HHS Public Access

Author manuscript

Nat Rev Neurosci. Author manuscript; available in PMC 2023 September 20.

Published in final edited form as:

Nat Rev Neurosci. 2022 July ; 23(7): 428–438. doi:10.1038/s41583-022-00589-2.

Taking stock of value in the orbitofrontal cortex

Eric B. Knudsen[1], **Joni D. Wallis**[2]

[1]Helen Wills Neuroscience Institute, University of California Berkeley, Berkeley, California, USA.

[2]Department of Psychology and Helen Wills Neuroscience Institute, University of California Berkeley, Berkeley, California, USA.

Abstract

People with damage to the orbitofrontal cortex (OFC) have specific problems making decisions, whereas their other cognitive functions are spared. Neurophysiological studies have shown that OFC neurons fire in proportion to the value of anticipated outcomes. Thus, a central role of the OFC is to guide optimal decision-making by signalling values associated with different choices. Until recently, this view of OFC function dominated the field. New data, however, suggest that the OFC may have a much broader role in cognition by representing cognitive maps that can be used to guide behaviour and that value is just one of many variables that are important for behavioural control. In this Review, we critically evaluate these two alternative accounts of OFC function and examine how they might be reconciled.

In the 1990s, Damasio brought attention to the fascinating deficits exhibited by people with damage to the orbitofrontal cortex (OFC)[1]. These individuals would make terrible choices and their everyday lives lurched from one catastrophic decision to another, yet they would perform normally on standard laboratory tests of cognition[2]. Damasio developed new tests based on real-world decision-making[3]. Participants were given play money and had to make choices across a series of gambles based on reward history. People with OFC damage performed poorly, quickly losing all their money as they failed to learn from the outcomes associated with each gamble. These impairments had gone unnoticed for years, simply because no one thought to assess individuals' decision-making in a neuropsychological exam. Inspired by Damasio's findings, many labs started to record OFC neurons in non-human primates during decision-making tasks. In a seminal study, Padoa-Schioppa and Assad presented monkeys with choices between different amounts of various juices[4]. OFC neurons encoded the amount of juice, weighted by how much the monkey liked it, consistent with how valuable the option was to the animal. This made a neat and compelling story: individuals with OFC damage made poor decisions because they were missing the neurons that would signal the value of each option.

wallis@berkeley.edu .

Author contributions

The authors contributed equally to all aspects of the article.

Competing interests

The authors declare no competing interests.

Until recently, this value hypothesis remained the dominant view of OFC function. However, a major gap in this theory is how the values of options are learned in the first place. Studies that have examined such learning have come to a radically different view of OFC function, arguing that the OFC is responsible for representing a 'cognitive map'. This term was first introduced by Tolman[5], who studied rats learning to navigate mazes. Tolman observed that animals were not just learning simple responses but rather were constantly constructing a mental map of the maze, which he referred to as a 'cognitive map'. More recently, the concept of a cognitive map has been expanded to describe any network of associations that specifies the relationships that underpin a task[6]. The neuronal instantiation of the cognitive map was originally ascribed to the hippocampus[7], but recent neuroimaging studies have found that the OFC is the only cortical region to be activated when participants use cognitive maps[8], suggesting that it may play an important role.

Thus, the OFC field currently finds itself at a crossroads in which two competing hypotheses — the value hypothesis and the cognitive map hypothesis — potentially provide explanations of OFC function. This Review critically evaluates each of these hypotheses and examines how they might be related. In addition, the cognitive map hypothesis raises an additional question: how do the OFC and hippocampus interact in terms of implementing the cognitive map?

The value hypothesis

There are several features of OFC neurons — observed through recordings of such neurons in non-human primates — on which nearly everyone agrees. First, OFC neurons respond to a large range of ways by which the desirability of an option varies. For example, OFC neurons encode the probability of receiving a reward[9], the amount of effort needed to earn a reward[10], the delay until reward delivery[10,11], the amount of secondary reinforcer[12], and whether an outcome is anticipated to be appetitive or aversive[13-15]. Thus, to at least some extent, the OFC value signal is abstract. Second, more neurons encode the value of the option that will be chosen than the option that will not be chosen[4,9,16]. This suggests that the OFC is involved in selecting which option will be chosen. Third, there is a relative lack of motor signals in the OFC, at least compared with other frontal areas[4,9]. This is consistent with its anatomy: it has stronger connections to areas involved in autonomic control than to those involved in musculoskeletal control[17-19]. Thus, OFC decisions take place in what has been dubbed a 'goods space', facilitating decisions independently of the actions necessary to obtain the outcome[20].

In addition to neurophysiology, there is a good deal of convergent evidence to support the role of the OFC in decision-making. Neuroimaging studies in humans have observed OFC activation in various decision-making tasks, including choices between different food items[21-23], erotica[24,25] and monetary gambles[26]. Individuals with OFC damage show decision-making deficits across a broad range of domains, including choosing between colours[27], apartments[28] and political candidates[29]. Perhaps the gold standard for assessing such deficits involves presenting participants with choices in which the outcomes vary along more than one dimension (for example, the type and amount of juice), thereby requiring participants to integrate across these dimensions to calculate the value of the option.

Impairments on such tasks occur following OFC damage in humans[30], during electrical microstimulation of the OFC in monkeys[31] and during optogenetic inactivation of the OFC in rodents[32].

The precise mechanisms by which OFC value responses are translated into a decision are less clear. Computational models have been proposed in which distinct populations of neurons are responsible for representing the value of the options on offer and then gradually encode the value of the chosen option[33,34]. Downstream areas, such as the lateral prefrontal cortex (PFC), are then thought to be responsible for translating the decision into an action[35-37]. Empirical data show that option values and the ultimately chosen value are encoded by distinct populations of OFC neurons[38]. However, there is little evidence that encoding of option values precedes encoding of the chosen value, with both signals arising at approximately the same time[38].

An alternative possibility for how OFC activity is used to make decisions has been raised by a recent study that focused on the dynamics of ensembles of OFC neurons (FIG. 1). Multiple OFC neurons were recorded simultaneously while monkeys were presented with offers of specific values[39]. A decoding algorithm was then trained to classify patterns of OFC activity according to the value being represented. Rather than discrete populations encoding different properties of the choice, OFC neurons seemed to vacillate between representing the value of each option in turn. In addition, more-valuable options were represented more frequently and for longer than less-valuable options, with the result that the higher value of the chosen option dominated neuronal encoding when averaged across trials. Thus, the apparent prevalence of neurons encoding the chosen option was an artefact of averaging neuronal activity across trials rather than a real feature of OFC neurons. Nevertheless, the vacillation did affect the decision: more vacillation produced longer choice reaction times and choices that were more likely to be suboptimal. A downstream area involved in motor planning could in principle use OFC vacillation dynamics to select the more valuable option by integrating the vacillation.

One challenge to the value hypothesis has been in defining when OFC value signals are required to make a decision. Some of the earliest studies showed that monkeys with OFC damage were not impaired at learning to choose between novel stimulus–outcome associations[40], although they were impaired when the task contingencies changed. Monkeys with OFC damage also have intact food preferences[41-44] but do not update these preferences when their motivational state changes[41,43]. A neuroimaging study in humans used the phenomenon of repetition suppression[45] to show that the OFC is preferentially involved when participants consider novel as opposed to familiar choices[22].

Explanations for which tasks will require the OFC have typically focused on how values are learned[46,47]. Reinforcement learning is a computational framework that was developed in psychology and artificial intelligence to formalize how agents learn to select actions to acquire reward and avoid punishment[48]. Importantly, there are two main reinforcement learning methods[49-51]. Model-free reinforcement learning is associated with habits and skills, and relies on trial-and-error learning, storing, or caching the values of past actions and inflexibly repeating actions that have led to higher values. Model-based reinforcement

learning is associated with goal-directed actions and involves the generation of predictions through a computationally expensive process that depends on an internal model of the environment. Importantly, such a model allows rapid updating of choices by a process of inference. For example, consider two distinct sequences of responses that lead to either a food reward or a liquid reward. If one's motivational state changed from hungry to thirsty, a model-free system would have to experience the devalued food reward and gradually reduce the value of those responses relative to that of the responses leading to liquid. By contrast, a model-based system could immediately infer which sequence of responses would lead to liquid.

Damage to the OFC typically produces deficits on tasks that require model-based inference but not on tasks that rely on model-free cached values[52,53]. The lack of such inference helps to explain why animals with OFC damage have difficulty updating their behaviour in response to changes in motivational state[41,43]. This framework does not necessarily contradict the notion that OFC is important for economic decisions. Indeed, such decisions are thought to be deliberative rather than automatic[47,54], and consequently more consistent with a model-based system.

Another puzzle is why so much cortical real estate is devoted to representing a single scalar quantity. One argument is that value signals must be constructed from decision primitives, such as reward magnitude or reward probability[20], analogous to the way that the visual system constructs a scene from visual primitives such as orientation and colour. Indeed, although there are many OFC neurons that encode abstract values that involve the integration of multiple decision primitives, there are just as many OFC neurons that respond to single decision dimensions[9,55]. It is also important to differentiate between the information that an area stores in its pattern of synaptic weights, and what the area represents in its current pattern of activity. For example, we have argued that OFC vacillation may arise from competing representations of option values, with options that predict higher values having a representational advantage[56]. In this conceptualization, previously experienced stimulus–outcome associations are stored in a distributed network in the OFC, with the strength of the synaptic connections proportional to the value of the outcome. When the animal faces a choice, these stimulus–outcome associations are simultaneously activated and compete for representation, resulting in vacillation. This conceptualization is consistent with both neuroimaging[57] and neuropsychological results[58], suggesting that the OFC may contain a representation of reward that is richer than simply the value of a reward.

The cognitive map hypothesis

Over the past decade, a radically different view of OFC function has developed. In part, this has been driven by increasing interest in how agents, both biological and artificial, learn[59]. In the 1970s, researchers discovered that hippocampal neurons in rats encoded the animal's location in space, consistent with Tolman's idea that animals formed maps of their environment[60,61]. However, damage to the hippocampus in humans does not simply produce problems with spatial navigation but rather produces much more profound deficits, preventing the formation of new episodic memories[62,63]. One way to reconcile the findings

in rodents and humans was to argue that the cognitive map was not just a spatial map. The same map could be used to specify how the different elements of a memory might be combined into a single episode, or it could specify how different spatial landmarks are combined to form a map of the environment[6,64-66].

One problem with this expanded role for the cognitive map is that it can be difficult to specify a priori what information the map should and should not contain[6]. There is a good deal of theoretical work that is currently focused on putting the cognitive map on a more formal footing. One approach has been to explore how animals can use Bayesian inference to infer how the world is structured[67,68]. Animals attempt to infer whether some evidence is consistent with an existing state of the world or implies the existence of a new state. Once states have been identified, the animal must then learn how they are related to one another. Here, formal models derived from graph theory have proved useful (FIG. 2a). Animals learn a graph (akin to a network) where each vertex (node in the graph) is a state of the world and the edges (the connections between vertices) specify the transition probabilities from one state of the world to another. The state-transition graph can be integrated with the reward location to calculate the value of individual states, which can then be used to guide optimal choice behaviour. From this point forward, we use ‘state-transition graph’ to refer to this network of states and ‘cognitive map’ to refer to the broader concept, which includes the integration of reward with the state-transition graph to determine optimal behaviour.

FIGURE 2b illustrates a graph that specifies the states that might describe a restaurant experience. Each arrow has a probability associated with it. For example, if we had to seat ourselves, we might expect a higher likelihood that we would have to order from an app rather than from a server. One advantage of learning these abstract transition structures is that the same graph can be applied to many different sensory problems[69]. Although restaurants look very different from one another, the way to navigate the restaurant experience can be distilled to the same graph. Note also that specific states are not necessarily defined by their sensory properties but by the temporal context or narrative in which they occur. At several points in the restaurant experience, you can be sat at an empty table but the appropriate response is to either ask the server for a menu or the bill depending on where you are in the sequence of states.

A second advantage of learning such state-transition graphs is that knowledge of the causal structure of the world is kept separate from value. This allows rapid changes in action selection when values change. Suppose that, over our dinner conversation, our dining companion inspires us with their healthy lifestyle. Our most valuable option shifts from choosing the high-calorie dessert to instead skipping dessert and getting the bill. We do not need to learn through trial-and-error which is the more valuable option but rather can use the state-transition graph to determine the appropriate response for obtaining the new, more valuable goal.

This distinction between trial-and-error learning and knowledge of state transitions has previously been described with respect to model-free and model-based reinforcement learning[49]. Model-free reinforcement does not have access to the state-transition graph and instead uses trial-and-error learning to estimate a single value, for each state, that reflects

the total reward that an agent can expect to earn in the future from that state, with rewards temporally closer to the agent weighted more highly than distant rewards. By contrast, model-based reinforcement learning iteratively searches through the state-transition graph to determine the path with the highest value, given the goal state. Model-free reinforcement learning is computationally simple but inflexible, whereas model-based reinforcement learning is flexible but computationally expensive — although this simple dichotomy has been criticized recently[51].

The successor representation is particularly relevant to the blurring of the lines between model-free and model-based learning algorithms[70-73]. Agents that incorporate the successor representation will derive advantages of both model-free and model-based learning. Parameters are cached during trial-and-error learning reducing computational complexity (as in model-free learning) but, rather than caching the value of states, the successor representation caches the likelihood of transitions between states, thereby incorporating a map-like structure (as in model-based learning). The separation between the learning of state transitions and rewards allows the agent to respond rapidly to changes in the value function (like model-based learning), although the agent can only learn changes in the state-transition graph slowly through trial-and-error (like model-free learning).

How do these ideas relate to the kinds of tasks that are dependent on the OFC? Consider a classic test of OFC function: stimulus-outcome reversal. In this task, a participant learns that one stimulus is associated with reward, whereas another stimulus is not. Once the participant has learned these associations, the experimenter reverses the contingencies. In a nice example of cross-species homology, rodents[74], monkeys[40,75] and humans[76] with OFC damage have no problem with the initial learning, but show severe impairments when the contingencies reverse, continuing to perseverate by choosing the previously rewarded outcome. The explanation for this deceptively simple deficit has undergone several revisions. Initial explanations were influenced by behaviourist theory. The learning process underlying the initial learning was assumed to be the same as the learning process following the reversal. As such, animals without OFC damage were proposed to learn the initial contingencies by trial-and-error, repeating rewarded choices and avoiding unrewarded choices, similar to a model-free learner. Following a reversal, the same process was suggested to allow them to learn the changed contingencies. To explain the critical paradox as to why the initial learning was intact but reversal learning was impaired in animals with OFC damage, vague concepts were invoked that did little more than describe the observed behaviour. For example, animals were argued to have a lack of ‘inhibitory control’ that prevented them from inhibiting the response to the previously rewarded option[40] or they were argued to lack the ‘behavioural flexibility’ that would allow them to change their choices[77].

Recent formulations of reversal learning have instead invoked the concept of a state-transition graph (FIG. 3). Importantly, in this formulation, reversal is not simply the unlearning of the original association and the learning of the new contingency. Instead, the participant learns that there are two states of the world, one in which stimulus A is rewarded and B is not, and one in which stimulus B is rewarded and A is not[78]. The OFC is argued to be crucial for this representation of the task[78]. This would explain why participants with

OFC damage are unimpaired on the original learning problem, which can be solved via simple associative learning, and why they are impaired on later reversals once the map of the task has been constructed and would speed learning in control participants[40]. However, recent neuropsychological studies in monkeys have suggested that the OFC might not be as crucial for reversal learning as originally thought[79,80]. It is possible that there are multiple factors that may contribute to whether the reversal task is learned through a cognitive map or through simpler strategies and, therefore, whether it requires the OFC. These include the length of training, whether there are concurrent discriminations and whether reward is delivered probabilistically[81].

To test the cognitive map hypothesis more directly, Schuck et al. devised a complex task that involved multiple contingencies to derive the correct answer[8]. Participants were presented with a picture of a house and a face, and they had to judge the age of one of the categories. When the age changed (for example, a young face followed by an old face) the participants had to switch to judging the other category. The task could be represented as a state-transition graph in which each unique combination of past and present category and perceptual configuration was defined as a task state (circles), with transitions between each state governed by the structure of the task (FIG. 4a). Importantly, the experimental conditions were balanced so that each state was equally likely to be correct at some point, ensuring that no one state was consistently more valuable than another. The authors trained a decoder to classify blood oxygen level-dependent (BOLD) activity based on which of the 16 task states the participant was currently in. Only a region in the ventromedial PFC (vmPFC), a region directly medial to the OFC, represented all the relevant task contingencies that would be necessary to specify a cognitive map (FIG. 4b), and the degree to which this information could be decoded from the vmPFC predicted participants' performance (FIG. 4c). We note that it is difficult to interpret negative neuroimaging results from the OFC since there is often a loss of signal due to susceptibility artefacts[82,83]; therefore, even though a signal was observed in the vmPFC, the study does not rule out a role for the OFC. Indeed, the authors concluded that a principal function of the OFC was to encode the cognitive map.

The encoding of a cognitive map seems to bear little relationship to value coding in the OFC described above. However, the preponderance of observations of value coding in the OFC described in the literature may be a self-fulfilling prophecy. Guided by the work in people with OFC damage, experimenters have tended to focus on very simple decision-making tasks in which the only parameters that were varied were related to value[20,84]. Under such circumstances, it is not surprising that value was the only parameter encoded by OFC neurons[4,9,38]. In fact, when experimenters have used more-complex tasks, OFC encoding has proven equally complex, yet these findings have often been ignored because they did not fit the dominant theoretical view of OFC function. For example, we showed that prefrontal neurons encoded abstract rules and that such activity was equally prevalent in the OFC as in other prefrontal areas[85]. In formal terms, rules specify the state-transition graph, consistent with the notion that the OFC encodes not just value, but rather the entire cognitive map. Similarly, OFC neurons encoded rules in the Wisconsin card sorting test[86], stay–shift strategies on a visuomotor conditional task[87], matches and mismatches during an olfactory-recognition task[88], and different blocks on an olfactory-choice task[89].

In addition, the state-transition graph is an important component of model-based learning, which, as discussed above, also seems to depend on OFC. The cognitive map might also be used to infer hypothetical outcomes. For example, consider a rock–paper–scissors game in which your opponent consistently chooses paper. A model-free learning system would have to iteratively select each outcome in turn to determine which outcome wins against paper. However, if we understand the state-transition graph of the game, we can choose the optimal scissors response immediately. Indeed, in monkeys trained on this game, OFC neurons encode not just the actual outcome of the animal's choice but also the hypothetical outcomes that would have resulted from making different choices[90], indicating that the OFC can use the state-transition graph to generate value predictions that have not actually been experienced.

Hypothetical outcomes are a key component of counterfactual emotions such as guilt, regret and relief. To generate such emotions, an animal must be able to infer what would have happened had they made a different choice. Counterfactual emotions can have a powerful sway on our decisions. For example, humans often deviate from rational decision theory because they try to minimize the likelihood of experiencing regret[91]. There is evidence that these counterfactual emotions depend on the OFC. For example, people with OFC lesions do not experience regret[92]. Rats that have rejected an average reward, only to be presented with a worse reward, will look back to the location of the rejected reward while their OFC neurons encode the value of the rejected reward[93].

In sum, the cognitive map hypothesis has the potential to explain a greater array of experimental findings from the OFC than can the value hypothesis. However, recall that the cognitive map has historically been associated with the hippocampus — which raises the question of what, if anything, is the OFC doing differently from the hippocampus?

The role of the hippocampus

The neural representation of the cognitive map was first associated with the hippocampus with the discovery of hippocampal place neurons[60], and the concept was later expanded to include abstract relational structures, in part to explain the role of the hippocampus in episodic memory in humans[64-66]. As described above, such abstract relationships are an important component of the state-transition graph because they enable the same graph to be used for very different sensory situations[6,69]. However, until recently, the empirical evidence for such relational encoding was rather weak. Many studies demonstrated that hippocampal neurons encoded non-spatial information[65,88,94] but they did not show the representational structure that would be necessary to demonstrate the existence of a map. A more recent study did examine the parametric coding of non-spatial information using a task that required rodents to navigate an auditory space[95]. However, in this case, the parametric representation was still that of a sensory stimulus rather than of an abstract cognitive parameter. Neuroimaging results have also provided evidence for relational coding. For example, when people are required to categorize stimuli that vary along two abstract dimensions, the hippocampus responds in a way that reflects the distance between the stimuli in this abstract two-dimensional space[96,97]. However, these results lack the resolution to determine what this coding looks like at the single-neuron level.

Value is one variable that is potentially useful for testing relational coding. It is an abstract, cognitive variable, yet it is also salient to animals. We developed a behavioural task that required monkeys to track the changing reward values associated with three pictures. We then recorded from the hippocampus during the performance of this task[98]. Hippocampal neurons encoded ‘value place fields’ that essentially specified the relationship between the values of the three reward-predictive pictures (FIG. 5a,b). Like space, value is relational[99,100]: we determine the value of an outcome relative to other outcomes. For example, we can experience a reward as negative if another choice would have led to an even larger reward[92]. Several researchers have noted that the tuning of hippocampal neurons seems to be optimized for encoding relationships[65,69,101]. Thus, it can be used to construct a spatial map of the location of objects relative to one another, or it can be used to construct a value map of the value of pictures relative to one another.

However, note that, in our experiment, we were using value to explore the relational code and, as such, we could not unconfound the state-transition graph from value, a point that we return to below. In fact, there is accumulating evidence that the hippocampus could also make an important contribution to value-based decision-making. Humans with hippocampal damage are impaired at making reward-based decisions[102]. In rats trained to run around a track on which reward was only delivered every four laps, different pools of hippocampal neurons would fire on each lap, potentially providing a mechanism to track the individual episodes that lead to reward[103]. Hippocampal replay events, in which place neurons fire in an ordered sequence, predict the distance to a reward location during both hippocampal theta oscillations[104] and during sharp-wave ripples[105].

Thus, we have two brain regions that strongly connect with one another[17,106] and seem to share many of the same functions. In addition, there is now considerable evidence that the two areas functionally interact with one another during the performance of many behavioural tasks. For example, we observed increased coherence in the theta oscillation between the OFC and the hippocampus when animals were adapting their choice behaviour in response to changing reward contingencies[107]. In addition, when human participants perform a complex, sequential decision-making task[8], hippocampal BOLD responses are consistent with the sequential replay of participants’ experience with the task and this, in turn, leads to an improved representation of the task in the OFC and better performance[108].

Causal manipulations have also revealed the importance of hippocampal input to the OFC. A prominent theta oscillation occurs in the OFC when animals learn the significance of reward-predictive cues[109]. To demonstrate the importance of this oscillation, we used closed-loop microstimulation[107]. We recorded the theta oscillation in real-time in monkeys and used it to control the application of microstimulation to the OFC. Stimulation applied at the peak of theta severely disrupted the oscillation as well as the ability of the animal to learn new reward contingencies. Because the theta oscillation probably has a hippocampal origin[110], we also applied closed-loop theta microstimulation to the hippocampus and similarly severely disrupted reward-based learning. Results in rodents have also demonstrated the importance of OFC–hippocampal interaction in learning. Optogenetic inactivation of the hippocampus impairs the OFC representation of the block structure of an olfactory-choice task[111].

However, despite the similarities in OFC and hippocampal functions, it is also clear that neuronal coding in the hippocampus is very different from that in the OFC. This is most apparent when comparing neuronal responses from the two areas during the performance of identical behavioural tasks. For example, we directly compared OFC and hippocampal neuronal responses on the reward-learning task described above[107]. OFC neurons typically encode the value of the chosen option in a linear way (FIG. 5c). Consequently, they are active at most points along the circular trajectory through value space, rather than at specific locations like hippocampal neurons. This results in a sparser representation of value among hippocampal neurons than among OFC neurons that can be quantified using information theory (FIG. 5d).

Along similar lines, McKenzie et al. trained rats on a task in which different environmental contexts predicted which rewards would be paired with which objects. Although all the information that was relevant to performing the task was present in both the OFC and hippocampus, context dominated neural encoding in the hippocampus[112], whereas value dominated neural encoding in the OFC[113]. Similar results have been obtained using other tasks. For example, in rodent studies of model-based learning of olfactory discriminations, although all the information relevant to performing the task is represented in the OFC, value explains by far the most variance in OFC neuronal firing[114]. Thus, an important outstanding issue is what exactly the specific contributions of the two brain areas are; we discuss this next.

Reconciling the hypotheses

At this point, we return to considering the state of the value hypothesis and the cognitive map hypothesis. One could be forgiven for assuming everything is settled. Researchers were like the fabled blind men, each feeling a different part of an elephant and coming to a different conclusion as to what it was. Hippocampal researchers typically used the rodent model system, in part because the hippocampus is easier to access in the rodent than in the primate. Rodents were tested in arenas and mazes and, therefore, space dominated the neural representation. By contrast, researchers studying the OFC often used monkeys, who sat, like gamblers in a casino, in a single position playing hundreds of trials to win rewards. Little wonder that, in this case, value dominated the neural representation. By taking a broader view of both areas, we can see that they are both responsible for implementing the cognitive map[115,116]. However, one problem with such a neat reconciliation is that the pendulum may have swung too far, with the OFC now being ascribed too broad a role.

An important consideration for any theory is whether it has predictive validity. A prediction of the cognitive map hypothesis is that any sufficiently complex task should benefit from a cognitive map and therefore require the OFC; however, this is not always the case. For example, consider a task designed by Baxter et al.[117]. Monkeys learned to apply one of two choice-response strategies to two sets of stimuli. Pictures in one set of stimuli (persistent) had to be chosen four times in a row to earn a reward, whereas choices for pictures in the second set (sporadic) were rewarded only after a persistent reward had been earned and were not rewarded again until another persistent reward had been earned. The task is conceptually like the Schuck et al.[8] task shown in FIG. 4a, in that different stimuli instructed different

responses but were equally likely to be correct depending on the experimental condition. Despite this similarity, animals with OFC lesions who had learned the task prior to the lesion could perform it without impairments. This is difficult to reconcile with the idea that the OFC is representing a cognitive map to facilitate performance of the task.

A more promising direction has been to focus on a potential division of labour between the OFC and hippocampus[116]. One possibility is that the hippocampus represents the state-transition graph, whereas the OFC uses knowledge of the reward's location to calculate the value of the vertices in the graph, thereby helping to guide the optimal choice at each decision point. A recent study by Basu et al. showed that single-neuron responses in the OFC were consistent with this function[118]. Rats were trained to forage between two locations on a linear track. OFC neurons were tuned to the location of the reward, and their firing rates gradually increased as the animal approached the rewarded location. The linear track can be described via a linear state-transition graph, and the OFC neuronal responses are consistent with encoding the value of those states. Furthermore, optogenetic inactivation of these neurons impaired the performance of the task.

Another task that is relevant to testing the 'division of labour' hypothesis is sensory preconditioning. Participants learn that certain states of the world co-occur and then use this knowledge to infer predictions about the value of potential outcomes. For example, if I learn that cue A and cue B occur together, and then I learn that A predicts reward, I might rightly infer that B will also predict reward. There is a growing body of work to support the notion that this type of learning is dependent on OFC–hippocampal interactions. For example, in rodents, OFC neurons encode cue associations as well as inferred outcomes[119], and pharmacological inactivation of the OFC specifically impairs learning that is based on the inferred outcomes[120]. Neuroimaging results in humans have shown that the BOLD response to paired cues (that is, to A and to B) becomes more similar in both the hippocampus and the OFC as the network of associations is learned[121]. Furthermore, the OFC represented the inferred outcome when cue B was presented, and this was accompanied by an increase in connectivity between the OFC and the hippocampus. Causal studies have also favoured the interpretation that the OFC uses the cognitive map to generate reward predictions. For example, transcranial magnetic stimulation of the OFC in humans specifically impairs the ability to infer outcomes from cue associations in the sensory preconditioning task[122].

Returning to the Baxter et al. task described above[117] with the idea that the OFC does not represent the cognitive map per se but instead computes the value at different vertices in the map, it becomes clearer why the complex task may not require the OFC. The task can be specified as a state-transition graph, but the value of each vertex remains fixed throughout the performance of the task and, therefore, the OFC is not needed to update the values. Other lesion results support this conclusion. Selective lesions of the PFC in macaques performing a version of the Wisconsin card sorting test demonstrate a double dissociation. Damage to the dorsolateral PFC leads to a problem in holding the state-transition graph in working memory, whereas damage to the OFC induces problems with rapidly updating the value of the graph nodes based on reward outcomes[123].

We note that we have remained agnostic to the precise nature of the interaction between the OFC and the hippocampus. Given the role of the hippocampus in forming new memories[124-126], we might predict that it would be important for learning the state-transition graph, which would then be used by the OFC to determine the value of the graph vertices. However, it is also possible that information from the OFC could be used to modify the state-transition graph in response to violated reward expectations. For example, neuroimaging has shown that OFC learning signals predict changes in the strength of the hippocampal representation of state-transition probabilities[127].

Given the proposed distinction between the OFC and the hippocampus proposed above, it is worth returning to how well it aligns with the value hypothesis and the cognitive map hypothesis. With respect to the value hypothesis, the value framework clearly only captures a portion of OFC function. However, many of the problems that are tackled by the value hypothesis are also relevant to the cognitive map hypothesis. For example, it remains unclear how an animal assigns a value to a specific outcome, and this value is an important component of calculating the value of the states in the state-transition graph. Likewise, the cognitive map requires animals to choose the optimal motor response based on the value of the different resultant states; the neural implementation of this process also remains unclear.

Turning to the cognitive map hypothesis, the ‘hard’ version of this hypothesis, that the OFC is responsible for its entire implementation, seems like too broad a role to adequately account for the experimental evidence. However, by limiting its role to calculating values of states using the state-transition graph, we can preserve the primary role of the OFC of encoding value information while also accounting for the role of the OFC in encoding state transitions. Although we argue that the hippocampus is primarily responsible for learning and modifying the state-transition graph, we also note that many researchers have argued that the hippocampus is responsible for initial learning of state-transition graphs but not for their long-term storage, which is argued to be the preserve of the cortex[124-126]. This raises the possibility that other cortical regions, such as the lateral PFC, may be responsible for the long-term storage of the state transition graph.

Future directions

In this final section, we outline a personal wish list of experimental directions. First, if the cognitive map hypothesis of OFC function is to challenge the value hypothesis, it is important that we establish the same breadth of experimental evidence to ensure a solid foundation on which to build this new idea. Most important is to establish causal evidence, particularly in humans. Studies of individuals with OFC damage remain largely focused on the value hypothesis. For example, it would be fascinating to know how people with OFC damage perform on the same tasks that involve cognitive maps and that drive strong OFC BOLD responses[8].

Second, it seems reasonable that the computational principles that are currently driving theoretical work on the cognitive map will ultimately prove useful for understanding its implementation in the brain, including in the OFC. For example, it has been suggested that the entorhinal cortex may be responsible for encoding the state-transition graph, whereas the

hippocampus then combines this information with specific sensory input, thereby enabling different maps to generalize to novel experiences[69]. Similarly, understanding how animals parse knowledge and infer the existence of different states of the world[128] will probably be extremely useful in interpreting behaviour and, therefore, the neural implementation of behaviour. For example, behavioural evidence suggests that bottlenecks in the state-transition graph can be used to organize knowledge, and this organization is reflected in the neural representation in frontal and temporal cortices[129] and the hippocampus[130].

Third, regarding the value hypothesis, we need to improve the ecological validity of our measurements. Real-world decisions often involve unique circumstances and information. By their nature they are ‘one-shot’. This makes them challenging to study in the laboratory, certainly with neural measurements and even behaviourally. Our measurements are noisy and frequently require multiple trials. Fortunately, modern neuroscience methods are beginning to provide us with access to the neural mechanisms that underlie one-shot cognitive processes. In particular, innovative recording probes open up the possibility of recording from thousands of OFC neurons simultaneously[131], and decoding algorithms provide the methods to measure the neural representation with single-trial resolution[56]. In addition, we clearly need to expand the range of tasks used to investigate OFC function, at least beyond the simple binary choice tasks that many of us study.

It is an exciting time to be researching the OFC. After a decade or so of somewhat incremental progress, the past few years have seen the emergence of dramatic new ideas regarding OFC function. Yet, despite this, it seems premature to completely abandon the value hypothesis, which can explain so many OFC findings. Instead, it seems likely that the infusion of new theoretical ideas from formal models of the cognitive map will help to better align our understanding of OFC value representations to naturalistic behaviour and ultimately explain why OFC damage is so catastrophic for everyday decision-making.

Acknowledgements

J.D.W. was supported by grants from the US National Institutes of Health (NIH) R01-MH117763, R01-MH121448 and R01-NS116623 during the writing of this manuscript.

Glossary

Secondary reinforcer
A reward or punishment whose value is learned (for example, money) through its association with a primary reinforcer whose value is innate (for example, food).

Repetition suppression
A reduction in the magnitude of the evoked blood oxygen level-dependent response when a stimulus is presented repeatedly.

Reinforcement learning
The process by which an agent learns to predict and maximize future reward.

Inference
The process of deriving logical conclusions from known premises.

Decision primitives
Parameters that are combined to calculate the value of a reward; for example, a food reward might include size, probability of occurrence and calories.

Episodic memories
Memories of personal experiences that are tied to specific times and places.

Bayesian inference
A statistical approach that uses Bayes theorem to determine how much to update one's belief given a new piece of evidence.

Graph theory
A branch of mathematics that focuses on understanding networks. A graph consists of vertices (also called nodes) that are connected by edges (also called lines).

Successor representation
A map of the environment that estimates the predictive relationships between different states of the environment.

Perseverate
To continue to repeat a previously rewarded action even when it no longer leads to reward.

Behaviourist theory
The theory that psychology can be objectively studied only through observable actions; it arose as a reaction to nineteenth-century psychology which focused on introspection.

Susceptibility artefacts
Artefacts that occur during MRI at air–tissue boundaries; they are particularly serious for brain areas close to sinuses.

Wisconsin card sorting test
A neuropsychological test in which participants sort cards according to rules such as shape or colour. Patients with frontal lobe damage have difficulty switching between rules.

Visuomotor conditional task
A task that requires subjects to follow a conditional 'if–then' rule, in which the 'if' is a visual stimulus and the 'then' is a motor response.

Rock–paper–scissors game
Two players simultaneously make one of three hand shapes: rock, paper or scissors. Rock beats scissors, scissors beat paper, and paper beats rock.

Place neurons
Hippocampal neurons that fire whenever an animal is in a specific location.

Sharp-wave ripples
Oscillations that are characteristics of electrical activity in the mammalian hippocampus: they are of large amplitude and high frequency (100–250 Hz).

Ecological validity

The degree to which a laboratory test predicts behaviour in real-world settings.

References

1. Damasio AR Descartes' Error: Emotion, Reason, and the Human Brain (Putman, 1994).
2. Eslinger PJ & Damasio AR Severe disturbance of higher cognition after bilateral frontal lobe ablation: patient EVR. Neurology 35, 1731–1741 (1985). [PubMed: 4069365]
3. Bechara A, Damasio AR, Damasio H & Anderson SW Insensitivity to future consequences following damage to human prefrontal cortex. Cognition 50, 7–15 (1994). [PubMed: 8039375]
4. Padoa-Schioppa C & Assad JA Neurons in the orbitofrontal cortex encode economic value. Nature 441, 223–226 (2006). [PubMed: 16633341]
5. Tolman EC Cognitive maps in rats and men. Psychol. Rev 55, 189–208 (1948). [PubMed: 18870876]
6. Behrens TEJ et al. What is a cognitive map? Organizing knowledge for flexible behavior. Neuron 100, 490–509 (2018). [PubMed: 30359611] This review explains how cognitive maps might be used to organize knowledge across multiple domains, not just space.
7. O'Keefe J & Nadel L The Hippocampus as a Cognitive Map (Oxford University Press, 1978).
8. Schuck NW, Cai MB, Wilson RC & Niv Y Human orbitofrontal cortex represents a cognitive map of state space. Neuron 91, 1402–1412 (2016). [PubMed: 27657452] A human neuroimaging study showing that the vmPFC might be particularly important for representing the cognitive map.
9. Kennerley SW, Dahmubed AF, Lara AH & Wallis JD Neurons in the frontal lobe encode the value of multiple decision variables. J. Cogn. Neurosci 21, 1162–1178 (2009). [PubMed: 18752411]
10. Hosokawa T, Kennerley SW, Sloan J & Wallis JD Single-neuron mechanisms underlying cost-benefit analysis in frontal cortex. J. Neurosci 33, 17385–17397 (2013). [PubMed: 24174671]
11. Roesch MR & Olson CR Neuronal activity in primate orbitofrontal cortex reflects the value of time. J. Neurophysiol 94, 2457–2471 (2005). [PubMed: 15958600]
12. Rich EL & Wallis JD Medial-lateral organization of the orbitofrontal cortex. J. Cogn. Neurosci 26, 1347–1362 (2014). [PubMed: 24405106]
13. Morrison SE & Salzman CD The convergence of information about rewarding and aversive stimuli in single neurons. J. Neurosci 29, 11471–11483 (2009). [PubMed: 19759296]
14. Roesch MR & Olson CR Neuronal activity related to reward value and motivation in primate frontal cortex. Science 304, 307–310 (2004). [PubMed: 15073380]
15. Wallis JD & Rich EL Challenges of interpreting frontal neurons during value-based decision-making. Front. Neurosci 5, 124 (2011). [PubMed: 22125508]
16. Kennerley SW, Behrens TE & Wallis JD Double dissociation of value computations in orbitofrontal and anterior cingulate neurons. Nat. Neurosci 14, 1581–1589 (2011). [PubMed: 22037498]
17. Carmichael ST & Price JL Limbic connections of the orbital and medial prefrontal cortex in macaque monkeys. J. Comp. Neurol 363, 615–641 (1995). [PubMed: 8847421]
18. Carmichael ST & Price JL Sensory and premotor connections of the orbital and medial prefrontal cortex of macaque monkeys. J. Comp. Neurol 363, 642–664 (1995). [PubMed: 8847422]
19. Cavada C, Company T, Tejedor J, Cruz-Rizzolo RJ & Reinoso-Suarez F The anatomical connections of the macaque monkey orbitofrontal cortex. A review. Cereb. Cortex 10, 220–242 (2000). [PubMed: 10731218]
20. Padoa-Schioppa C Neurobiology of economic choice: a good-based model. Annu. Rev. Neurosci 34, 333–359 (2011). [PubMed: 21456961]
21. Siep N et al. Hunger is the best spice: an fMRI study of the effects of attention, hunger and calorie content on food reward processing in the amygdala and orbitofrontal cortex. Behav. Brain Res 198, 149–158 (2009). [PubMed: 19028527]
22. Barron HC, Dolan RJ & Behrens TE Online evaluation of novel choices by simultaneous representation of multiple memories. Nat. Neurosci 16, 1492–1498 (2013). [PubMed: 24013592]

23. Hare TA, O'Doherty J, Camerer CF, Schultz W & Rangel A Dissociating the role of the orbitofrontal cortex and the striatum in the computation of goal values and prediction errors. J. Neurosci 28, 5623–5630 (2008). [PubMed: 18509023]
24. Sescousse G, Redoute J & Dreher JC The architecture of reward value coding in the human orbitofrontal cortex. J. Neurosci 30, 13095–13104 (2010). [PubMed: 20881127]
25. Li Y, Sescousse G, Amiez C & Dreher JC Local morphology predicts functional organization of experienced value signals in the human orbitofrontal cortex. J. Neurosci 35, 1648–1658 (2015). [PubMed: 25632140]
26. Peters J & Buchel C Overlapping and distinct neural systems code for subjective value during intertemporal and risky decision making. J. Neurosci 29, 15727–15734 (2009). [PubMed: 20016088]
27. Fellows LK & Farah MJ The role of ventromedial prefrontal cortex in decision making: Judgment under uncertainty or judgment per se? Cereb. Cortex 17, 2669–2674 (2007). [PubMed: 17259643]
28. Fellows LK Deciding how to decide: ventromedial frontal lobe damage affects information acquisition in multi-attribute decision making. Brain 129, 944–952 (2006). [PubMed: 16455794]
29. Xia C, Stolle D, Gidengil E & Fellows LK Lateral orbitofrontal cortex links social impressions to political choices. J. Neurosci 35, 8507–8514 (2015). [PubMed: 26041918]
30. Camille N, Griffiths CA, Vo K, Fellows LK & Kable JW Ventromedial frontal lobe damage disrupts value maximization in humans. J. Neurosci 31, 7527–7532 (2011). [PubMed: 21593337]
31. Ballesta S, Shi W, Conen KE & Padoa-Schioppa C Values encoded in orbitofrontal cortex are causally related to economic choices. Nature 10.1038/s41586-020-2880-x (2020). Causal evidence for the role of the OFC in economic choice using electrical microstimulation.
32. Gardner MPH et al. Processing in lateral orbitofrontal cortex is required to estimate subjective preference during initial, but not established, economic choice. Neuron 108, 526–537.e4 (2020). [PubMed: 32888408]
33. Rustichini A & Padoa-Schioppa C A neurocomputational model of economic decisions. J. Neurophysiol 114, 1382–1398 (2015). [PubMed: 26063776]
34. Hunt LT et al. Mechanisms underlying cortical activity during value-guided choice. Nat. Neurosci 15, 470–476 (2012). [PubMed: 22231429]
35. Cai X & Padoa-Schioppa C Contributions of orbitofrontal and lateral prefrontal cortices to economic choice and the good-to-action transformation. Neuron 81, 1140–1151 (2014). [PubMed: 24529981]
36. Hunt LT, Behrens TE, Hosokawa T, Wallis JD & Kennerley SW Capturing the temporal evolution of choice across prefrontal cortex. eLife 10.7554/eLife. 11945 (2015).
37. Yim MY, Cai X & Wang XJ Transforming the choice outcome to an action plan in monkey lateral prefrontal cortex: a neural circuit model. Neuron 103, 520–532.e5 (2019). [PubMed: 31230761]
38. Padoa-Schioppa C Neuronal origins of choice variability in economic decisions. Neuron 80, 1322–1336 (2013). [PubMed: 24314733]
39. Rich EL & Wallis JD Decoding subjective decisions from orbitofrontal cortex. Nat. Neurosci 19, 973–980 (2016). [PubMed: 27273768] Study that first measured population-level dynamics in the OFC during decision-making with single-trial resolution.
40. Iversen SD & Mishkin M Perseverative interference in monkeys following selective lesions of the inferior prefrontal convexity. Exp. Brain Res 11, 376–386 (1970). [PubMed: 4993199]
41. Baxter MG, Parker A, Lindner CC, Izquierdo AD & Murray EA Control of response selection by reinforcer value requires interaction of amygdala and orbital prefrontal cortex. J. Neurosci 20, 4311–4319 (2000). [PubMed: 10818166]
42. Roberts AC & Wallis JD Inhibitory control and affective processing in the prefrontal cortex: neuropsychological studies in the common marmoset. Cereb. Cortex 10, 252–262 (2000). [PubMed: 10731220]
43. Izquierdo A, Suda RK & Murray EA Bilateral orbital prefrontal cortex lesions in rhesus monkeys disrupt choices guided by both reward value and reward contingency. J. Neurosci 24, 7540–7548 (2004). [PubMed: 15329401]
44. Pickens CL et al. Different roles for orbitofrontal cortex and basolateral amygdala in a reinforcer devaluation task. J. Neurosci 23, 11078–11084 (2003). [PubMed: 14657165]

45. Kourtzi Z & Kanwisher N Representation of perceived object shape by the human lateral occipital complex. Science 293, 1506–1509 (2001). [PubMed: 11520991]
46. Stalnaker TA, Cooch NK & Schoenbaum G What the orbitofrontal cortex does not do. Nat. Neurosci 18, 620–627 (2015). [PubMed: 25919962]
47. Padoa-Schioppa C & Schoenbaum G Dialogue on economic choice, learning theory, and neuronal representations. Curr. Opin. Behav. Sci 5, 16–23 (2015). [PubMed: 26613099]
48. Sutton RS & Barto AG Toward a modern theory of adaptive networks: expectation and prediction. Psychol. Rev 88, 135–170 (1981). [PubMed: 7291377]
49. Daw ND, Niv Y & Dayan P Uncertainty-based competition between prefrontal and dorsolateral striatal systems for behavioral control. Nat. Neurosci 8, 1704–1711 (2005). [PubMed: 16286932]
50. Doll BB, Simon DA & Daw ND The ubiquity of model-based reinforcement learning. Curr. Opin. Neurobiol 22, 1075–1081 (2012). [PubMed: 22959354]
51. Collins AGE & Cockburn J Beyond dichotomies in reinforcement learning. Nat. Rev. Neurosci 21, 576–586 (2020). [PubMed: 32873936] An interesting review that warns against the simple dichotomy of model-free versus model-based mechanisms that has come to dominate neuroscientific studies of reinforcement learning.
52. Gremel CM & Costa RM Orbitofrontal and striatal circuits dynamically encode the shift between goal-directed and habitual actions. Nat. Commun 4, 2264 (2013). [PubMed: 23921250]
53. McDannald MA, Lucantonio F, Burke KA, Niv Y & Schoenbaum G Ventral striatum and orbitofrontal cortex are both required for model-based, but not model-free, reinforcement learning. J. Neurosci 31, 2700–2705 (2011). [PubMed: 21325538]
54. Redish AD Vicarious trial and error. Nat. Rev. Neurosci 17, 147–159 (2016). [PubMed: 26891625]
55. Kennerley SW & Wallis JD Evaluating choices by single neurons in the frontal lobe: outcome value encoded across multiple decision variables. Eur. J. Neurosci 29, 2061–2073 (2009). [PubMed: 19453638]
56. Wallis JD Decoding cognitive processes from neural ensembles. Trends Cogn. Sci 10.1016/j.tics.2018.09.002 (2018).
57. Klein-Flugge MC, Barron HC, Brodersen KH, Dolan RJ & Behrens TE Segregated encoding of reward-identity and stimulus-reward associations in human orbitofrontal cortex. J. Neurosci 33, 3202–3211 (2013). [PubMed: 23407973]
58. Burke KA, Franz TM, Miller DN & Schoenbaum G The role of the orbitofrontal cortex in the pursuit of happiness and more specific rewards. Nature 454, 340–344 (2008). [PubMed: 18563088]
59. Botvinick M, Wang JX, Dabney W, Miller KJ & Kurth-Nelson Z Deep reinforcement learning and its neuroscientific implications. Neuron 107, 603–616 (2020). [PubMed: 32663439]
60. O'Keefe J & Dostrovsky J The hippocampus as a spatial map. Preliminary evidence from unit activity in the freely-moving rat. Brain Res. 34, 171–175 (1971). [PubMed: 5124915]
61. Moser EI, Kropff E & Moser MB Place cells, grid cells, and the brain's spatial representation system. Annu. Rev. Neurosci 31, 69–89 (2008). [PubMed: 18284371]
62. Scoville WB & Milner B Loss of recent memory after bilateral hippocampal lesions. J. Neurol. Neurosurg. Psychiatry 20, 11–21 (1957). [PubMed: 13406589]
63. Eichenbaum H What H.M. taught us. J. Cogn. Neurosci 25, 14–21 (2013). [PubMed: 22905817]
64. Cohen NJ & Eichenbaum H The theory that wouldn't die: a critical look at the spatial mapping theory of hippocampal function. Hippocampus 1, 265–268 (1991). [PubMed: 1669304]
65. Howard MW et al. A unified mathematical framework for coding time, space, and sequences in the hippocampal region. J. Neurosci 34, 4692–4707 (2014). [PubMed: 24672015]
66. Eichenbaum H & Cohen NJ Can we reconcile the declarative memory and spatial navigation views on hippocampal function? Neuron 83, 764–770 (2014). [PubMed: 25144874]
67. Courville AC, Daw ND & Touretzky DS Bayesian theories of conditioning in a changing world. Trends Cogn. Sci 10, 294–300 (2006). [PubMed: 16793323]
68. Gershman SJ, Blei DM & Niv Y Context, learning, and extinction. Psychol. Rev 117, 197–209 (2010). [PubMed: 20063968]

69. Whittington JCR et al. The Tolman-Eichenbaum machine: unifying space and relational memory through generalization in the hippocampal formation. Cell 183, 1249–1263.e23 (2020). [PubMed: 33181068]
70. Dayan P Improving generalization for temporal difference learning: the successor representation. Neural Comput. 5, 613–624 (1993).
71. Stachenfeld KL, Botvinick MM & Gershman SJ The hippocampus as a predictive map. Nat. Neurosci 20, 1643–1653 (2017). [PubMed: 28967910] First demonstration that the firing properties of hippocampal place neurons are consistent with encoding the successor representation.
72. Momennejad I Learning structures: predictive representations, replay, and generalization. Curr. Opin. Behav. Sci 32, 155–166 (2020). [PubMed: 35419465]
73. Gershman SJ The successor representation: its computational logic and neural substrates. J. Neurosci 38, 7193–7200 (2018). [PubMed: 30006364]
74. Schoenbaum G, Setlow B, Nugent SL, Saddoris MP & Gallagher M Lesions of orbitofrontal cortex and basolateral amygdala complex disrupt acquisition of odor-guided discriminations and reversals. Learn. Mem 10, 129–140 (2003). [PubMed: 12663751]
75. Dias R, Robbins TW & Roberts AC Dissociation in prefrontal cortex of affective and attentional shifts. Nature 380, 69–72 (1996). [PubMed: 8598908]
76. Fellows LK & Farah MJ Ventromedial frontal cortex mediates affective shifting in humans: evidence from a reversal learning paradigm. Brain 126, 1830–1837 (2003). [PubMed: 12821528]
77. Clarke H, Walker S, Dalley J, Robbins T & Roberts A Cognitive inflexibility after prefrontal serotonin depletion is behaviorally and neurochemically specific. Cereb. Cortex 17, 18–27 (2007). [PubMed: 16481566]
78. Wilson RC, Takahashi YK, Schoenbaum G & Niv Y Orbitofrontal cortex as a cognitive map of task space. Neuron 81, 267–279 (2014). [PubMed: 24462094]
79. Rudebeck PH & Murray EA Dissociable effects of subtotal lesions within the macaque orbital prefrontal cortex on reward-guided behavior. J. Neurosci 31 , 10569–10578 (2011). [PubMed: 21775601]
80. Rudebeck PH, Saunders RC, Prescott AT, Chau LS & Murray EA Prefrontal mechanisms of behavioral flexibility, emotion regulation and value updating. Nat. Neurosci 16, 1140–1145 (2013). [PubMed: 23792944]
81. Izquierdo A, Brigman JL, Radke AK, Rudebeck PH & Holmes A The neural basis of reversal learning: an updated perspective. Neuroscience 345, 12–26 (2017). [PubMed: 26979052]
82. Glover GH & Law CS Spiral-in/out BOLD fMRI for increased SNR and reduced susceptibility artifacts. Magn. Reson. Med 46, 515–522 (2001). [PubMed: 11550244]
83. Wallis JD Cross-species studies of orbitofrontal cortex and value-based decision-making. Nat. Neurosci 15, 13–19 (2012).
84. Wallis JD & Kennerley SW Heterogeneous reward signals in prefrontal cortex. Curr. Opin. Neurobiol 20, 191–198 (2010). [PubMed: 20303739]
85. Wallis JD, Anderson KC & Miller EK Single neurons in prefrontal cortex encode abstract rules. Nature 411, 953–956 (2001). [PubMed: 11418860]
86. Sleezer BJ, Castagno MD & Hayden BY Rule encoding in orbitofrontal cortex and striatum guides selection. J. Neurosci 36, 11223–11237 (2016). [PubMed: 27807165]
87. Tsujimoto S, Genovesio A & Wise SP Comparison of strategy signals in the dorsolateral and orbital prefrontal cortex. J. Neurosci 31, 4583–4592 (2011). [PubMed: 21430158]
88. Ramus SJ & Eichenbaum H Neural correlates of olfactory recognition memory in the rat orbitofrontal cortex. J. Neurosci 20, 8199–8208 (2000). [PubMed: 11050143]
89. Stalnaker TA, Raheja N & Schoenbaum G Orbitofrontal state representations are related to choice adaptations and reward predictions. J. Neurosci 41, 1941–1951 (2021). [PubMed: 33446521]
90. Abe H & Lee D Distributed coding of actual and hypothetical outcomes in the orbital and dorsolateral prefrontal cortex. Neuron 70, 731–741 (2011). [PubMed: 21609828]
91. Bell DE Regret in decision making under uncertainty. Oper. Res 30, 961–981 (1982).
92. Camille N et al. The involvement of the orbitofrontal cortex in the experience of regret. Science 304, 1167–1170 (2004). [PubMed: 15155951]

93. Steiner AP & Redish AD Behavioral and neurophysiological correlates of regret in rat decision-making on a neuroeconomic task. Nat. Neurosci 17, 995–1002 (2014). [PubMed: 24908102]

94. Wood ER, Dudchenko PA & Eichenbaum H The global record of memory in hippocampal neuronal activity. Nature 397, 613–616 (1999). [PubMed: 10050854]

95. Aronov D, Nevers R & Tank DW Mapping of a non-spatial dimension by the hippocampal-entorhinal circuit. Nature 543, 719–722 (2017). [PubMed: 28358077]

96. Theves S, Fernandez G & Doeller CF The hippocampus encodes distances in multidimensional feature space. Curr. Biol 29, 1226–1231.e3 (2019). [PubMed: 30905602]

97. Park SA, Miller DS, Nili H, Ranganath C & Boorman ED Map making: constructing, combining, and inferring on abstract cognitive maps. Neuron 107, 1226–1238.e8 (2020). [PubMed: 32702288]

98. Knudsen EB & Wallis JD Hippocampal neurons construct a map of an abstract value space. Cell 184, 4640–4650.e10 (2021). [PubMed: 34348112] Study showing that hippocampal neurons also encode abstract, cognitive information using the same relational code that the hippocampus uses for encoding space.

99. Rangel A & Clithero JA Value normalization in decision making: theory and evidence. Curr. Opin. Neurobiol 22, 970–981 (2012). [PubMed: 22939568]

100. Louie K, Grattan LE & Glimcher PW Reward value-based gain control: divisive normalization in parietal cortex. J. Neurosci 31, 10627–10639 (2011). [PubMed: 21775606]

101. Eichenbaum H, Dudchenko P, Wood E, Shapiro M & Tanila H The hippocampus, memory, and place cells: is it spatial memory or a memory space? Neuron 23, 209–226 (1999). [PubMed: 10399928]

102. Bakkour A et al. The hippocampus supports deliberation during value-based decisions. eLife 10.7554/eLife.46080 (2019).

103. Sun C, Yang W, Martin J & Tonegawa S Hippocampal neurons represent events as transferable units of experience. Nat. Neurosci 23, 651–663 (2020). [PubMed: 32251386]

104. Wikenheiser AM & Redish AD Hippocampal theta sequences reflect current goals. Nat. Neurosci 18, 289–294 (2015). [PubMed: 25559082]

105. Pfeiffer BE & Foster DJ Hippocampal place-cell sequences depict future paths to remembered goals. Nature 497, 74–79 (2013). [PubMed: 23594744]

106. Barbas H & Blatt GJ Topographically specific hippocampal projections target functionally distinct prefrontal areas in the rhesus monkey. Hippocampus 5, 511–533 (1995). [PubMed: 8646279]

107. Knudsen EB & Wallis JD Closed-loop theta stimulation in the orbitofrontal cortex prevents reward-based learning. Neuron 106, 537–547.e4 (2020). [PubMed: 32160515]

108. Schuck NW & Niv Y Sequential replay of nonspatial task states in the human hippocampus. Science 10.1126/science.aaw5181 (2019).

109. van Wingerden M, Vinck M, Lankelma JV & Pennartz CM Learning-associated gamma-band phase-locking of action-outcome selective neurons in orbitofrontal cortex. J. Neurosci 30, 10025–10038 (2010). [PubMed: 20668187]

110. Buzsaki G Theta oscillations in the hippocampus. Neuron 33, 325–340 (2002). [PubMed: 11832222]

111. Wikenheiser AM, Marrero-Garcia Y & Schoenbaum G Suppression of ventral hippocampal output impairs integrated orbitofrontal encoding of task structure. Neuron 95, 1197–1207.e3 (2017). [PubMed: 28823726]

112. McKenzie S et al. Hippocampal representation of related and opposing memories develop within distinct, hierarchically organized neural schemas. Neuron 83, 202–215 (2014). [PubMed: 24910078]

113. Farovik A et al. Orbitofrontal cortex encodes memories within value-based schemas and represents contexts that guide memory retrieval. J. Neurosci 35, 8333–8344 (2015). [PubMed: 26019346]

114. 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). [PubMed: 30827919]

115. Niv Y Learning task-state representations. Nat. Neurosci 22, 1544–1553 (2019). [PubMed: 31551597]

116. Wikenheiser AM & Schoenbaum G Over the river, through the woods: cognitive maps in the hippocampus and orbitofrontal cortex. Nat. Rev. Neurosci 17, 513–523 (2016). [PubMed: 27256552] Paper that proposed that the OFC might be important for integrating reward information with the hippocampal map.

117. Baxter MG, Gaffan D, Kyriazis DA & Mitchell AS Ventrolateral prefrontal cortex is required for performance of a strategy implementation task but not reinforcer devaluation effects in rhesus monkeys. Eur. J. Neurosci 29, 2049–2059 (2009). [PubMed: 19453635]

118. Basu R et al. The orbitofrontal cortex maps future navigational goals. Nature 599, 449–452 (2021). [PubMed: 34707289] Recent neurophysiological study showing that OFC neurons potentially encode the value of specific states in the state-transition graph.

119. Sadacca BF et al. Orbitofrontal neurons signal sensory associations underlying model-based inference in a sensory preconditioning task. eLife 10.7554/eLife.30373 (2018).

120. Jones JL et al. Orbitofrontal cortex supports behavior and learning using inferred but not cached values. Science 338, 953–956 (2012). [PubMed: 23162000]

121. Wang F, Schoenbaum G & Kahnt T Interactions between human orbitofrontal cortex and hippocampus support model-based inference. PLoS Biol. 18, e3000578 (2020). [PubMed: 31961854]

122. Wang F, Howard JD, Voss JL, Schoenbaum G & Kahnt T Targeted stimulation of an orbitofrontal network disrupts decisions based on inferred, not experienced outcomes. J. Neurosci 40, 8726–8733 (2020). [PubMed: 33051355]

123. Buckley MJ et al. Dissociable components of rule-guided behavior depend on distinct medial and prefrontal regions. Science 325, 52–58 (2009). [PubMed: 19574382]

124. Preston AR & Eichenbaum H Interplay of hippocampus and prefrontal cortex in memory. Curr. Biol 23, R764–R773 (2013). [PubMed: 24028960]

125. 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). [PubMed: 26726963]

126. Squire LR & Alvarez P Retrograde amnesia and memory consolidation: a neurobiological perspective. Curr. Opin. Neurobiol 5, 169–177 (1995). [PubMed: 7620304]

127. Boorman ED, Rajendran VG, O'Reilly JX & Behrens TE Two anatomically and computationally distinct learning signals predict changes to stimulus-outcome associations in hippocampus. Neuron 89, 1343–1354 (2016). [PubMed: 26948895]

128. Sanders H, Wilson MA & Gershman SJ Hippocampal remapping as hidden state inference. eLife 10.7554/eLife.51140 (2020).

129. Schapiro AC, Rogers TT, Cordova NI, Turk-Browne NB & Botvinick MM Neural representations of events arise from temporal community structure. Nat. Neurosci 16, 486–492 (2013). [PubMed: 23416451]

130. Schapiro AC, Turk-Browne NB, Norman KA & Botvinick MM Statistical learning of temporal community structure in the hippocampus. Hippocampus 26, 3–8 (2016). [PubMed: 26332666]

131. Jun JJ et al. Fully integrated silicon probes for high-density recording of neural activity. Nature 551, 232–236 (2017). [PubMed: 29120427]

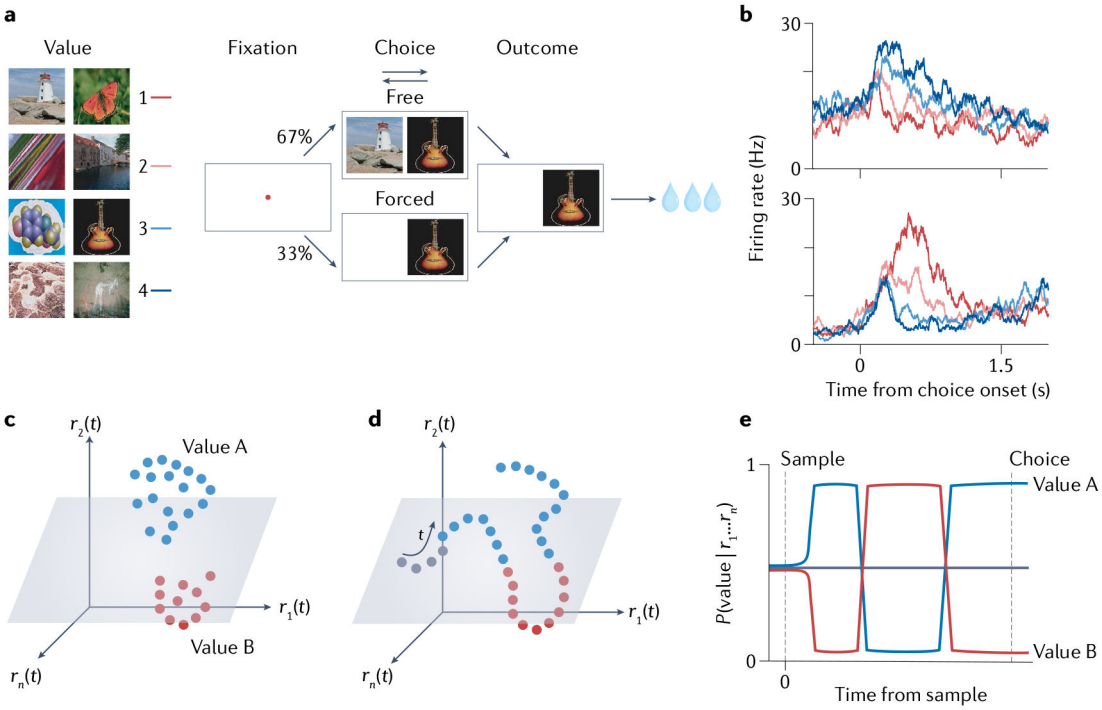

Fig. 1 |. During decision-making, the OFC flip-flops between representing the value of either option.

a | Monkeys learn the value of different pictures and then perform a choice task. After fixating on a red dot, one (forced choice) or two (free choice) pictures are presented and the animal must make its selection. **b** | Individual orbitofrontal cortex (OFC) neurons typically encode the value of the chosen picture[9]. About half of the value-encoding neurons have a positive relationship between firing rate and value (top), whereas the other half have a negative relationship (bottom). **c** | A classifier was trained to recognize patterns of neural activity that are elicited by specific picture values on forced-choice trials. In the example, the neural ensemble consists of n neurons ($r_1 \ldots r_n$) and the firing rate at time t is plotted. Each data point indicates a pattern of activity that was elicited on different trials, colour-coded according to the value of the outcome of the choice. A hyperplane can successfully separate the two groups of trials. **d** | The trained classifier is then used to decode activity on the free-choice trial. Each successive dot corresponds to the activity of each neuron at successive time steps, t, within a trial. When ensemble activity lies above the hyperplane, the picture's value is decoded as value A (blue dots), whereas below, it is decoded as value B (red dots). **e** | Decisions are characterized by OFC neural ensembles 'flip-flopping' between representing the value of either option in turn[39]. Panels **a** and **b** are adapted with permission from REF.[16], Springer Nature Ltd.

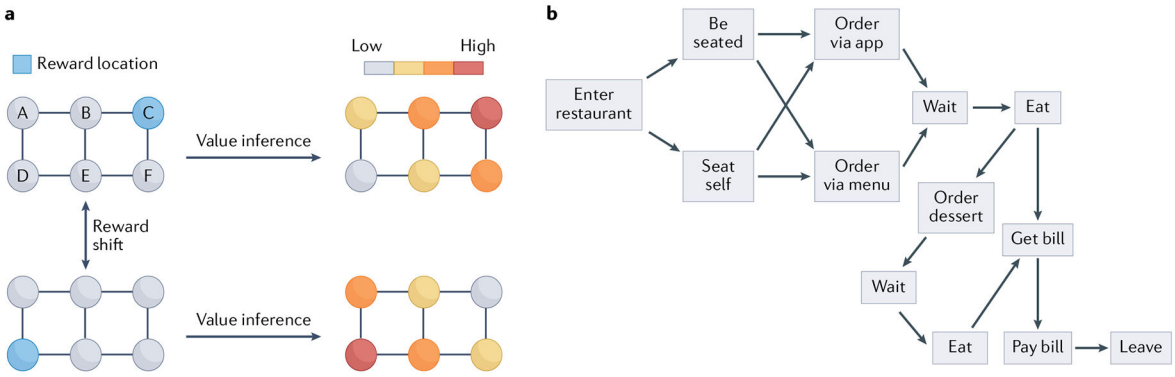

Fig. 2 |. The relationship between the state-transition graph and value.

a | The state-transition graph (left) specifies the relationship between states and the transition probability; for example, state A can directly lead to B or D, but not to the other three states of C, E or F. The state-transition graph can be integrated with the reward location to calculate the value of individual states (right). **b** | A potential state-transition graph for having a meal in a restaurant. Different states are linked to other states with a specific probability of occurrence. The same graph can be used for many different restaurants, and the value of states can be rapidly updated in response to changes in goals.

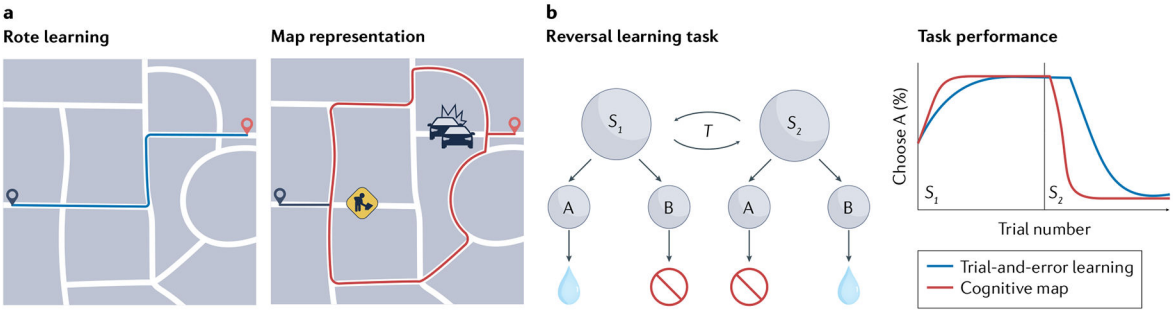

Fig. 3 | . An illustrative example of the flexibility of cognitive maps.

a | Imagine you need to get across town (start: black pin, goal: red pin). This can be done simply by following a rote set of directions (for example, take the second left followed by the first right). However, this method cannot cope with unexpected events such as construction works or an accident. By contrast, a map-like representation allows on-the-fly adjustments, providing a much more flexible way to navigate. **b** | Cognitive maps can also be applied to behavioural tasks and provide the same kinds of advantages for high-level behaviour that maps provide for spatial navigation. Consider the classic reversal task with two reward-predictive cues, A and B. This task can be mapped as two distinct states (S_1 and S_2) that describe the likelihood of each cue predicting reward, with some probability of transitioning between the two states, T. Without such a map, the task could still be completed through trial-and-error learning (bottom, blue trace). However, the map enables more-flexible switching between states (bottom, red trace) because the participant can inferthe outcomes associated with each cue from a single trial.

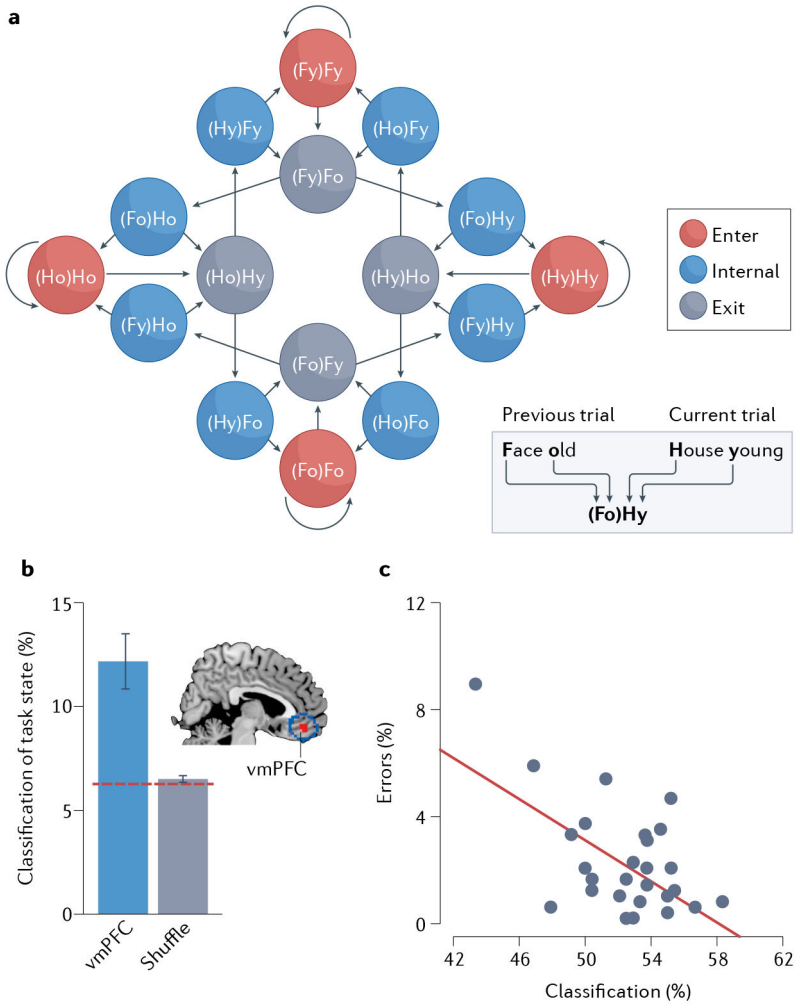

Fig. 4 |. Responses in human vmPFC reflect encoding of state information.

a | Map of the task. Participants saw houses and faces and had to judge the age of one of the categories. When the age changed, they had to switch to judging the other category. Each combination of task contingencies is defined as a state (circles), with transitions between each state governed by the structure of the task. **b** | The researchers trained an algorithm to decode which of the 16 states was currently in effect. The only region in which this information could be decoded above chance was the ventromedial prefrontal cortex (vmPFC). **c** | The degree to which state information could be extracted from vmPFC blood oxygen level-dependent signals (x-axis) predicted the number of incorrect decisions participants made (y-axis). Figure adapted with permission from REF.[8], Elsevier.

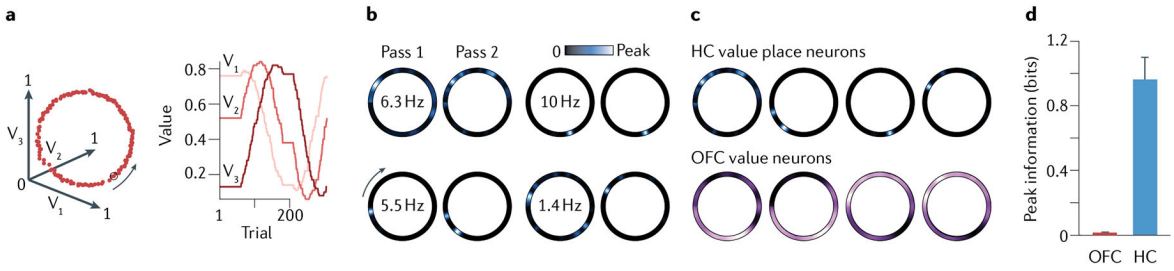

Fig. 5 | Value place neurons in the primate hippocampus.

Monkeys were trained to choose between pairs of presented pictures[98]. There were three pictures in total and each was associated with a probability of receiving juice that gradually changed overtime, requiring the animal to track the changing contingencies. **a** | A value space can be constructed in which each axis is the value of one of the pictures, such that a point in this space defines the value of the pictures relative to one another. The changing reward contingencies result in a circular trajectory through this value space. On the right, the same picture values are plotted independently. **b** | Four examples of firing patterns of hippocampal (HC) neurons, showing how the firing rate varied across two successive completions of the circular trajectory. Peak firing rate values are shown for each neuron. Neurons encoded specific locations in value space. **c** | Top: Activity of four example HC value place neurons plotted following conventions in part **b.** Bottom: Four example orbitofrontal cortex (OFC) neurons, plotted in the same manner. **d** | Average peak spatial information encoded by OFC value neurons (left, purple) and HC value place neurons (right) on the circular trajectory. Figure adapted with permission from REF.[98], Elsevier.