversy over the role of the OFC in learning and decision making

We have proposed a role for the orbitofrontal cortex (OFC) in schema learning and instantiation. An alternative prominent view suggests that the OFC represents economic value, which is the expected reward for different options in a task[154,155,256]. This thesis has been supported by a variety of studies across species showing that OFC neurons encode economic value during decision making[300–302], neuroimaging studies in humans showing OFC activation across decision tasks[250,303,304] and studies across species showing lesions to the OFC impair value-based decision making[305–308].

However, the hallmark deficit due to OFC lesions are impairments in reversing previously learned associations, that is, a slowness in changing the preferred option when reward contingencies reverse. In these tasks, lesions to the OFC in rodents and primates spare initial learning, while consistently slowing the shift of option preference after reversals[309–311]. As loss of encoding of the value of the different options should have impaired initial learning as well, this finding has been taken to suggest a role for OFC in inhibitory control in particular[312]. An alternative view suggests that these findings, and many others (including ones attributed to representation of economic value), are consistent with the OFC representing latent states or schemas. In this view, slower reversal in the absence of a functioning OFC is attributed to trial-and-error relearning of the now-correct action rather than the more rapid shift between latent states (or schemas) representing different action policies that are appropriate for different contexts (pre- and post-reversal)[10].

One resolution of this controversy suggests that neurons in the OFC show representation of economic value in electrophysiological studies in animals in tasks in which the expected reward for different options is a crucial part of the latent state representation. However, when reward and states are dissociated, or when the task does not include rewards (and therefore economic value) at all, findings across species suggest that neural activity in the OFC, and potentially mostly the mOFC–vmPFC (an area encompassing the medial part of the OFC and the ventromedial prefrontal cortex, see main text), represents latent states or schemas[8,9,160,248,313,314]. Another possibility is that the overall level of activation or rate of neural firing in the OFC represents economic value, whereas different subpopulations of OFC neurons represent latent states[249,258,315] (see mPFC representations of both schemas and RL states).



of task states and the relationships between them for conceptual spaces[161,162], sequential structures[9,163–165] and social knowledge[166]. Some theories suggest that these mOFC–vmPFC representations are particularly important when states cannot be determined on the basis of perceptual input alone but are latent (such as latent causes; see Hierarchical RL and latent cause inference may contribute to learning and instantiating schema hierarchies) and require the retrieval of information from memory[10,157]. Empirically, multivoxel activity patterns in mOFC–vmPFC are consistent with Bayesian inference of the current (latent) state when this inference requires integrating retrieved prior memories and current observations[7]. Another study successfully classified task-state representations from the mOFC–vmPFC that included information from the current and the previous trial, and thus relied on memory[9]. The mid-mPFC also mediates the retrieval and recombination of memories needed to make choices about novel stimuli[167,168].

Schemas hold knowledge of what typically occurs in an event, and therefore schema instantiation requires retrieving information from memory. Consistent with the mPFC representing latent task states that rely on memory, it is also involved in mediating schemas[5,15]. For instance, lesions to the mOFC–vmPFC impaired the appropriate deployment of schema knowledge[14,169–171]. Moreover, recent studies showed that multivoxel activity patterns in the mid-mPFC were more similar for events that belonged to the same schema (such as visiting different restaurants) compared with different schemas (such as visiting a restaurant versus an airport). This was true even when computing similarity across video and audio stimuli, which suggested a schema representation beyond perceptual features in the mid-mPFC[12,172–174]. Together with the finding that mPFC representations of task states follow Bayesian inference[7], these findings strengthen the proposal that schemas are instantiated via Bayesian latent cause inference[16].

Fig. 2 | Dimensionality reduction in the mPFC. Empirical results from human functional MRI studies of four disparate paradigms support our proposal that the medial prefrontal cortex (mPFC) functions to reduce dimensionality in schemas. **a**, In a reinforcement learning task (top right), participants learned that one of three category dimensions (faces, *F*; landmarks, *L*; tools, *T*) is relevant for obtaining reward. The model that best explained participants' behaviour biased attention towards that category both during choice (middle) and during learning from the outcome (bottom), suggesting participants had a dimensionality-reduced representation of the reinforcement learning task. In this model, the expected reward for each choice (V) was calculated by weighting the value of each component image (v) by the attention to its category (Φ). After obtaining reward, the value of component images was updated according to the prediction error, δ, reflecting the difference between expected and obtained reward, scaled by a learning rate η and the same attention weight Φ). Activation (functional MRI blood oxygen level-dependent response) in the mid-mPFC (top left) correlated with predicted rewards, as estimated by that dimensionality-reduced representation. **b**, In a categorization task, participants categorized bugs on the basis of one, two or three dimensions (antennae size, leg size and mandible shape). Ventromedial prefrontal cortex (vmPFC) high-dimensional multivoxel activity patterns were measured during categorization and orthogonal components explaining variance in multivoxel activity patterns were extracted. Dimensionality reduction ('compression') was quantified as the number of components that explained 90% of the variance in vmPFC activity patterns (with fewer components interpreted as stronger compression). As participants learned the categories over time, simpler categorizations had stronger compression (plot), suggesting that dimensionality reduction in vmPFC tracked the dimensions of the categories. **c**, In an associative memory task, participants first encoded associations between trial-unique objects and several shared scenes. During retrieval trials, objects that had appeared during encoding with the same scene (overlapping) had greater similarity in their neural representations in mid-mPFC compared with those that appeared with different scenes (non-overlapping). This similarity only emerged following a period of consolidation ('remote'), suggesting that during consolidation, specific episodes (trial-unique objects) became grouped on the basis of a shared feature (the scene), whereas details about the items within each episode were reduced. **d**, In a schema instantiation task, participants watched movie clips showing different instances of schemas (such as a schema of a restaurant, bottom). mPFC representations generalized across different instances of the same schemas (such as all cafe clips, right), as indicated by increased similarity within schemas compared with across schemas. This suggested that the mPFC had reduced dimensions and a lack of specific details for each schema instance viewed. Similar mPFC representations were also found across visual versus auditory modalities, but not when the order of events was compromised (left, the red reflects the *z*-score of the contrast, comparing intact versus scrambled order activation), which suggested that although the modality dimension is reduced in mPFC schema representations, the sequential information is preserved. Part **a** is adapted with permission from ref. 134, Elsevier. Part **b** is adapted from ref. 175, Springer Nature Limited. Part **c** is adapted with permission from ref. 83, Elsevier. Part **d** is adapted from ref. 12, CC BY 4.0 (left and bottom) and ref. 173, Springer Nature Limited (right).

The mPFC might represent schemas and reinforcement learning states through dimensionality reduction

Studies show dimensionality reduction in the mPFC across paradigms[12,173] (Fig. 2). For instance, in a RL task, mid-mPFC activation correlated with predicted rewards computed on the basis of attending to one relevant task dimension out of three available[134] (Fig. 2a). More directly, Mack et al.[175] used a categorization task and extracted the number of orthogonal components that accounted for variance in mOFC–vmPFC multivoxel activity patterns. The results showed more compression (namely, fewer components required to explain variance) through learning, especially for categorizations that required fewer dimensions (Fig. 2b). Other studies exposed participants to item–scene associations, with some items sharing the same scene[83,172,176]. After a period of consolidation, the neural representations of items that shared the same scene, but not different scenes, showed stronger similarity to each other in the mid-mPFC, as if specific episodes (each item–scene pair) had been grouped on the basis of a shared feature (the scene) and details about the items within each episode were downweighted (Fig. 2c). Studies grouping episodes on the basis of similar attentional goals obtained similar results[177]. Last, studies that showed reduced representation of episodic details in the mPFC also suggest dimensionality reduction during schema instantiation[12,173] (Fig. 2d). Many of these studies showed specificity to the mPFC[9,12,83,175]. This is consistent with the conclusion that the mPFC is performing dimensionality reduction in these RL and schema-related tasks, rather than the alternative explanation that dimensionality reduction is performed elsewhere in the brain and is then communicated to the mPFC.

Lesion studies further demonstrate that the mOFC–vmPFC has a causal role in the appropriate deployment of schemas[5,14,169], although whether dimensionality reduction underlies observed impairments is an open question. Individuals with mOFC–vmPFC lesions are slower to match concepts to a relevant schema and reject concepts that belong to an irrelevant competing schema, and are also more likely to rate concepts from an irrelevant schema as belonging to a currently relevant schema[14,169]. Individuals with mOFC–vmPFC lesions also do not demonstrate the typical memory enhancement induced from schema-related stimuli[171,178]. For instance, whereby neurotypical participants often show false memory of words as having appeared in a list of schematically related words, this false memory effect is reduced in individuals with vmPFC–mOFC lesions[179]. Damage to the mid-mPFC leads to poor performance in the Wisconsin Card Sorting Test[180–183], which requires participants to sort cards that have multiple dimensions (colour, number and shape) based on only one dimension (for example, colour). One hypothesis is that these lesion study findings are due to impaired dimensionality reduction – relevant schema dimensions are not prioritized, leading to behavioural and memory impairments in some cases[14,169,171,178,180] and to fewer memory intrusions in others[179]. However, this dimensionality-reduction hypothesis awaits direct testing in future lesion studies.

Consistent with our proposal that dimensionality reduction prioritizes both goal-relevant and repeating but goal-irrelevant dimensions, evidence suggests that the mPFC represents both types of information[10,157,160,184,185], and represents task structure even when it is goal irrelevant[164,166,185,186]. For example, the mOFC–vmPFC represented a cognitive map of a two-dimensional social hierarchy even in a task that asked participants to make inferences based only on one dimension[166]. In general, mOFC–vmPFC neural representations of items that share a dimension (such as the context of learning) are more similar than representations of items that do not share a dimension (even when participants perform an unrelated task)[187,188]. The mPFC also showed differential activity during encoding of semantically congruent versus incongruent information not only when participants judged congruency[189,190], but also when they judged grammatical correctness of word stimuli (when semantic congruency was goal irrelevant)[191–193].

Temporal order seems to be a consistently important dimension in mPFC representations of schemas. Indeed, scrambling the order of events in a schema disrupted their representation in the

Box 5 | The contribution of cortical and subcortical networks to schemas and states

Brain areas other than the medial prefrontal cortex (mPFC) probably contribute to learning, representing and instantiating both schemas and states. In particular, neural activity in lateral PFC is probably essential to representations of both schemas and states[159]. The lateral PFC is implicated in an array of executive functions[316–318]. Its role in executive function has been established across species and paradigms, including the Wisconsin Card Sorting Test that requires adaptive sorting of cards based on changing rules or dimensions[181,182,319] and tasks that require cognitive control of episodic and semantic memory[208,320–325]. Collectively, one can describe lateral PFC function as selecting and maintaining neural activity patterns that represent goals and the means to achieve them[326,327]. Consistent with this, it has been recently proposed that the lateral PFC reformats task knowledge in schemas and states (represented by neural activity in the medial orbitofrontal cortex–ventromedial PFC and the medial temporal lobe; Box 2) to perform action selection and produce behaviour[159].

The lateral PFC, and specifically the dorsolateral PFC, is a part of the frontostriatal loop that includes the dorsal striatum and is considered a substrate for model-based reinforcement learning (RL)[30] (also called the 'associative' or 'cognitive' loop[328,329]). Some hierarchical RL models propose that within this loop, the dorsolateral PFC represents hierarchical action policies[88] (see Hierarchical RL and latent cause inference may contribute to learning and instantiating schema hierarchies). Others suggest that the lateral PFC performs Bayesian inference to select the current context, which in turn determines the action policy in the striatum[95,121]. Thus, frontostriatal loops, together with midbrain dopaminergic modulation of learning in corticostriatal synapses[330,331] might mediate learning and executing of hierarchical schemas[88,97,98,332,333]. Consistent with our view that salience can trigger event segmentation and the creation of subgoals for hierarchical RL and schemas, the dorsal striatum is activated at event boundaries and this activity correlated with later memory of events[197,290]. The striatum, as well as midbrain areas, also respond to prediction errors during learning of semantic knowledge[334], state transitions[44] and novelty more broadly, even in the absence of explicit rewards[106,111,335]. Other key regions are the hippocampus and its prefrontal interactions[5,15,17,208] (Box 2), as well as regions in the temporal and parietal lobes[5].

mPFC[12] (Fig. 2d). In rodents, lesions to the mPFC impaired temporal memory[194]. In humans, mOFC–vmPFC lesions (specifically, BA25, subcallosal vmPFC) impaired schema knowledge, while sparing category knowledge[14]; arguably, the temporal order of events is a critical aspect of the former, but not the latter. Studies in both humans and rodents found representation of sequential order in the mPFC (see refs. 195–197). These representations might be supported by strong anatomical connections from mPFC to the hippocampus[198–201] – widely thought to represent temporal and sequential information[122,202–205] – as mPFC–hippocampal functional connectivity supports learning and memory of sequential information[206,207]. Such representations of temporal order are consistent with representation of schemas because schemas contain information about the sequence of typical occurrences in an event, just as state-transition probabilities in model-based RL reflect a sequential transition from one state to another (Fig. 1 and Box 1).

The different functions of the mPFC versus the hippocampus in supporting temporal information included in schemas is not yet clear[17,112,122,208]. One general idea is that the hippocampus is crucial for initial encoding of memories (Box 2), and that as memories become schematized, they become more dependent on the mPFC and independent of the hippocampus[59,60,62]. Although this shift from hippocampal to mPFC encoding may occur for temporal information as well, studies show the hippocampus represents temporal and sequential information both for novel episodes and for well-learned regularities[122,207,209,210]. Thus, another suggestion is that the mPFC represents sequential order memory, whereas the hippocampus represents continuous and gradually changing temporal context[211]. Another not mutually exclusive idea is that mPFC representations of schemas provide dimensions for retrieval of temporal memory of specific episodes represented by the hippocampus[212], as well as other brain areas[174,213].

In sum, evidence across paradigms suggests that both in schemas and in RL the mPFC might mediate dimensionality reduction. However, questions remain regarding whether this is a causal role, and what principles guide dimensionality reduction in the mPFC.

Gradients of dimensionality reduction and memory reactivation along mPFC axes

One potential role of representations of schemas and states in the mPFC is to guide the retrieval of knowledge via memory reactivation in posterior brain regions[5,16,132]. For instance, mPFC neural activity precedes hippocampal and ventral temporal lobe activity during memory retrieval[214–217]. Mid-mPFC activity also correlates with the persistence (across time) of ventral temporal lobe and hippocampal representations of items experienced within the same context[218]. More directly, a recent study showed that the extent of the representation of a schema in the mPFC correlated with the strength of the posterior medial cortex representation of specific events consistent with that schema (for example, a specific visit to a restaurant)[174]. Lesion studies demonstrate causality: in a rodent reversal-learning task that required resolving interference to infer the correct state, mPFC lesions impaired hippocampal representations that mediated interference resolution[219]. In humans, mOFC–vmPFC lesions impair the evaluation of retrieved memories[220,221], which can result in confabulation – retrieval of memories that are irrelevant to a specific context or schema[170,222]. In addition, the mPFC might route the involvement of cortical versus hippocampal systems based on how memories relate to the current schema instantiation[5,15]: connectivity with posterior cortical regions mediates memory of schema-consistent information, whereas connectivity with the hippocampus mediates memory of schema-inconsistent information[13,15,17,189,190,223–227].

These memory reactivation findings were reported in different loci in the mPFC. In this Perspective, we hypothesize that the level of dimensionality reduction, or the degree to which memories are schematic and lack specific details, might underlie the gradual involvement of subregions along the anterior to posterior and the ventral to dorsal axes of the mPFC (regarding the medial to lateral axis, the lateral OFC has been extensively discussed elsewhere as representing states based on observable information[10,156,184], and we discuss the lateral PFC in Box 5). Our proposal is motivated by gradual changes in the anatomical structure and connectivity along the mPFC. There is a wide agreement on a gradual transition from agranular to granular cortex along the posterior–anterior axis of the mPFC[228–230]. Along the ventral–dorsal axis, studies in humans and monkeys generally show different connectivity profile of mid-mPFC versus mOFC–vmPFC[198–201] (more below), and a recent study found that these changes are gradual[231].

We therefore propose that abstract representations might recruit more anterior parts of the mPFC, whereas detailed memories might recruit the posterior mPFC and its connectivity with the hippocampus. Research suggests that the anterior part of the PFC is involved in representing future or counterfactual states and actions, but not current ones[232–236]. Studies on prospective planning and predictions also show a gradient of predictions in mPFC, whereby predictions of the far future are represented more anteriorly and predictions of the near future are represented more posteriorly[72,74]. Potentially, the farther one prospects to the future, the more abstract and less concrete and detailed are one's thoughts[237–239], and therefore they are represented more anteriorly. This might also be true for counterfactual thoughts compared with actions and events that have materialized. Last, functional MRI studies of dimensionality reduction in abstract tasks report a more anterior cluster of voxels in the mOFC–vmPFC[134,175] compared with studies addressing retrieval of autobiographic memories of specific events[240]. This posterior versus anterior localization in the mPFC aligns with the connectivity of the more posterior (and ventral) part of the mPFC and the hippocampus[198–201], which is crucial for the encoding and retrieval of detailed memories, potentially through the allocation of distinct representations that serve to disambiguate similar stimuli[241–245] and states[246–248].

Specificity versus abstraction might also underlie graded involvement from ventral to dorsal mPFC, supported by differences in functional connectivity. For instance, more ventral parts of the mPFC have been shown to represent values of specific stimuli (for example, the value of a chocolate outcome) rather than more generic (scalar) value representations that are invariant to the specific stimulus (for example, when two different outcomes are equally valued)[249,250]. Further, although retrieval of specific autobiographical memories tends to involve a ventral mPFC cluster of voxels[240], a study addressing rule learning that required abstraction across multiple episodes showed mid-mPFC activation[251]. Whereas the mOFC–vmPFC is connected to the hippocampus, which is important for detailed memories, the mid-mPFC is connected with the posterior medial cortex[201,231,252], which represents events over large timescales by potentially abstracting away more specific details[174,253–255]. Of note, the studies mentioned here employed a variety of learning protocols and stimuli, so they are not all directly comparable. Nevertheless, they are in line with our proposal that the extent of dimensionality reduction underlies differential involvement of mPFC subregions.

Although our hypothesis is consistent with existing data, we note that other proposals exist for the function of the mPFC, including representing economic value[8,155,256], the evaluation of retrieved memories[221], confidence in value estimates[257] or retrieved memories[221], or signalling the congruency of perceptual information with the current schema[15] (Box 4). In our view, recent studies that examined multivoxel activity patterns support the hypothesis that the mPFC represents states or schemas because these studies show different neural representations for stimuli or memories associated with different states or schemas, even when these had the same economic value, were retrieved with similar levels of confidence or were similarly congruent with the current schema[9,83,174,249,258]. Thus, different representations may be multiplexed in the mPFC: different populations of neurons might represent different states and schemas that, in turn, activate different neuron populations in posterior brain regions and lead to the retrieval of different memories[15,59], whereas the overall level of activity in mPFC neurons might signal value, congruency with existing schemas or other monitoring signals.

Conclusion and future directions

In this Perspective, we outlined how RL and event schemas might be related. We proposed that schemas might be learned via RL-related mechanisms such as prediction errors, hierarchical decomposition of tasks and dimensionality reduction. We then hypothesized that dimensionality reduction might underlie the involvement of the mPFC in both schemas and RL, and postulated that the extent of abstraction might determine the locus of involvement along anatomical mPFC axes. Although this hypothesis is consistent with the literature, it needs to be further tested and compared with alternative proposals for the functional role of the mPFC and OFC (Box 4).

Like episodic memories, schemas change through periods of consolidation[58,60,62,259,260]. An open question is how the computations we outlined contribute to these changes. A few recent studies directly tested aspects of this question by exposing subjects to unique locations, and then testing for memory of both the specific locations and schema-like summary statistics of the distribution from which locations were drawn[119,261–264]. In both rodents and humans, these tests showed that episodic memories of specific locations deteriorate over time. In rodents, the formation of schemas (measured through memory of the location distribution) required the passage of time[262], whereas in humans, schemas formed rapidly, even before memory consolidation, and then remained stable[119] or also deteriorated[263]

Glossary

Action policy
A probability distribution across possible actions in a state.

Agent
An entity (human or artificial) that learns and behaves in the environment.

Dimensionality reduction
The projection of data into a lower-dimensional subspace, in which dimensions reflect a weighted combination of the original dimensions.

Event boundaries
Changes in situations or salient occurrences that cause the segmentation of experience into discrete event representations during perception and in memory.

Latent cause inference
A computational theory of how observations are grouped into clusters ('latent causes') according to their similarity.

Multivoxel activity patterns
Blood oxygen level-dependent activity patterns in functional MRI that span across multiple voxels (three-dimensional units of measurement in functional MRI) thought to reflect the neural representations of stimuli, states, contexts and so on.

State
In reinforcement learning, a state is a summary of features that describe the current situation.

State prediction error
A signal that accompanies a transition to a state that is different than expected.

State-transition probability
The probability of transitioning from one state to another.

over time. An exciting future direction could be to combine such paradigms with tasks that include hierarchical representations and action sequences[71,97,121]. Computational models of these tasks would allow assessing the specific mechanisms we proposed here for schema learning and instantiation, both early in learning (Box 2) and through consolidation.

Although RL algorithms generally suffer from a curse of dimensionality, namely they scale poorly to rich high-dimensional everyday life events[6,125], recent RL models combined hierarchical RL principles and neural network models to learn multidimensional environments and successfully completed very complex tasks (see refs. 265–267). Whether such advanced RL algorithms are used by humans to learn schemas of multidimensional and complex environments is a question for future research[268]. We propose that ideas and findings from schema research, such as those we have laid out in this Perspective, could potentially inform further development of these models. For example, exciting work now combines neural networks, latent cause inference and RL to explain schemas and event segmentation of complex environments in humans[22,269].

In addition to these and other outstanding questions we outlined throughout this Perspective, key questions for future research include what is the organizing principle of mPFC function, whether it indeed performs dimensionality reduction and what are the computations that brain-wide networks perform in learning hierarchical schemas and states (Box 5). Broadly, rather than trying to map neural substrates (regions or networks) to cognitive constructs, the approach we promoted in this Perspective, and that has been proposed many times before (see refs. 270–272), identifies computations (such as dimensionality reduction) that serve putative cognitive constructs such as schemas or RL (but that can also serve additional cognitive functions; Box 4) and asks how these are implemented by different brain areas. The road from the brain to cognition and behaviour thus goes through computation.

References

1. Alba, J. W. & Hasher, L. Is memory schematic? *Psychol. Bull.* **93**, 203–231 (1983).
2. Ghosh, V. E. & Gilboa, A. What is a memory schema? A historical perspective on current neuroscience literature. *Neuropsychologia* **53**, 104–114 (2014).
3. Schank, R. & Abelson, R. in *Scripts Plans Goals and Understanding: an Inquiry into Human Knowledge Structures* 36–68 (Lawrence Erlbaum, 1977).
A classic paper offering a philosophical investigation into human knowledge organization, goals, and event schemas ('scripts').
4. Bartlett, F. C. *Remembering: a Study in Experimental and Social Psychology* (Cambridge Univ. Press, 1932).
One of the earliest demonstrations not only that memories are reconstructed, biased and transformed by schemas, but also of the notion that memories are not fixed copies of the past but rather are malleable and are influenced by expectations and social norms.
5. Gilboa, A. & Marlatte, H. Neurobiology of schemas and schema-mediated memory. *Trends Cogn. Sci.* **21**, 618–631 (2017).
This article provides one of the only overviews of brain regions and networks involved in schema learning and instantiation and the influence of schemas on memory.
6. Sutton, R. S. & Barto, A. G. *Reinforcement Learning: an Introduction* (MIT, 2018).
7. Chan, S. C. Y., Niv, Y. & Norman, K. A. A probability distribution over latent causes, in the orbitofrontal cortex. *J. Neurosci.* **36**, 7817–7828 (2016).
8. Klein-Flügge, M. C., Bongioanni, A. & Rushworth, M. F. S. Medial and orbital frontal cortex in decision-making and flexible behavior. *Neuron* https://doi.org/10.1016/j.neuron.2022.05.022 (2022).
9. Schuck, N. W. et al. Human orbitofrontal cortex represents a cognitive map of state space article human orbitofrontal cortex represents a cognitive map of state space. *Neuron* **91**, 1402–1412 (2016).
This study provides a compelling demonstration of a neural representation of complex task structure in the mOFC–vmPFC in a task with no rewards, thus obviating a variety of alternative explanations of mOFC–vmPFC activity.
10. Wilson, R. C., Takahashi, Y. K., Schoenbaum, G. & Niv, Y. Orbitofrontal cortex as a cognitive map of task space. *Neuron* **81**, 267–279 (2014).
11. Zhou, J. et al. Evolving schema representations in orbitofrontal ensembles during learning. *Nature* **590**, 606–611 (2020).
Using single-unit recordings during a sequential task, the authors show how neural representations of schemas evolve in the rodent OFC, including demonstration of dimensionality reduction, generalization of neural representations across instances and schema-dependent enhanced learning.
12. Baldassano, C., Hasson, U. & Norman, K. A. Representation of real-world event schemas during narrative perception. *J. Neurosci.* **38**, 9689–9699 (2018).
This study uses sophisticated hidden Markov models to segment events in functional MRI data from participants watching dynamic movie stimuli and demonstrates, in the human mPFC, abstract representations of schemas that include information about the order of events.
13. Bonasia, K. et al. Prior knowledge modulates the neural substrates of encoding and retrieving naturalistic events at short and long delays. *Neurobiol. Learn. Mem.* **153**, 26–39 (2018).
14. Giuliano, A. E., Bonasia, K., Ghosh, V. E., Moscovitch, M. & Gilboa, A. Differential influence of ventromedial prefrontal cortex lesions on neural representations of schema and semantic category knowledge. *J. Cogn. Neurosci*. https://doi.org/10.1162/jocn_a_01746 (2021).
A unique study showing different impairments in deploying schema versus category knowledge in individuals with mOFC–vmPFC lesions.
15. van Kesteren, M. T. R., Ruiter, D. J., Fernández, G. & Henson, R. N. How schema and novelty augment memory formation. *Trends Neurosci.* **35**, 211–219 (2012).
16. Varga, N., Morton, N. & Preston, A. in *The Oxford Handbook of Human Memory* (eds Kahana, M. J. & Wagner, A. D.) (Oxford Univ. Press, 2022).
17. Preston, A. R. & Eichenbaum, H. Interplay of hippocampus and prefrontal cortex in memory. *Curr. Biol.* **23**, R764–R773 (2013).
A perspective on hippocampal and prefrontal involvement in memory encoding, consolidation, and retrieval as a function of schematization and memory integration.
18. Rumelhart, D. E. & Ortony, A. in *Schooling and the Acquisition of Knowledge* 99–135 https://doi.org/10.4324/9781315271644-10 (1977).
19. Piaget, J. *The Origins of Intelligence in Children* (International Univ. Press, 1952).
20. Zacks, J. M. Event perception and memory. *Annu. Rev. Psychol.* **71**, 165–191 (2020).
21. Elman, J. L. & McRae, K. A model of event knowledge. *Psychol. Rev.* **126**, 252–291 (2019).
22. Franklin, N. T., Norman, K. A., Ranganath, C., Zacks, J. M. & Gershman, S. J. Structured event memory: a neuro-symbolic model of event cognition. *Psychol. Rev.* **127**, 327–361 (2020).
A computational model accounting for event segmentation findings in humans using latent cause inference and event schemas.
23. Collins, A. M. & Loftus, E. F. Spreading activation theory of semantic processing. *Psychol. Rev.* **82**, 407–428 (1975).
24. Collins, A. M. & Quillian, M. R. Retrieval time from semantic memory. *J. Verbal Learn. Verbal Behav.* **8**, 240–247 (1969).
25. Kenett, Y. N., Levi, E., Anaki, D. & Faust, M. The semantic distance task: quantifying semantic distance with semantic network path length. *J. Exp. Psychol. Learn. Mem. Cogn.* https://doi.org/10.1037/xlm0000391 (2017).
26. Kumar, A. A. Semantic memory: a review of methods, models, and current challenges. *Psychon. Bull. Rev*. **28**, 40–80 (2021).
27. Love, B. C., Medin, D. L. & Gureckis, T. M. SUSTAIN: a network model of category learning. *Psychol. Rev.* **111**, 309–332 (2004).
28. McClelland, J. L., McNaughton, B. L. & Oreilly, R. C. Why there are complementary learning-systems in the hippocampus and neocortex—insights from the success and failures of connectionist models of learning and memory. *Psychol. Rev.* **102**, 419–457 (1995).
29. Murphy, G. L. *The Big Book of Concepts* (MIT, 2004).
30. Daw, N. D., Niv, Y. & Dayan, P. Uncertainty-based competition between prefrontal and dorsolateral striatal systems for behavioral control. *Nat. Neurosci.* **8**, 1704–1711 (2005).
31. Daw, N. D., Gershman, S. J., Seymour, B., Dayan, P. & Dolan, R. J. Model-based influences on humans' choices and striatal prediction errors. *Neuron* **69**, 1204–1215 (2011).
32. Doya, K., Samejima, K., Katagiri, K. & Kawato, M. Multiple model-based reinforcement learning. *Neural Comput.* **1369**, 1347–1369 (2002).
33. Niv, Y. & Schoenbaum, G. Dialogues on prediction errors. *Trends Cogn. Sci.* **12**, 265–272 (2008).
34. Rescorla, R. A. Pavlovian conditioning. It's not what you think it is. *Am. Psychol.* **43**, 151–160 (1988).
35. Rescorla, R. A. & Wagner, A. R. in Classical Conditioning II: Current Research and Theory (eds A. H. Black & W. F. Prokasy), Ch. 3, 64–99 (Appleton-Century-Crofts 1972).
36. Kamin, L. J. Predictability, surprise, attention, and conditioning. *Symposium on Punishment* (Princeton University 1967).
37. Kamin, L. J. in *Miami Symposium on the Prediction of Behavior, 1967: Aversive Stimulation* (ed. Jones, M. R.) 9–31 (Univ. Miami Press, 1968).
A seminal paper describing the discovery of 'blocking', which prompted the idea that associative learning is driven by prediction errors, not simple co-occurrence of events.
38. d'Acremont, M., Schultz, W. & Bossaerts, P. The human brain encodes event frequencies while forming subjective beliefs. *J. Neurosci.* **33**, 10887–10897 (2013).
39. Niv, Y. & Langdon, A. Reinforcement learning with Marr. *Curr. Opin. Behav. Sci.* **11**, 67–73 (2016).
40. Ergo, K., De Loof, E. & Verguts, T. Reward prediction error and declarative memory. *Trends Cogn. Sci.* **24**, 388–397 (2020).

41. Greve, A., Cooper, E., Tibon, R. & Henson, R. N. Knowledge is power: prior knowledge aids memory for both congruent and incongruent events, but in different ways. *J. Exp. Psychol. Gen.* **148**, 325–341 (2019).
42. Rouhani, N., Norman, K. A. & Niv, Y. Dissociable effects of surprising rewards on learning and memory. *J. Exp. Psychol. Learn. Mem. Cogn.* **44**, 1430–1443 (2018).
43. Rouhani, N. & Niv, Y. Signed and unsigned reward prediction errors dynamically enhance learning and memory. *eLife* **10**, e61077 (2021).
44. Sharpe, M. J. et al. Dopamine transients are sufficient and necessary for acquisition of model-based associations. *Nat. Neurosci.* **20**, 735–742 (2017).
45. Sharpe, M. J. et al. Dopamine transients do not act as model-free prediction errors during associative learning. *Nat. Commun.* **11**, 106 (2020).
This study demonstrates in rats that dopaminergic prediction errors are sufficient and necessary for learning state transitions even in the absence of a reward, contrary to previous assumptions that dopamine represents reward prediction errors used only for learning the reward value of stimuli.
46. Momennejad, I. et al. The successor representation in human reinforcement learning. *Nat. Hum. Behav.* **1**, 680–692 (2017).
47. Sharp, P. B., Dolan, R. J. & Eldar, E. Disrupted state transition learning as a computational marker of compulsivity. *Psychol. Med.* https://doi.org/10.1017/S0033291721003846 (2021) .
48. Sharp, P. B., Russek, E. M., Huys, Q. J., Dolan, R. J. & Eldar, E. Humans perseverate on punishment avoidance goals in multigoal reinforcement learning. *eLife* **11**, e74402 (2022).
49. Akam, T. et al. The anterior cingulate cortex predicts future states to mediate model-based action selection. *Neuron* **109**, 149–163.e7 (2021).
This study uniquely demonstrates, in mice, that the anterior cingulate cortex encodes state prediction errors and is necessary for using these for model-based reinforcement learning.
50. Bartolo, R. & Averbeck, B. B. Prefrontal cortex predicts state switches during reversal learning. *Neuron* **106**, 1044–1054.e4 (2020).
51. Bein, O., Plotkin, N. A. & Davachi, L. Mnemonic prediction errors promote detailed memories. *Learn. Mem.* **28**, 422–434 (2021).
52. Greve, A., Cooper, E., Kaula, A., Anderson, M. C. & Henson, R. Does prediction error drive one-shot declarative learning? *J. Mem. Lang.* **94**, 149–165 (2017).
53. Kafkas, A. & Montaldi, D. Expectation affects learning and modulates memory experience at retrieval. *Cognition* **180**, 123–134 (2018).
One of the first behavioural demonstrations that state prediction errors enhance memory in the absence of rewards.
54. Henson, R. N. & Gagnepain, P. Predictive, interactive multiple memory systems. *Hippocampus* **20**, 1315–1326 (2010).
55. Quent, J. A., Henson, R. N. & Greve, A. A predictive account of how novelty influences declarative memory. *Neurobiol. Learn. Mem.* **179**, 107382 (2021).
56. Kim, G., Lewis-Peacock, J. A., Norman, K. A. & Turk-Browne, N. B. Pruning of memories by context-based prediction error. *Proc. Natl Acad. Sci. USA* **111**, 8997–9002 (2014).
This study uses functional MRI multivoxel pattern classification to estimate state predictions in the brain and shows that state prediction errors correlate with forgetting of states that cued erroneous predictions, consistent with update of the internal model.
57. Kim, H., Schlichting, M. L., Preston, A. R. & Lewis-Peacock, J. A. Predictability changes what we remember in familiar temporal contexts. *J. Cogn. Neurosci.* **32**, 124–140 (2020).
58. Antony, J. W. & Schapiro, A. C. Active and effective replay: systems consolidation reconsidered again. *Nat. Rev. Neurosci.* **20**, 506–507 (2019).
59. Frankland, P. W. & Bontempi, B. The organization of recent and remote memories. *Nat. Rev. Neurosci.* **6**, 119–130 (2005).
60. Gilboa, A. & Moscovitch, M. No consolidation without representation: correspondence between neural and psychological representations in recent and remote memory. *Neuron* **109**, 2239–2255 (2021).
An insightful perspective arguing that a memory of an event is retained in multiple forms that differentially rely on hippocampal and cortical structures, and that their expression is a function of factors such as time, schematization and task demands, rather than system level consolidation per se.
61. Moscovitch, M., Cabeza, R., Winocur, G. & Nadel, L. Episodic memory and beyond: the hippocampus and neocortex in transformation. *Annu. Rev. Psychol.* **67**, 105–134 (2016).
62. Squire, L. R. & Alvarez, P. Retrograde amnesia and memory consolidation: a neurobiological perspective. *Curr. Opin. Neurobiol.* **5**, 169–177 (1995).
63. Anderson, J. R. & Reder, L. M. The fan effect: new results and new theories. *J. Exp. Psychol. Gen.* **128**, 186–197 (1999).
64. Bein, O., Trzewik, M. & Maril, A. The role of prior knowledge in incremental associative learning: an empirical and computational approach. *J. Mem. Lang.* **107**, 1–24 (2019).
65. Bein, O., Gasser, C., Amer, T., Maril, A. & Davachi, L. Predictions transform memories: how expected versus unexpected events are integrated or separated in memory. *Neurosci. Biobehav. Rev.* **153**, 105368 (2023).
66. Bellana, B., Mansour, R., Ladyka-Wojcik, N., Grady, C. L. & Moscovitch, M. The influence of prior knowledge on the formation of detailed and durable memories. *J. Mem. Lang.* **121**, 104264 (2021).
67. Hemmer, P. & Steyvers, M. Integrating episodic memories and prior knowledge at multiple levels of abstraction. *Psychon. Bull. Rev.* **16**, 80–87 (2009).
68. Lee, W.-T., Hazeltine, E. & Jiang, J. Interference and integration in hierarchical task learning. *J. Exp. Psychol. Gen.* **151**, 3028–3044 (2022).
69. Reder, L. M., Paynter, C., Diana, R. A., Ngiam, J. & Dickison, D. in *Psychology of Learning and Motivation* (eds Benjamin A. S. & Ross, B. H.) **48**, 271–312 (Elsevier, 2007).
70. Tompary, A. & Thompson-Schill, S. L. Semantic influences on episodic memory distortions. *J. Exp. Psychol. Gen.* https://doi.org/10.1037/xge0001017 (2021).
71. Gasser, C. & Davachi, L. Cross-modal facilitation of episodic memory by sequential action execution. *Psychol. Sci*. https://doi.org/10.1177/09567976231158292 (2023).
72. Brunec, I. K. & Momennejad, I. Predictive representations in hippocampal and prefrontal hierarchies. *J. Neurosci.* **42**, 299–312 (2022).
Using a spatial navigation task, this study provides one of very few demonstrations to date of gradients of predictive representations in the human mPFC.
73. Gravina, M. T. & Sederberg, P. B. The neural architecture of prediction over a continuum of spatiotemporal scales. *Curr. Opin. Behav. Sci.* **17**, 194–202 (2017).
74. Lee, C. S., Aly, M. & Baldassano, C. Anticipation of temporally structured events in the brain. *eLife* **10**, e64972 (2021).
75. Antony, J. W., Van Dam, J., Massey, J. R., Barnett, A. J. & Bennion, K. A. Long-term, multi-event surprise correlates with enhanced autobiographical memory. *Nat. Hum. Behav.* https://doi.org/10.1038/s41562-023-01631-8 (2023).
76. Antony, J. W. et al. Behavioral, physiological, and neural signatures of surprise during naturalistic sports viewing. *Neuron* **109**, 1–14 (2021).
77. Sinclair, A. H. & Barense, M. D. Surprise and destabilize: prediction error influences episodic memory reconsolidation. *Learn. Mem.* **25**, 369–381 (2018).
This study shows, in humans, that prediction errors lead to memory update in a complex and dynamic environment.
78. Sinclair, A. H., Manalili, G. M., Brunec, I. K., Adcock, R. A. & Barense, M. D. Prediction errors disrupt hippocampal representations and update episodic memories. *Proc. Natl Acad. Sci. USA* **118**, e2117625118 (2021).
79. Antony, J. W. et al. Behavioral, physiological, and neural signatures of surprise during naturalistic sports viewing. *Neuron* https://doi.org/10.1016/j.neuron.2020.10.029 (2020).
80. Tulving, E. in *Organization of Memory* (eds Tulving, E. & Donaldson, W.) 381–403 (Academic, 1972).
81. Tulving, E. Episodic memory from mind to brain. *Annu. Rev. Psychol.* **53**, 1–25 (2002).
82. Renoult, L., Irish, M., Moscovitch, M. & Rugg, M. D. From knowing to remembering: the semantic-episodic distinction. *Trends Cogn. Sci.* **23**, 1041–1057 (2019).
83. Tompary, A. & Davachi, L. Consolidation promotes the emergence of representational overlap in the hippocampus and medial prefrontal cortex. *Neuron* **96**, 228–241 (2017).
This human functional MRI study shows schematization of associative memories in the brain, demonstrating potential dimensionality reduction in the mPFC as memories consolidate.
84. Badre, D. Cognitive control, hierarchy, and the rostro-caudal organization of the frontal lobes. *Trends Cogn. Sci.* **12**, 193–200 (2008).
85. Barto, A. G. & Mahadevan, S. Recent advances in hierarchical reinforcement learning. *Discrete Event Dyn. Syst. Theory Appl.* **13**, 41–77 (2003).
86. Botvinick, M. M. Hierarchical models of behavior and prefrontal function. *Trends Cogn. Sci.* **12**, 201–208 (2008).
87. Botvinick, M. M. Hierarchical reinforcement learning and decision making. *Curr. Opin. Neurobiol.* **22**, 956–962 (2012).
88. Botvinick, M. M., Niv, Y. & Barto, A. G. Hierarchically organized behavior and its neural foundations: a reinforcement learning perspective. *Cognition* **113**, 262–280 (2009).
This review offers a computational model for learning options in hierarchical reinforcement learning and how it might be implemented in the brain.
89. Collins, A. G. E. in *Goal-Directed Decision Making* (eds Morris, R., Bornstein, A. & Shenhav, A.) 105–123 (Academic, 2018).
90. Correa, C. G., Ho, M. K., Callaway, F., Daw, N. D. & Griffiths, T. L. Humans decompose tasks by trading off utility and computational cost. *PLoS Comput. Biol.* **19**, e1011087 (2023).
A recent behavioural and computational investigation into how humans use goals and subgoals to construct efficient hierarchical task structures during planning.
91. Tomov, M. S., Yagati, S., Kumar, A., Yang, W. & Gershman, S. J. Discovery of hierarchical representations for efficient planning. *PLoS Comput. Biol.* **16**, e1007594 (2020).
92. Sutton, R. S., Precup, D. & Singh, S. Between MDPs and semi-MDPs: a framework for temporal abstraction in reinforcement learning. *Artif. Intell.* **112**, 181–211 (1999).
93. Holroyd, C. B. & Yeung, N. Motivation of extended behaviors by anterior cingulate cortex. *Trends Cogn. Sci.* **16**, 122–128 (2012).
94. Solway, A. et al. Optimal behavioral hierarchy. *PLoS Comput. Biol.* **10**, e1003779 (2014).
95. Collins, A. G. E. & Frank, M. J. Cognitive control over learning: creating, clustering, and generalizing task-set structure. *Psychol. Rev.* **120**, 190–229 (2013).
96. Liu, R. G. & Frank, M. J. Hierarchical clustering optimizes the tradeoff between compositionality and expressivity of task structures for flexible reinforcement learning. *Artif. Intell.* **312**, 103770 (2022).
97. Diuk, C., Tsai, K., Wallis, J., Botvinick, M. & Niv, Y. Hierarchical learning induces two simultaneous, but separable, prediction errors in human basal ganglia. *J. Neurosci.* **33**, 5797–5805 (2013).
98. Ribas-Fernandes, J. J. F. et al. A neural signature of hierarchical reinforcement learning. *Neuron* **71**, 370–379 (2011).
99. Frost, R., Armstrong, B. C., Siegelman, N. & Christiansen, M. H. Domain generality versus modality specificity: the paradox of statistical learning. *Trends Cogn. Sci.* **19**, 117–125 (2015).
100. Saffran, J. R., Aslin, R. N. & Newport, E. L. Statistical learning by 8-month-old infants. *Science* **274**, 1926–1928 (1996).

101. Saffran, J. R. & Wilson, D. P. From syllables to syntax: multilevel statistical learning by 12-month-old infants. *Infancy* **4**, 273–284 (2003).
102. Schapiro, A. C. & Turk-Browne, N. in *Brain Mapping: an Encyclopedic Reference* (ed. Toga, A. W.), Vol. 3, 501–506 (Elsevier, 2015).
103. Turk-Browne, N. B., Jungé, J. & Scholl, B. J. The automaticity of visual statistical learning. *J. Exp. Psychol. Gen.* **134**, 552–564 (2005).
104. Balaguer, J., Spiers, H., Hassabis, D. & Summerfield, C. Neural mechanisms of hierarchical planning in a virtual subway network. *Neuron* **90**, 893–903 (2016).
105. Bunzeck, N., Dayan, P., Dolan, R. J. & Duzel, E. A common mechanism for adaptive scaling of reward and novelty. *Hum. Brain Mapp.* **31**, 1380–1394 (2010).
106. Bunzeck, N. & Düzel, E. Absolute coding of stimulus novelty in the human substantia nigra/VTA. *Neuron* **51**, 369–379 (2006).
107. Cowan, E. T., Schapiro, A. C., Dunsmoor, J. E. & Murty, V. P. Memory consolidation as an adaptive process. *Psychon. Bull. Rev*. https://doi.org/10.3758/s13423-021-01978-x (2021) .
108. Düzel, E., Bunzeck, N., Guitart-Masip, M. & Düzel, S. Novelty-related motivation of anticipation and exploration by dopamine (NOMAD): implications for healthy aging. *Neurosci. Biobehav. Rev.* **34**, 660–669 (2010).
109. Kamiński, J. et al. Novelty-sensitive dopaminergic neurons in the human substantia nigra predict success of declarative memory formation. *Curr. Biol.* **28**, 1333–1343.e4 (2018).
110. Murty, V. P. & Adcock, R. A. Enriched encoding: reward motivation organizes cortical networks for hippocampal detection of unexpected events. *Cereb. Cortex* **24**, 2160–2168 (2014).
111. Wittmann, B. C., Bunzeck, N., Dolan, R. J. & Düzel, E. Anticipation of novelty recruits reward system and hippocampus while promoting recollection. *Neuroimage* **38**, 194–202 (2007).
112. Clewett, D., DuBrow, S. & Davachi, L. Transcending time in the brain: how event memories are constructed from experience. *Hippocampus* **29**, 162–183 (2019). **A comprehensive article reviewing behavioural and neural studies of temporal memory and event segmentation.**
113. Shin, Y. S. & DuBrow, S. Structuring memory through inference-based event segmentation. *Top. Cogn. Sci.* **13**, 106–127 (2021).
114. Clewett, D., Gasser, C. & Davachi, L. Pupil-linked arousal signals track the temporal organization of events in memory. *Nat. Commun.* **11**, 4007 (2020).
115. DuBrow, S. & Davachi, L. The influence of context boundaries on memory for the sequential order of events. *J. Exp. Psychol. Gen.* **142**, 1277–1286 (2013).
116. Ezzyat, Y. & Davachi, L. Similarity breeds proximity: pattern similarity within and across contexts is related to later mnemonic judgments of temporal proximity. *Neuron* **81**, 1179–1189 (2014).
117. Rouhani, N., Norman, K. A., Niv, Y. & Bornstein, A. M. Reward prediction errors create event boundaries in memory. *Cognition* **203**, 104269 (2020).
118. DuBrow, S., Rouhani, N., Niv, Y. & Norman, K. A. Does mental context drift or shift? *Curr. Opin. Behav. Sci.* **17**, 141–146 (2017).
119. Tompary, A., Zhou, W. & Davachi, L. Schematic memories develop quickly, but are not expressed unless necessary. *Sci. Rep.* **10**, 16968 (2020).
120. Collins, A. G. E. & Frank, M. J. Neural signature of hierarchically structured expectations predicts clustering and transfer of rule sets in reinforcement learning. *Cognition* **152**, 160–169 (2016).
121. Eckstein, M. K. & Collins, A. G. E. Computational evidence for hierarchically structured reinforcement learning in humans. *Proc. Natl Acad. Sci. USA* **117**, 29381–29389 (2020).
122. Davachi, L. & DuBrow, S. How the hippocampus preserves order: the role of prediction and context. *Trends Cogn. Sci.* **19**, 92–99 (2015).
123. Éltető, N., Nemeth, D., Janacsek, K. & Dayan, P. Tracking human skill learning with a hierarchical Bayesian sequence model. *PLoS Comput. Biol.* **18**, e1009866 (2022).
124. Gershman, S. J. & Niv, Y. Learning latent structure: carving nature at its joints. *Curr. Opin. Neurobiol.* **20**, 251–256 (2010).
125. Niv, Y. Learning task-state representations. *Nat. Neurosci.* **22**, 1544–1553 (2019). **A review of the neural and computational underpinnings of representation learning in reinforcement learning, focusing on latent cause inference and dimensionality reduction as methods for learning to represent a task as a series of states that include only task-relevant information.**
126. Gershman, S. J. A unifying probabilistic view of associative learning. *PLoS Comput. Biol.* **11**, e1004567 (2015).
127. Gershman, S. J., Monfils, M.-H., Norman, K. A. & Niv, Y. The computational nature of memory modification. *eLife* **6**, e23763 (2017).
128. Bar-Gad, I., Morris, G. & Bergman, H. Information processing, dimensionality reduction and reinforcement learning in the basal ganglia. *Prog. Neurobiol.* **71**, 439–473 (2003).
129. Schuck, N. W., Wilson, R. & Niv, Y. in *Goal-Directed Decision Making: Computations and Neural Circuits* (eds. Morris, R., Bornstein, A. & Shenhav, A.) 259–278 (Elsevier, 2018).
130. Langdon, A. J., Song, M. & Niv, Y. Uncovering the 'state': tracing the hidden state representations that structure learning and decision-making. *Behav. Process.* **167**, 103891 (2019).
131. Daniel, R., Radulescu, A. & Niv, Y. Intact reinforcement learning but impaired attentional control during multidimensional probabilistic learning in older adults. *J. Neurosci.* **40**, 1084–1096 (2020).
132. De Martino, B. & Cortese, A. Goals, usefulness and abstraction in value-based choice. *Trends Cogn. Sci.* **27**, 65–80 (2023).
133. Farashahi, S., Rowe, K., Aslami, Z., Lee, D. & Soltani, A. Feature-based learning improves adaptability without compromising precision. *Nat. Commun.* **8**, 1768 (2017).
134. Leong, Y. C., Radulescu, A., Daniel, R., DeWoskin, V. & Niv, Y. Dynamic interaction between reinforcement learning and attention in multidimensional environments. *Neuron* **93**, 451–463 (2017).
135. Niv, Y. et al. Reinforcement learning in multidimensional environments relies on attention mechanisms. *J. Neurosci.* **35**, 8145–8157 (2015).
136. Radulescu, A., Daniel, R. & Niv, Y. The effects of aging on the interaction between reinforcement learning and attention. *Psychol. Aging* **31**, 747–757 (2016).
137. Radulescu, A., Niv, Y. & Ballard, I. Holistic reinforcement learning: the role of structure and attention. *Trends Cogn. Sci.* **23**, 278–292 (2019).
138. Conway, C. M. How does the brain learn environmental structure? Ten core principles for understanding the neurocognitive mechanisms of statistical learning. *Neurosci. Biobehav. Rev.* **112**, 279–299 (2020).
139. Forest, T. A., Siegelman, N. & Finn, A. S. Attention shifts to more complex structures with experience. *Psychol. Sci.* **33**, 2059–2072 (2022).
140. Zhao, J., Al-Aidroos, N. & Turk-Browne, N. B. Attention is spontanesouly biased toward regularities. *Psychol. Sci.* **24**, 667–677 (2013).
141. Bein, O. et al. Delineating the effect of semantic congruency on episodic memory: the role of integration and relatedness. *PLoS ONE* **10**, e0115624 (2015).
142. Gronau, N. To grasp the world at a glance: the role of attention in visual and semantic associative processing. *J. Imaging* **7**, 191 (2021).
143. Gronau, N., Neta, M. & Bar, M. Integrated contextual representation for objects' identities and their locations. *J. Cogn. Neurosci.* **20**, 371–388 (2008).
144. Gronau, N. & Shachar, M. Contextual consistency facilitates long-term memory of perceptual detail in barely seen images. *J. Exp. Psychol. Hum. Percept. Perform.* **41**, 1095–1111 (2015).
145. Melega, G. & Sheldon, S. Conceptual relatedness promotes memory generalization at the cost of detailed recollection. *Sci. Rep.* **13**, 15575 (2023).
146. Sherman, B. E. & Turk-Browne, N. B. Statistical prediction of the future impairs episodic encoding of the present. *Proc. Natl Acad. Sci. USA* https://doi.org/10.1101/851147 (2020).
147. Tolman, E. C. Cognitive maps in rats and men. *Psychol. Rev.* **55**, 189–208 (1948).
148. Ho, M. K., Abel, D., Griffiths, T. L. & Littman, M. L. The value of abstraction. *Curr. Opin. Behav. Sci.* **29**, 111–116 (2019).
149. Li, L., Walsh, T. J. & Littman, M. L. Towards a unified theory of state abstraction for MDPs. In *AI&M* (2006).
150. Konidaris, G. On the necessity of abstraction. *Curr. Opin. Behav. Sci.* **29**, 1–7 (2019).
151. Soto, F. A., Gershman, S. J. & Niv, Y. Explaining compound generalization in associative and causal learning through rational principles of dimensional generalization. *Psychol. Rev.* **121**, 526–558 (2014).
152. Gershman, S. J., Radulescu, A., Norman, K. A. & Niv, Y. Statistical computations underlying the dynamics of memory updating. *PLoS Comput. Biol.* **10**, e1003939 (2014).
153. Delgado, M. R. et al. Viewpoints: dialogues on the functional role of the ventromedial prefrontal cortex. *Nat. Neurosci.* **19**, 1545–1552 (2016).
154. Knudsen, E. B. & Wallis, J. D. Taking stock of value in the orbitofrontal cortex. *Nat. Rev. Neurosci.* **23**, 428–438 (2022).
155. Padoa-Schioppa, C. & Conen, K. E. Orbitofrontal cortex: a neural circuit for economic decisions. *Neuron* **96**, 736–754 (2017).
156. Stalnaker, T. A., Cooch, N. K. & Schoenbaum, G. What the orbitofrontal cortex does not do. *Nat. Neurosci.* **18**, 620–627 (2015).
157. Boorman, E. D., Witkowski, P. P., Zhang, Y. & Park, S. A. The orbital frontal cortex, task structure, and inference. *Behav. Neurosci.* **135**, 291–300 (2021).
158. Möhring, L. & Gläscher, J. Prediction errors drive dynamic changes in neural patterns that guide behavior. *Cell Rep.* **42**, 112931 (2023).
159. Vaidya, A. R. & Badre, D. Abstract task representations for inference and control. *Trends Cogn. Sci.* **26**, 484–498 (2022).
160. Zhou, J., Gardner, M. P. H. & Schoenbaum, G. Is the core function of orbitofrontal cortex to signal values or make predictions? *Curr. Opin. Behav. Sci.* **41**, 1–9 (2021).
161. Constantinescu, A. O., O'Reilly, J. X. & Behrens, T. E. J. Organizing conceptual knowledge in humans with a gridlike code. *Science* **352**, 1464–1467 (2016).
162. Viganò, S. & Piazza, M. Distance and direction codes underlie navigation of a novel semantic space in the human brain. *J. Neurosci.* **40**, 2727–2736 (2020).
163. Baram, A. B., Muller, T. H., Nili, H., Garvert, M. M. & Behrens, T. E. J. Entorhinal and ventromedial prefrontal cortices abstract and generalize the structure of reinforcement learning problems. *Neuron* **109**, 713–723.e7 (2021).
164. Klein-Flügge, M. C., Wittmann, M. K., Shpektor, A., Jensen, D. E. A. & Rushworth, M. F. S. Multiple associative structures created by reinforcement and incidental statistical learning mechanisms. *Nat. Commun.* **10**, 4835 (2019).
165. Schapiro, A. C., Rogers, T. T., Cordova, N. I., Turk-Browne, N. B. & Botvinick, M. M. Neural representations of events arise from temporal community structure. *Nat. Neurosci.* **16**, 486–492 (2013).
166. Park, S. A., Miller, D. S., Nili, H., Ranganath, C. & Boorman, E. D. Map making: constructing, combining, and inferring on abstract cognitive maps. *Neuron* **107**, 1226–1238 (2020).
167. Barron, H. C., Dolan, R. J. & Behrens, T. E. J. Online evaluation of novel choices by simultaneous representation of multiple memories. *Nat. Neurosci.* **16**, 1492–1498 (2013).
168. Bongioanni, A. et al. Activation and disruption of a neural mechanism for novel choice in monkeys. *Nature* **591**, 270–274 (2021).
169. Ghosh, V. E., Moscovitch, M., Melo Colella, B. & Gilboa, A. Schema representation in patients with ventromedial PFC lesions. *J. Neurosci.* **34**, 12057–12070 (2014).
170. Gilboa, A. Strategic retrieval, confabulations, and delusions: theory and data. *Cogn. Neuropsychiatry* **15**, 145–180 (2010).

171. Spalding, K. N., Jones, S. H., Duff, M. C., Tranel, D. & Warren, D. E. Investigating the neural correlates of schemas: ventromedial prefrontal cortex is necessary for normal schematic influence on memory. *J. Neurosci.* **35**, 15746–15751 (2015).
172. Zheng, L., Gao, Z., McAvan, A. S., Isham, E. A. & Ekstrom, A. D. Partially overlapping spatial environments trigger reinstatement in hippocampus and schema representations in prefrontal cortex. *Nat. Commun.* **12**, 6231 (2021).
173. Reagh, Z. M. & Ranganath, C. Flexible reuse of cortico-hippocampal representations during encoding and recall of naturalistic events. *Nat. Commun.* **14**, 1279 (2023).
174. Masís-Obando, R., Norman, K. A. & Baldassano, C. Schema representations in distinct brain networks support narrative memory during encoding and retrieval. *eLife* **11**, e70445 (2022).
175. Mack, M. L., Preston, A. R. & Love, B. C. Ventromedial prefrontal cortex compression during concept learning. *Nat. Commun.* **11**, 46 (2020).
Using principal component analysis and functional MRI multivoxel activity patterns, this study provides compelling evidence for dimensionality reduction in the human mOFC–vmPFC during categorization.
176. Audrain, S. & McAndrews, M. P. Schemas provide a scaffold for neocortical integration of new memories over time. *Nat. Commun.* **13**, 5795 (2022).
177. Günseli, E. & Aly, M. Preparation for upcoming attentional states in the hippocampus and medial prefrontal cortex. *eLife* **9**, e53191 (2020).
178. Tse, D. et al. Schema-dependent gene activation and memory encoding in neocortex. *Science* **333**, 891–895 (2011).
179. Warren, D. E., Jones, S. H., Duff, M. C. & Tranel, D. False recall is reduced by damage to the ventromedial prefrontal cortex: implications for understanding the neural correlates of schematic memory. *J. Neurosci.* **34**, 7677–7682 (2014).
180. Brown, V. J. & Tait, D. S. in *Translational Neuropsychopharmacology* (eds Robbins, T. W. & Sahakian, B. J.) **28** (Springer, 2015).
181. Stuss, D. T. et al. Wisconsin Card Sorting Test performance in patients with focal frontal and posterior brain damage: effects of lesion location and test structure on separable cognitive processes. *Neuropsychologia* **38**, 388–402 (2000).
182. Drewe, E. A. The effect of type and area of brain lesion on Wisconsin Card Sorting Test performance. *Cortex* **10**, 159–170 (1974).
183. Bissonette, G. B., Powell, E. M. & Roesch, M. R. Neural structures underlying set-shifting: roles of medial prefrontal cortex and anterior cingulate cortex. *Behav. Brain Res.* **250**, 91–101 (2013).
184. Rushworth, M. F. S., Noonan, M. P., Boorman, E. D., Walton, M. E. & Behrens, T. E. Frontal cortex and reward-guided learning and decision-making. *Neuron* **70**, 1054–1069 (2011).
185. Moneta, N., Garvert, M. M., Heekeren, H. R. & Schuck, N. W. Task state representations in vmPFC mediate relevant and irrelevant value signals and their behavioral influence. *Nat. Commun.* **14**, 3156 (2023).
186. Lebreton, M., Jorge, S., Michel, V., Thirion, B. & Pessiglione, M. An automatic valuation system in the human brain: evidence from functional neuroimaging. *Neuron* **64**, 431–439 (2009).
187. Cowan, E. et al. Sleep spindles promote the restructuring of memory representations in ventromedial prefrontal cortex through enhanced hippocampal-cortical functional connectivity. *J. Neurosci.* **40**, 1909–1919 (2020).
188. Schlichting, M. L., Mumford, J. A. & Preston, A. R. Learning-related representational changes reveal dissociable integration and separation signatures in the hippocampus and prefrontal cortex. *Nat. Commun.* **6**, 1–10 (2015).
189. Bein, O., Reggev, N. & Maril, A. Prior knowledge influences on hippocampus and medial prefrontal cortex interactions in subsequent memory. *Neuropsychologia* **64**, 320–330 (2014).
190. van Kesteren, M. T. R. et al. Differential roles for medial prefrontal and medial temporal cortices in schema-dependent encoding: from congruent to incongruent. *Neuropsychologia* **51**, 2352–2359 (2013).
191. Reggev, N., Bein, O. & Maril, A. Distinct neural suppression and encoding effects for conceptual novelty and familiarity. *J. Cogn. Neurosci.* **28**, 1455–1470 (2016).
192. Yacoby, A., Reggev, N. & Maril, A. Examining the transition of novel information toward familiarity. *Neuropsychologia* **161**, 107993 (2021).
193. Yacoby, A., Reggev, N. & Maril, A. Lack of source memory as a potential marker of early assimilation of novel items into current knowledge. *Neuropsychologia* **185**, 108569 (2023).
194. Chiba, A. A., Kesner, R. P. & Gibson, C. J. Memory for temporal order of new and familiar spatial location sequences: role of the medial prefrontal cortex. *Learn. Mem.* **4**, 311–317 (1997).
195. Tiganj, Z., Jung, M. W., Kim, J. & Howard, M. W. Sequential firing codes for time in rodent medial prefrontal cortex. *Cereb. Cortex* **27**, 5663–5671 (2017).
196. Hsieh, L.-T. & Ranganath, C. Cortical and subcortical contributions to sequence retrieval: schematic coding of temporal context in the neocortical recollection network. *Neuroimage* **121**, 78–90 (2015).
197. Ezzyat, Y. & Davachi, L. What constitutes an episode in episodic memory? *Psychol. Sci.* **22**, 243–252 (2011).
198. Barbas, H. & Blatt, G. J. Topographically specific hippocampal projections target functionally distinct prefrontal areas in the rhesus monkey. *Hippocampus* **5**, 511–533 (1995).
199. Cavada, C., Compañy, T., Tejedor, J., Cruz-Rizzolo, R. J. & Reinoso-Suárez, F. The anatomical connections of the macaque monkey orbitofrontal cortex. A review. *Cereb. Cortex* **10**, 220–242 (2000).
200. Öngür, D. & Price, J. L. The organization of networks within the orbital and medial prefrontal cortex of rats, monkeys and humans. *Cereb. Cortex* **10**, 206–219 (2000).
201. Price, J. L. Definition of the orbital cortex in relation to specific connections with limbic and visceral structures and other cortical regions. *Ann. N. Y. Acad. Sci.* **1121**, 54–71 (2007).
202. Eichenbaum, H. Time cells in the hippocampus: a new dimension for mapping memories. *Nat. Rev. Neurosci.* **15**, 732–744 (2014).
203. Ranganath, C. & Hsieh, L.-T. The hippocampus: a special place for time. *Ann. N. Y. Acad. Sci.* **1369**, 93–110 (2016).
204. Kesner, R. P. & Hunsaker, M. R. The temporal attributes of episodic memory. *Behav. Brain Res.* **215**, 299–309 (2010).
205. Buzsáki, G. & Tingley, D. Space and time: the hippocampus as a sequence generator. *Trends Cogn. Sci.* **22**, 853–869 (2018).
206. DuBrow, S. & Davachi, L. Temporal binding within and across events. *Neurobiol. Learn. Mem.* **134**, 107–114 (2016).
207. Schapiro, A. C., Turk-Browne, N. B., Norman, K. A. & Botvinick, M. M. Statistical learning of temporal community structure in the hippocampus. *Hippocampus* **26**, 3–8 (2016).
208. Eichenbaum, H. Prefrontal–hippocampal interactions in episodic memory. *Nat. Rev. Neurosci.* **18**, 547–558 (2017).
209. Fortin, N. J., Agster, K. L. & Eichenbaum, H. B. Critical role of the hippocampus in memory for sequences of events. *Nat. Neurosci.* **5**, 458–462 (2002).
210. Allen, T. A., Salz, D. M., McKenzie, S. & Fortin, N. J. Nonspatial sequence coding in CA1 neurons. *J. Neurosci.* **36**, 1547–1563 (2016).
211. Reeders, P. C., Hamm, A. G., Allen, T. A. & Mattfeld, A. T. Medial prefrontal cortex and hippocampal activity differentially contribute to ordinal and temporal context retrieval during sequence memory. *Learn. Mem.* **28**, 134–147 (2021).
212. Ranganath, C. Time, memory, and the legacy of Howard Eichenbaum. *Hippocampus* **29**, 146–161 (2019).
213. Jayachandran, M. et al. Prefrontal pathways provide top-down control of memory for sequences of events. *Cell Rep.* **28**, 640–654.e6 (2019).
214. Place, R., Farovik, A., Brockmann, M. & Eichenbaum, H. Bidirectional prefrontal–hippocampal interactions support context-guided memory. *Nat. Neurosci.* **19**, 992–994 (2016).
215. McCormick, C., Barry, D. N., Jafarian, A., Barnes, G. R. & Maguire, E. A. vmPFC drives hippocampal processing during autobiographical memory recall regardless of remoteness. *Cereb. Cortex* **30**, 5972–5987 (2020).
216. Bar, M. et al. Top-down facilitation of visual recognition. *Proc. Natl Acad. Sci. USA* **103**, 449–454 (2006).
217. Bar, M. The proactive brain: memory for predictions. *Philos. Trans. R. Soc. Lond. B* **364**, 1235–1243 (2009).
218. Ezzyat, Y. & Davachi, L. Neural evidence for representational persistence within events. *J. Neurosci.* **41**, 7909–7920 (2021).
219. Guise, K. G. & Shapiro, M. L. Medial prefrontal cortex reduces memory interference by modifying hippocampal encoding. *Neuron* **94**, 183–192.e8 (2017).
220. Hebscher, M., Barkan-Abramski, M., Goldsmith, M., Aharon-Peretz, J. & Gilboa, A. Memory, decision-making, and the ventromedial prefrontal cortex (vmPFC): the roles of subcallosal and posterior orbitofrontal cortices in monitoring and control processes. *Cereb. Cortex* **26**, 4590–4601 (2016).
221. Hebscher, M. & Gilboa, A. A boost of confidence: the role of the ventromedial prefrontal cortex in memory, decision-making, and schemas. *Neuropsychologia* **90**, 46–58 (2016).
222. Gilboa, A., Alain, C., He, Y., Stuss, D. T. & Moscovitch, M. Ventromedial prefrontal cortex lesions produce early functional alterations during remote memory retrieval. *J. Neurosci.* **29**, 4871–4881 (2009).
223. Alonso, A., van der Meij, J., Tse, D. & Genzel, L. Naïve to expert: considering the role of previous knowledge in memory. *Brain Neurosci. Adv.* **4**, 239821282094868 (2020).
224. Sommer, T. The emergence of knowledge and how it supports the memory for novel related information. *Cereb. Cortex* **27**, 1906–1921 (2017).
225. Sommer, T., Hennies, N., Lewis, P. A. & Alink, A. The assimilation of novel information into schemata and its efficient consolidation. *J. Neurosci.* **42**, 5916–5929 (2022).
In a Herculean effort, the authors tracked the emergence of schemas in the brain over a year, and showed changes in hippocampal and prefrontal activations during schema instantiation and the influence of learned schemas on new learning.
226. van Kesteren, M. T. R., Fernández, G., Norris, D. G. & Hermans, E. J. Persistent schema-dependent hippocampal-neocortical connectivity during memory encoding and postencoding rest in humans. *Proc. Natl Acad. Sci. USA* **107**, 7550–7555 (2010).
227. van Kesteren, M. T. R., Rijpkema, M., Ruiter, D. J. & Fernández, G. Retrieval of associative information congruent with prior knowledge is related to increased medial prefrontal activity and connectivity. *J. Neurosci.* **30**, 15888–15894 (2010).
228. Mackey, S. & Petrides, M. Quantitative demonstration of comparable architectonic areas within the ventromedial and lateral orbital frontal cortex in the human and the macaque monkey brains. *Eur. J. Neurosci.* **32**, 1940–1950 (2010).
229. Mackey, S. & Petrides, M. Architecture and morphology of the human ventromedial prefrontal cortex. *Eur. J. Neurosci.* **40**, 2777–2796 (2014).
230. Uylings, H. B. M. et al. 3-D cytoarchitectonic parcellation of human orbitofrontal cortex: correlation with postmortem MRI author links open overlay panel. *Psychiatry Res.* **183**, 1–20 (2010).
231. Jackson, R. L., Bajada, C. J., Lambon Ralph, M. A. & Cloutman, L. L. The graded change in connectivity across the ventromedial prefrontal cortex reveals distinct subregions. *Cereb. Cortex* **30**, 165–180 (2020).

232. Boorman, E. D., Behrens, T. E. J., Woolrich, M. W. & Rushworth, M. F. S. How green is the grass on the other side? Frontopolar cortex and the evidence in favor of alternative courses of action. *Neuron* **62**, 733–743 (2009).
233. Boorman, E. D., Behrens, T. E. & Rushworth, M. F. Counterfactual choice and learning in a neural network centered on human lateral frontopolar cortex. *PLoS Biol.* **9**, e1001093 (2011).
234. Haynes, J.-D. et al. Reading hidden intentions in the human brain. *Curr. Biol.* **17**, 323–328 (2007).
235. Momennejad, I. & Haynes, J.-D. Human anterior prefrontal cortex encodes the 'what' and 'when' of future intentions. *Neuroimage* **61**, 139–148 (2012).
236. Momennejad, I. & Haynes, J.-D. Encoding of prospective tasks in the human prefrontal cortex under varying task loads. *J. Neurosci.* **33**, 17342–17349 (2013).
237. Liberman, N., Sagristano, M. D. & Trope, Y. The effect of temporal distance on level of mental construal. *J. Exp. Soc. Psychol.* **38**, 523–534, (2002).
238. Trope, Y. & Liberman, N. Temporal construal. *Psychol. Rev.* **110**, 403–421 (2003).
239. Trope, Y. & Liberman, N. Construal-level theory of psychological distance. *Psychol. Rev.* **117**, 440–463 (2010).
240. Takashima, A. et al. Declarative memory consolidation in humans: a prospective functional magnetic resonance imaging study. *Proc. Natl Acad. Sci. USA* **103**, 756–761 (2006).
241. Baker, S. et al. The human dentate gyrus plays a necessary role in discriminating new memories. *Curr. Biol.* **26**, 2629–2634 (2016).
242. Bakker, A., Kirwan, C. B., Miller, M. & Stark, C. E. L. Pattern separation in the human hippocampal CA3 and dentate gyrus. *Science* **319**, 1640–1642 (2008).
243. Berron, D. et al. Strong evidence for pattern separation in human dentate gyrus. *J. Neurosci.* **36**, 7569–7579 (2016).
244. Leutgeb, J. K., Leutgeb, S., Moser, M.-B. & Moser, E. I. Pattern separation in the dentate gyrus and CA3 of the hippocampus. *Science* **315**, 961–966 (2007).
245. Yassa, M. A. & Stark, C. E. L. Pattern separation in the hippocampus. *Trends Neurosci.* **34**, 515–525 (2011).
246. Ballard, I. C., Wagner, A. D. & McClure, S. M. Hippocampal pattern separation supports reinforcement learning. *Nat. Commun.* **10**, 1073 (2019).
247. Jiang, J., Wang, S.-F., Guo, W., Fernandez, C. & Wagner, A. D. Prefrontal reinstatement of contextual task demand is predicted by separable hippocampal patterns. *Nat. Commun.* **11**, 2053 (2020).
248. Zhou, J. et al. Complementary task structure representations in hippocampus and orbitofrontal cortex during an odor sequence task. *Curr. Biol.* **29**, 3402–3409.e3 (2019).
249. Howard, J. D., Gottfried, J. A., Tobler, P. N. & Kahnt, T. Identity-specific coding of future rewards in the human orbitofrontal cortex. *Proc. Natl Acad. Sci. USA* **112**, 5195–5200 (2015).
250. McNamee, D., Rangel, A. & O'Doherty, J. P. Category-dependent and category-independent goal-value codes in human ventromedial prefrontal cortex. *Nat. Neurosci.* **16**, 479–485 (2013).
251. Sweegers, C. C. G., Takashima, A., Fernández, G. & Talamini, L. M. Neural mechanisms supporting the extraction of general knowledge across episodic memories. *Neuroimage* **87**, 138–146 (2014).
252. Kahnt, T., Chang, L. J., Park, S. Q., Heinzle, J. & Haynes, J.-D. Connectivity-based parcellation of the human orbitofrontal cortex. *J. Neurosci.* **32**, 6240–6250 (2012).
253. Baldassano, C. et al. Discovering event structure in continuous narrative perception and memory. *Neuron* **95**, 709–721.e5 (2017).
254. Chen, J. et al. Shared memories reveal shared structure in neural activity across individuals. *Nat. Neurosci.* **20**, 115–125 (2017).
255. Hasson, U., Chen, J. & Honey, C. J. Hierarchical process memory: memory as an integral component of information processing. *Trends Cogn. Sci.* **19**, 304–313 (2015).
256. Levy, D. J. & Glimcher, P. W. The root of all value: a neural common currency for choice. *Curr. Opin. Neurobiol.* **22**, 1027–1038 (2012).
257. Barron, H. C., Garvert, M. M. & Behrens, T. E. J. Reassessing VMPFC: full of confidence? *Nat. Neurosci.* **18**, 1064–1066 (2015).
258. Howard, J. D. & Kahnt, T. To be specific: the role of orbitofrontal cortex in signaling reward identity. *Behav. Neurosci.* **135**, 210–217 (2021).
259. Liu, Y., Mattar, M. G., Behrens, T. E. J., Daw, N. D. & Dolan, R. J. Experience replay is associated with efficient nonlocal learning. *Science* **372**, eabf1357 (2021).
260. Schuck, N. W. & Niv, Y. Sequential replay of nonspatial task states in the human hippocampus. *Science* **364**, eaaw5181 (2019).
261. Graves, K. N., Antony, J. W. & Turk-Browne, N. B. Finding the pattern: on-line extraction of spatial structure during virtual navigation. *Psychol. Sci.* **31**, 1183–1190 (2020).
262. Richards, B. A. et al. Patterns across multiple memories are identified over time. *Nat. Neurosci.* **17**, 981–986 (2014).
263. Antony, J. W. et al. Spatial gist extraction during human memory consolidation. *J. Exp. Psychol. Learn. Mem. Cogn.* **48**, 929–941 (2022).
264. Graves, K. N. et al. Remembering the pattern: a longitudinal case study on statistical learning in spatial navigation and memory consolidation. *Neuropsychologia* **174**, 108341 (2022).
265. Ye, D. et al. Towards playing full MOBA games with deep reinforcement learning. *Adv. Neural Inf. Process. Syst.* **33**, 621–632 (2020).
266. Lohokare, A., Shah, A. & Zyda, M. Deep learning bot for league of legends. In *Proc. AAAI Conference on Artificial Intelligence and Interactive Digital Entertainment* **16**, 322–324 (2020).
267. Zhang, Z. et al. Hierarchical reinforcement learning for multi-agent MOBA game. Preprint at http://arxiv.org/abs/1901.08004 (2019).
268. Wise, T., Emery, K. & Radulescu, A. Naturalistic reinforcement learning. *Trends Cogn. Sci.* **28**, 144–158 (2024).
269. Lu, Q., Hasson, U. & Norman, K. A. A neural network model of when to retrieve and encode episodic memories. *eLife* **11**, e74445 (2022).
270. Schultz, W., Dayan, P. & Montague, P. R. A neural substrate of prediction and reward. *Science* **275**, 1593–1599 (1997).
271. Shohamy, D. & Turk-Browne, N. B. Mechanisms for widespread hippocampal involvement in cognition. *J. Exp. Psychol. Gen.* **142**, 1159–1170 (2013).
272. Marr, D. Simple memory a theory for archicortex. *Philos. Trans. R. Soc. Lond.* **262**, 23–81 (1971).
273. McClelland, J. L., McNaughton, B. L. & O'Reilly, R. C. Why there are complementary learning systems in the hippocampus and neocortex: insights from the successes and failures of connectionist models of learning and memory. *Psychol. Rev.* **102**, 419–457 (1995).
274. Shteingart, H., Neiman, T. & Loewenstein, Y. The role of first impression in operant learning. *J. Exp. Psychol. Gen.* **142**, 476–488 (2013).
275. Sherman, B. E. et al. Intracranial entrainment reveals statistical learning across levels of abstraction. *J. Cogn. Neurosci.* **35**, 1312–1328 (2023).
276. Eichenbaum, H. Hippocampus: cognitive processes and neural representations that underlie declarative memory. *Neuron* **44**, 109–120 (2004).
277. Norman, K. A. & O'Reilly, R. C. Modeling hippocampal and neocortical contributions to recognition memory: a complementary-learning-systems approach. *Psychol. Rev.* **110**, 611–646 (2003).
278. Baraduc, P., Duhamel, J.-R. & Wirth, S. Schema cells in the macaque hippocampus. *Science* **363**, 635–639 (2019).
279. Crivelli-Decker, J. Goal-oriented representations in the human hippocampus during planning and navigation. *Nat. Commun.* **14**, 2946 (2023).
280. Farzanfar, D., Spiers, H. J., Moscovitch, M. & Rosenbaum, R. S. From cognitive maps to spatial schemas. *Nat. Rev. Neurosci.* https://doi.org/10.1038/s41583-022-00655-9 (2022).
281. Garvert, M. M., Dolan, R. J. & Behrens, T. E. J. A map of abstract relational knowledge in the human hippocampal–entorhinal cortex. *eLife* **6**, e17086 (2017).
282. Garvert, M. M., Saanum, T., Schulz, E., Schuck, N. W. & Doeller, C. F. Hippocampal spatio-predictive cognitive maps adaptively guide reward generalization. *Nat. Neurosci.* **26**, 615–626 (2023).
283. Kumaran, D., Summerfield, J. J., Hassabis, D. & Maguire, E. A. Tracking the emergence of conceptual knowledge during human decision making. *Neuron* **63**, 889–901 (2009).
284. McKenzie, S. et al. Hippocampal representation of related and opposing memories develop within distinct, hierarchically organized neural schemas. *Neuron* **83**, 202–215 (2014).
285. Schapiro, A. C., Gregory, E., Landau, B., McCloskey, M. & Turk-Browne, N. B. The necessity of the medial temporal lobe for statistical learning. *J. Cogn. Neurosci.* **26**, 1736–1747 (2014).
286. Theves, S., Fernandez, G. & Doeller, C. F. The hippocampus encodes distances in multidimensional feature space. *Curr. Biol.* **29**, 1226–1231.e3 (2019).
287. Theves, S., Neville, D., Fernández, G. & Doeller, C. F. Learning and representation of hierarchical concepts in hippocampus and prefrontal cortex. *J. Neurosci.* **41**, 7675–7686 (2021).
288. Schapiro, A. C., Turk-Browne, N. B., Botvinick, M. M. & Norman, K. A. Complementary learning systems within the hippocampus: a neural network modelling approach to reconciling episodic memory with statistical learning. *Philos. Trans. R. Soc. Lond. B* **372**, 20160049 (2017).
289. Barnett, A. J. et al. Hippocampal–cortical interactions during event boundaries support retention of complex narrative events. *Neuron* https://doi.org/10.1016/j.neuron.2023.10.010 (2023).
290. Ben-Yakov, A. & Dudai, Y. Constructing realistic engrams: poststimulus activity of hippocampus and dorsal striatum predicts subsequent episodic memory. *J. Neurosci.* **31**, 9032–9042 (2011).
291. Ben-Yakov, A. & Henson, R. N. The hippocampal film editor: sensitivity and specificity to event boundaries in continuous experience. *J. Neurosci.* **38**, 10057–10068 (2018).
292. Behrens, T. E. J. et al. What is a cognitive map? Organizing knowledge for flexible behavior. *Neuron* **100**, 490–509 (2018).
293. Bellmund, J. L. S., Gärdenfors, P., Moser, E. I. & Doeller, C. F. Navigating cognition: spatial codes for human thinking. *Science* **362**, eaat6766 (2018).
294. Schiller, D. et al. Memory and space: towards an understanding of the cognitive map. *J. Neurosci.* **35**, 13904–13911 (2015).
295. Tavares, R. M. et al. A map for social navigation in the human brain. *Neuron* **87**, 231–243 (2015).
296. Theves, S., Fernández, G. & Doeller, C. F. The hippocampus maps concept space, not feature space. *J. Neurosci.* **40**, 7318–7325 (2020).
297. McClelland, J. L., McNaughton, B. L. & Lampinen, A. K. Integration of new information in memory: new insights from a complementary learning systems perspective. *Philos. Trans. R. Soc. Lond. B* **375**, 20190637 (2020).
298. Tenenbaum, J. B., Kemp, C., Griffiths, T. L. & Goodman, N. D. How to grow a mind: statistics, structure, and abstraction. *Science* **331**, 1279–1285 (2011).
299. Kemp, C., Tenenbaum, J. B., Griffiths, T. L., Yamada, T. & Ueda, N. Learning systems of concepts with an infinite relational model. *AAAI* **3**, 5 (2006).

300. Roesch, M. R. & Olson, C. R. Neuronal activity in primate orbitofrontal cortex reflects the value of time. *J. Neurophysiol.* **94**, 2457–2471 (2005).
301. Padoa-Schioppa, C. & Assad, J. A. Neurons in the orbitofrontal cortex encode economic value. *Nature* **441**, 223–226 (2006).
302. Padoa-Schioppa, C. & Assad, J. A. The representation of economic value in the orbitofrontal cortex is invariant for changes of menu. *Nat. Neurosci.* **11**, 95–102 (2008).
303. Kahnt, T., Heinzle, J., Park, S. Q. & Haynes, J.-D. The neural code of reward anticipation in human orbitofrontal cortex. *Proc. Natl Acad. Sci. USA* **107**, 6010–6015 (2010).
304. Elliott, R., Agnew, Z. & Deakin, J. F. W. Medial orbitofrontal cortex codes relative rather than absolute value of financial rewards in humans. *Eur. J. Neurosci.* **27**, 2213–2218 (2008).
305. Camille, N., Griffiths, C. A., Vo, K., Fellows, L. K. & Kable, J. W. Ventromedial frontal lobe damage disrupts value maximization in humans. *J. Neurosci.* **31**, 7527–7532 (2011).
306. Izquierdo, A., Suda, R. K. & Murray, E. A. Bilateral orbital prefrontal cortex lesions in rhesus monkeys disrupt choices guided by both reward value and reward contingency. *J. Neurosci.* **24**, 7540–7548 (2004).
307. Fellows, L. K. & Farah, M. J. The role of ventromedial prefrontal cortex in decision making: judgment under uncertainty or judgment per se? *Cereb. Cortex* **17**, 2669–2674 (2007).
308. Pelletier, G. & Fellows, L. K. A critical role for human ventromedial frontal lobe in value comparison of complex objects based on attribute configuration. *J. Neurosci.* **39**, 4124–4132 (2019).
309. Dias, R., Robbins, T. W. & Roberts, A. C. Dissociation in prefrontal cortex of affective and attentional shifts. *Nature* **380**, 69–72 (1996).
310. Izquierdo, A., Brigman, J. L., Radke, A. K., Rudebeck, P. H. & Holmes, A. The neural basis of reversal learning: an updated perspective. *Neuroscience* **345**, 12–26 (2017).
311. Rudebeck, P. H. & Rich, E. L. Orbitofrontal cortex. *Curr. Biol.* **28**, R1083–R1088 (2018).
312. Brockett, A. T. & Roesch, M. R. The ever-changing OFC landscape: what neural signals in OFC can tell us about inhibitory control. *Behav. Neurosci.* **135**, 129–137 (2021).
313. Gardner, M. P. H. & Schoenbaum, G. The orbitofrontal cartographer. *Behav. Neurosci.* **135**, 267–276 (2021).
314. Zhou, J. et al. Rat orbitofrontal ensemble activity contains multiplexed but dissociable representations of value and task structure in an odor sequence task. *Curr. Biol.* **29**, 897–907.e3 (2019).
315. Basu, R. et al. The orbitofrontal cortex maps future navigational goals. *Nature* **599**, 449–452 (2021).
316. Mansouri, F. A., Freedman, D. J. & Buckley, M. J. Emergence of abstract rules in the primate brain. *Nat. Rev. Neurosci.* **21**, 595–610 (2020).
317. Koechlin, E. & Summerfield, C. An information theoretical approach to prefrontal executive function. *Trends Cogn. Sci.* **11**, 229–235 (2007).
318. Goldman-Rakic, P. S. Architecture of the prefrontal cortex and the central executive. *Ann. N. Y. Acad. Sci.* **769**, 71–84 (1995).
319. Milner, B. Effects of different brain lesions on card sorting: the role of the frontal lobes. *Arch. Neurol.* **9**, 90–100 (1963).
320. Amer, T. & Davachi, L. Extra-hippocampal contributions to pattern separation. *eLife* **12**, e82250 (2023).
321. Bi, Y. Dual coding of knowledge in the human brain. *Trends Cogn. Sci.* **25**, 883–895 (2021).
322. Binder, J. R. & Desai, R. H. The neurobiology of semantic memory. *Trends Cogn. Sci.* **15**, 527–536 (2011).
323. Brod, G., Werkle-Bergner, M. & Shing, Y. L. The influence of prior knowledge on memory: a developmental cognitive neuroscience perspective. *Front. Behav. Neurosci.* **7**, 139 (2013).
324. Thompson-Schill, S. L., D'Esposito, M., Aguirre, G. K. & Farah, M. J. Role of left inferior prefrontal cortex in retrieval of semantic knowledge: a reevaluation. *Proc. Natl Acad. Sci. USA* **94**, 14792–14797 (1997).
325. Badre, D. & Wagner, A. D. Left ventrolateral prefrontal cortex and the cognitive control of memory. *Neuropsychologia* **45**, 2883–2901 (2007).
326. Miller, E. K. & Cohen, J. D. An integrative theory of prefrontal cortex function. *Annu. Rev. Neurosci.* **24**, 167–202 (2001).
327. Giallanza, T., Campbell, D., Cohen, J. D. & Rogers, T. T. An integrated model of semantics and control. *Psychol. Rev.* https://doi.org/10.31234/osf.io/jq7ta (2023).
328. Joel, D. & Weiner, I. The connections of the dopaminergic system with the striatum in rats and primates: an analysis with respect to the functional and compartmental organization of the striatum. *Neuroscience* **96**, 451–474 (2000).
329. Aarts, E., van Holstein, M. & Cools, R. Striatal dopamine and the interface between motivation and cognition. *Front. Psychol.* **2**, 163 (2011).
330. Barto, A. G. in *Models of Information Processing in the Basal Ganglia* (eds Houk, J. C., Davis, J. L. & Beiser, D. G.) 215–232 (MIT, 1994).
331. Niv, Y. Reinforcement learning in the brain. *J. Math. Psychol.* **53**, 139–154 (2009).
332. Frank, M. J. & Badre, D. Mechanisms of hierarchical reinforcement learning in corticostriatal circuits 1: computational analysis. *Cereb. Cortex* **22**, 509–526 (2012).
333. Badre, D. & Frank, M. J. Mechanisms of hierarchical reinforcement learning in cortico-striatal circuits 2: evidence from fMRI. *Cereb. Cortex* **22**, 527–536 (2012).
334. Pine, A., Sadeh, N., Ben-Yakov, A., Dudai, Y. & Mendelsohn, A. Knowledge acquisition is governed by striatal prediction errors. *Nat. Commun.* **9**, 1673 (2018).
335. Wittmann, B. C. et al. Reward-related FMRI activation of dopaminergic midbrain is associated with enhanced hippocampus-dependent long-term memory formation. *Neuron* **45**, 459–467 (2005).

Acknowledgements

This work was funded by a NIMH T32 fellowship to O.B. (T32MH065214).

Author contributions

O.B. researched data for the article. All authors contributed substantially to discussion of the content, wrote the article, reviewed and/or edited the manuscript before submission.

Competing interests

The authors declare no competing interests.

Additional information

Peer review information *Nature Reviews Neuroscience* thanks C. Baldassano, who co-reviewed with J. Huang; and the other, anonymous, reviewers for their contribution to the peer review of this work.