



HHS Public Access

Author manuscript

Behav Neurosci. Author manuscript; available in PMC 2022 April 01.

Published in final edited form as:

Behav Neurosci. 2021 April ; 135(2): 267–276. doi:10.1037/bne0000463.

The orbitofrontal cartographer

Matthew P.H. Gardner, Geoffrey Schoenbaum

NIDA-IRP, 251 Bayview, Baltimore, MD 21224

Abstract

Theories of orbitofrontal cortex (OFC) function have evolved substantially over the last few decades. There is now a general consensus that the OFC is important for predicting aspects of future events and for using these predictions to guide behavior. Yet the precise content of these predictions and the degree to which OFC contributes to agency contingent upon them has become contentious, with several plausible theories advocating different answers to these questions. In this review we will focus on three of these ideas - the economic value, credit assignment, and cognitive map hypotheses – describing both their successes and failures. We will propose that these failures hint at a more nuanced and perhaps unique role for the OFC, particularly the lateral subdivision, in supporting the proposed functions when an underlying model or map of the causal structures in the environment must be constructed or updated.

Theories of orbitofrontal cortex (OFC) function have evolved substantially over the last few decades. Convergence of observations from neural recordings and OFC manipulations in behavioral studies as well as evidence calling into question hallmarks of OFC function have led to a revolution in our understanding of the function of this brain region (Gallagher et al., 1999; Murray et al., 2007; Ostlund and Balleine, 2007; Padoa-Schioppa and Assad, 2006; Rolls, 1996; Rudebeck and Murray, 2011b; Wallis, 2007). There is now a general consensus that the OFC is important for predicting aspects of future events and for using these predictions to guide behavior. Yet the precise content of these predictions and the degree to which OFC contributes to agency contingent upon them has become contentious, with several plausible theories advocating different answers to these questions. These theories range from suggestions that OFC is important for tracking the associative relationships between stimuli (Rolls, 1996; Schoenbaum et al., 2009), to assigning credit (Noonan et al., 2010), to computing the value of options during economic decisions (Padoa-Schioppa, 2011; Padoa-Schioppa and Conen, 2017), to contributing to cognitive mapping (Wilson et al., 2014), or even as arbitrating between model-free or model-based approaches (Panayi et al., 2021). Although these theories overlap in important ways, such as their reliance on the use of models of the environment, they make substantially different claims as to the specific function of the OFC.

In this review we will focus on three of these ideas - the economic value, credit assignment, and cognitive map hypotheses – describing both their successes and also how, in their current form, they fail to fully incorporate several recent studies. These failures hint at a

more nuanced and perhaps unique role for the OFC¹, particularly the lateral subdivision, when an underlying model or map of the causal structures in the environment must be constructed or updated. We will propose that the OFC endows behavioral flexibility by allowing models, or cognitive maps, to be updated in other brain regions when new associative information is encountered that must be integrated with an existing model or when existing associations undergo changes.

Orbitofrontal cortex and economic choice

One theory that has gained enormous traction in the past decade to explain OFC function is the economic choice hypothesis (Fellows, 2007; Levy and Glimcher, 2012; Montague and Berns, 2002; Padoa-Schioppa, 2011). This idea grew out of observations that OFC damage is correlated with deficits in decision-making tasks and that OFC neurons are particularly tuned to predict value as defined by choice. Myriad electrophysiology recording and fMRI studies have indicated that OFC strongly encodes predictions of stimulus features that track with value (Breiter et al., 2001; FitzGerald et al., 2009; Gottfried et al., 2003; Gremel and Costa, 2013; O'Doherty et al., 2001; Padoa-Schioppa and Assad, 2006; Rich and Wallis, 2016; Roesch et al., 2006; Schoenbaum and Eichenbaum, 1995; Thorpe et al., 1983; Wallis and Miller, 2003; Zhou et al., 2019a). While generally it is difficult to dissociate such signals from correlated outcome features, such as the amount of food available, economic choice tasks are designed to specifically get at this question of whether 'pure' value, or more formally utility, is encoded by neural activity. Studies utilizing these types of tasks have revealed that some neurons in OFC, mainly area 13 of non-human primates, appear to encode value independently of physical features of rewarding stimuli, at least within a limited parametric space (Padoa-Schioppa and Assad, 2006, 2008). Although not without critics (Blanchard et al., 2015), these findings have provided a strong basis for the theory that OFC's function is to compute such pure or economic value. This theory holds that OFC determines, in real-time, the utility of available options in the environment to support optimized decision making. To do this, OFC receives high dimensional information about each option from other brain regions and distills this high-dimensional information down onto a single axis of subjective value (Padoa-Schioppa, 2011; Rustichini et al., 2017). This putatively allows a subject to make rational choices among complex and unrelated options in order to optimize decision-making behavior. According to this model, the OFC should be important whenever the subject is using real-time information to determine the value of available options.

Yet while there are clearly correlates of economic value within OFC in canonical economic choice tasks, it has been less clear whether such activity is causally required for economic choice. This changed only recently with several studies in different species employing the canonical economic choice task developed by Padoa-Schioppa et al. (Padoa-Schioppa et al.,

¹For this review, our intent was to focus on what is considered to be the lateral region of the orbitofrontal cortex. It is clear that the various subregions within the orbitofrontal cortex may have different functions, and there is substantial debate as to the homology of OFC across the species mentioned within this review. For a more detailed discussion of this issue, as well as for discussion of the homologies across species, please see Izquierdo, A. (2017). Functional heterogeneity within rat orbitofrontal cortex in reward learning and decision making. *Journal of Neuroscience* 37, 10529–10540. Rudebeck, P.H., and Murray, E.A. (2011a). Balkanizing the primate orbitofrontal cortex: distinct subregions for comparing and contrasting values. *Annals of the New York Academy of Sciences* 1239, 1–13.

2006). In the basic version of this task, subjects are confronted with two different reward options. The value of each option must be determined by integrating across dimensions of information characterizing the available options; this is typically done by varying outcome-type and amount (Padoa-Schioppa et al., 2006), though some forms of the task require integration over probability and amount (Constantinople et al., 2019). Subjects then choose their preferred option. By presenting many trials over a range of offer-types, a psychometric curve can be determined based on the subject's choice behavior. Experienced subjects typically show a relatively steep psychometric curve indicating a stable utility function. According to the economic choice account, the OFC should be essential for normal choice behavior under these conditions, since it is calculating the utility of underlying the options. Thus, orbitofrontal manipulations should result in dramatic impairments consistent with a loss of the ability to integrate across features to calculate this quantity, resulting in shallow curves, default choice strategies based on number or a side-bias, and most importantly choices among items that are irrational or that fail to obey rules of transitivity.

While these predictions have found limited support in some causal studies (Ballesta et al., 2020; Fellows, 2011; Kuwabara et al., 2020), a substantial proportion of such studies have failed to find effects of lateral OFC manipulations on economic choice behavior in rats (Constantinople et al., 2019; Gardner et al., 2019; Gardner et al., 2017; Gardner et al., 2020). Indeed in at least one case, inactivation of OFC seemed to make choices more rational (Barrus & Winstanley 2017). Consistent with these findings, a recent study failed to find robust encoding of value predictions in the macaque OFC, areas 11/13, at the time of choice in a task designed to determine whether reward predictions accumulate with additional evidence (Lin et al., 2020). Work from our lab, carefully designed to duplicate as much as possible the task used to obtain the single-unit recording data forming the basis of this model, found absolutely no effect at all of optogenetic inactivation of either lateral (Gardner et al., 2017) or medial (Gardner et al., 2018a) OFC in rats on any of several measures of choice performance, including the established indifference point for specific goods pairs, the steepness of the relationship, or the transitivity of the rats' choices among goods in several pairs. The precise factors determining when the OFC is necessary for economic choice, and when it is not, are not clear from these experiments; however, at a minimum, these findings indicate that having to integrate across multiple dimensions to calculate values to guide a choice does not fundamentally require the OFC, despite the clear neural correlates of value that exist in OFC under such conditions.

Orbitofrontal cortex and credit assignment

Another theory that has been advanced to explain orbitofrontal function is the credit assignment hypothesis (Noonan et al., 2010; Walton et al., 2010; Walton et al., 2011). This idea is based on work aimed at understanding OFC's role in tracking dynamic contingencies as well as risky, or probabilistic, decisions. These tasks typically require tracking the value of a stimulus over multiple trials to determine the expected value of a reward, making them perfect candidates for modeling using reinforcement learning (RL). Within the RL framework, prediction errors are determined by mismatches between the value of expected and actual reward. These prediction errors must be attributed correctly for the organism to best predict what specifically led to the reward and update behavior appropriately on future

latent, states not easily distinguished by external input. The critical importance of encoding the current state within a state space over model-free alternatives that only use states defined by explicit stimuli is the ability to use non-observable information to determine the full nature of the current situation.

To better understand the importance of encoding such partially-observable or latent states, it's important to consider the assumptions of the Markov decision process that underlies the state spaces within reinforcement learning theories. In order to discern two situations that have identical sensory features but are different based on how the subject arrived at that point in time, Markov processes require the formation of different latent states in order to capture the historical knowledge, or memory, of the path to that point. For example, in a game of tag, it's critical to maintain a memory of whether one is 'it' (or 'not it') to know whether to run towards (or away from) fellow game players. Markov models solve this problem by creating different states based on the internal memory of being 'it'. These hierarchical states then set the stage for appropriate actions to take given the current observations – if you're 'it' then chase everyone; if you're not 'it' then run away from the whoever is 'it'. Because memories are captured by distinctive states, a Markov model of a game of tag would require unique states for each person who is 'it' in order for the model to recognize who to run away from.

Here it is perhaps worth emphasizing that, as currently formulated, the cognitive map hypothesis proposes only that the current state representations reside in the lateral OFC (Niv, 2019; Sharpe et al., 2019). This is subtly different than the proposal that the full model, or the mappings between state representations used to make prospective decisions, is housed in OFC. As a result, the output from OFC under the cognitive map hypothesis could be used for model-based prospective guidance by providing knowledge of the current state, but it also could also perhaps support model-free learning by helping to specify the correct state to which to attribute cached values. In the cognitive map hypothesis, as applied to the OFC, because the organism has access to the current state, there is the potential to associate cached values with latent, or non-explicit states. Indeed, the idea that OFC provides a 'pointer' to the current state is able to explain a range of observations of OFC's importance for model-based behaviors as well as for tasks requiring solutions based on the existence of latent states.

To illustrate how representing latent states as part of a cognitive map can explain behavioral deficits associated with OFC dysfunction, it's useful to return to reversal learning, which has long been considered to be the archetypal behavior affected by OFC dysfunction. As mentioned before, theories that suggest OFC is involved in updating values of states, such as the credit assignment hypothesis, assume that values are learned and incrementally changed on a trial-by-trial basis following a reversal. The cognitive map hypothesis instead handily explains deficits in reversals by developing hidden hierarchical states that tell the subject which contingencies are in effect. Once the model is learned, subjects should be able to quickly flip between the two hierarchical states as the problem becomes a latent state detection issue rather than a learning issue.

Although experimentally discerning whether animals are reversing behavior due to state detection or through cached-value learning is tricky in these tasks, state detection should allow more rapid behavioral changes within tasks employing well-defined, or discernable, states. In fact, a recent study showed that mice performing a two choice probabilistic foraging task do show switches in behavior strongly consistent with a latent state model (Vertechi et al., 2019). Consistent with the cognitive map hypothesis, inactivation of lateral OFC switched mice from rapidly changing behavior during changes in task state, putatively dependent on encoding of a latent state, to a slower behavioral change consistent with model-free, or trial-by-trial, updating of the location values.

Of course, a strategy of determining state likelihood is not always the optimal solution for reversals. In another study, rats with lateral OFC lesions experienced reversals of probabilistic rewards which occurred at a very high or very low frequency. High frequency reversals with probabilistic rewards overall provide less information about which set of reward contingencies are currently active. This lack of clear available evidence results in a noise hindered use of a state-based strategy, and as such, a model-free strategy of directly tracking reward values in a single state remains an effective approach. Interestingly, OFC lesions did not impair high frequency reversals, yet rats with OFC lesions were impaired during low frequency reversals when state likelihoods are considerably more well-defined and exploitable for optimal behavior (Riceberg and Shapiro, 2012). Recordings within OFC were consistent with this finding as encoding of expected values were more stable within the low frequency reversals (Riceberg and Shapiro, 2017). Consistent with this use of latent states during well-defined transitions, immediate changes in state have been observed in single OFC neurons after just a single trial following a contingency reversal (Stalnaker et al., 2014). Critical to the argument that an unobservable state had switched, stimuli that were not experienced following a state change until after a few trials showed immediate changes in updated activity.

Support also comes from work by Stolyarova et al., who ran a two-choice probabilistic decision-making task in which they showed that lateral OFC was required for discerning small yet discrete shifts in the expected value of two options from the expected variance, indicating that OFC might be parsing the reward likelihoods into discrete states that must be detected to induce a state switch (Stolyarova and Izquierdo, 2017). Gershman et al. found neural evidence of such state encoding within OFC in a similar task in which different states were employed only when a stimulus feature provided task-relevant information (Gershman and Niv, 2013). OFC also encodes the current location within more complex state spaces that include hidden states (Schuck et al., 2016; Zhou et al., 2019a; Zhou et al., 2019b). Interestingly, current information regarding hidden or unobservable states is encoded in OFC most strongly at the time when the latent state is required for optimal performance in the task (Zhou 2018).

This idea can also be expanded to explain the more pervasive role for OFC in non-deterministic or probabilistic discrimination tasks, typically used to support credit assignment accounts. What has been considered the strongest evidence for the credit assignment theory comes from data showing that subjects performing a three-arm bandit task require an intact OFC (Noonan et al., 2010) and intact connections with other reward

related areas such as basolateral amygdala and nucleus accumbens (Groman et al., 2019). Within this task, three different options with different probabilities of reward are available on each trial. Over time, the expected value of at least one of these options changes stochastically, leading to changes in the best option at any given time. Lesions of OFC impairing behavior on this task were interpreted as showing that OFC was needed to appropriately match the errors occurring from trial-to-trial with the appropriate action.

Yet there is another solution to this task that meshes more with the cognitive map hypothesis, which is based on partially observable Markov decision processes (POMDP) (Averbeck, 2015; Rao, 2010). Simple RL frameworks treat the n-arm bandit task as a single state RL problem with trial-by-trial updates of each the n-actions that can be taken; put more simply, the values of each option are tracked independently over time. POMDP models, by contrast, base decisions on the expected value determined from a distribution of many different states. These different states are based upon the information gained about expected values of each action over each trial, and different states exist which predict that any of the actions might lead to the highest expected value. This can also be modeled using a discrete state space rather than as a continuous space (Rao, 2010) and so can be considered in the framework of the cognitive map hypothesis, in which OFC is attempting to determine the current latent state. As more information is gained about the options, the posterior probability distribution becomes more refined, leading to better predictions of the action with the highest likelihood of being the best option. Neural activity in OFC is consistent with this idea, with representations of components of the posterior probabilities as well as the full posterior distribution having been observed in OFC (Chan et al., 2016; Ting et al., 2015; Vilares et al., 2012). Even more nuanced predictions of POMDP models such as novelty bonuses in n-arm bandit tasks are clearly represented in OFC (Costa and Averbeck, 2020). Thus, while studies employing probabilistic discrimination tasks implicate OFC in tracking the values across trials in a way consistent with standard TDRL models, these findings can also be explained by the hypothesis that OFC is involved in the encoding of partially observable states with different reward probabilities.

Finally, the cognitive map hypothesis also incorporates broad evidence that OFC is important for sensory-sensory learning (Hart et al., 2020; Howard et al., 2020; Jones et al., 2012; McDannald et al., 2014; McDannald et al., 2011; McDannald et al., 2005; O'Doherty et al., 2000; Ostlund and Balleine, 2007; West et al., 2011). Perhaps the strongest example of this is OFC's importance for sensory preconditioning. In this task design, two stimuli (A and B) are presented in sequential order (A-B) during a conditioning phase in which no rewards are given. Following this reward-free conditioning phase the second stimulus (B) is associated with reward. Note that A is never present during this B-reward conditioning phase. Finally, responding to A is assessed in a probe test in which reward is not present. Most subjects show higher responding to cue A in comparison to a control stimulus C (typically within-subjects) for which pairs of stimuli (C-D) were trained in the initial preconditioning stage but for which D was not paired with reward. OFC is required during the probe test (Jones et al., 2012) as well as during only the initial preconditioning training when no rewards are present (Hart et al., 2020). This importance of OFC for the initial learning cannot be explained by an account that focuses on calculating utility, since the OFC is online during the final probe test, and the importance of OFC in the final probe test cannot

be explained by an account that focuses on credit assignment, since learning is not required in the probe test. On the other hand, the cognitive map hypothesis can account for both results, since the OFC must be online to develop the map defining the key relationships and to use it later to guide behavior.

The cognitive mapping model is similarly able to explain observations that the OFC is critical for behavior in diverse settings such as identity unblocking (McDannald et al., 2011), in which learning occurs even when there is no reward prediction error present, specific Pavlovian-to-instrumental transfer (Ostlund and Balleine, 2007), in which Pavlovian cues predicting specific sensory features of reinforcers bias subjects towards choosing specific actions (Lichtenberg et al., 2017; Ostlund and Balleine, 2007), and the differential outcomes effect (McDannald et al., 2005), in which learning about different actions is accentuated when the actions result in reinforcers with different features. All of these settings require value-neutral mapping of associative information either to drive learning or behavior; the involvement of the OFC is hard or impossible to accommodate through economic choice or credit-assignment theories, but fits well with the idea that OFC mediates cognitive mapping.

Indeed, the idea of a cognitive map in operation is perhaps best epitomized by a form of sensory-specific learning that has become the most defining behavioral assay of OFC function - Pavlovian reinforcer revaluation. OFC's importance in reinforcer revaluation has been observed in rodents, non-human primates, brain damaged and most recently even neurotypical humans (Gallagher et al., 1999; Gardner et al., 2019; Howard et al., 2020; Izquierdo et al., 2004; Panayi and Killcross, 2018; Parkes et al., 2018; Pickens et al., 2005; West et al., 2011). While details differ, these tasks generally involve a conditioned stimulus that is trained to predict a rewarding stimulus such as food. After learning, the value of the reinforcer is changed outside of the experimental setting and not in the presence of the predictive stimulus. For food, this is typically done by satiating the animal or by pairing with illness. When the subject is subsequently exposed to the conditioned stimulus, the subject must infer through the sensory properties of the cue and its association with the particular food that the predictive value of the cue has changed. Disruption of OFC consistently leads to deficits in this inferential step, leading to inappropriate responding to the conditioned stimulus.

The cognitive map hypothesis explains reinforcer revaluation in the same way it has traditionally been interpreted through modeled-based mechanisms - suggesting that inactivation of OFC disrupts the ability to use a full model relating the stimulus to the sensory properties of the food to their updated value. Loss of the model-based mechanism is thought to result in behavior relying on cached-value systems that are unaware of the updated value of the outcome following revaluation. This explains why subjects with OFC damage or in whom OFC is inactivated continue to respond as if the change in outcome value never occurred. Notably, we have reported a similar effect of revaluation on economic choice behavior; while OFC inactivation has no effect on baseline behavior, if OFC is inactivated after satiation on one of the outcomes on offer, subsequent choices fail to reflect this updated value (Gardner et al., 2019).

Yet while the cognitive map hypothesis appears to explain a broad array of data regarding OFC function, it too has areas where it comes up short of explaining the effects of OFC manipulations on behavior. One of these areas is reversal learning. As mentioned earlier, deficits in reversals have long been associated with OFC dysfunction. Yet recent research has cast doubt on whether orbitofrontal cortex is ubiquitously necessary for reversal learning, particularly for multiple reversals or those that occur after the first established reversal (Jang et al., 2015; Kazama and Bachevalier, 2009; Keiflin et al., 2013; Panayi and Killcross, 2018; Rudebeck and Murray, 2011b; Schoenbaum et al., 2002). While cognitive map theory successfully explains why OFC would be required for reversals in general, it is not clear why this involvement would be highly variable or affected by prior experience; if the OFC keeps track of the latent states, then this function should be equally necessary each time a reversal is encountered. Thus impairments should be reliable and persistent.

Other results of OFC inactivation are also at odds with predictions of the cognitive map hypothesis, particularly in tasks where models must be used but OFC inactivation does not seem to have a strong effect. One example is the Daw two-step task (Daw et al., 2011). This task was explicitly designed to distinguish choices based on cached values of states versus choices reflecting a knowledge of a full model. A simplistic explanation of the design of the task is that subjects are presented with two options, each of which leads to additional states, with some probability of transition. These second-stage states are rewarded with different probabilities. If subjects use a cached-value system, they update the first stage states based on whether they were rewarded on a given trial. If subjects rely on a full model for their decisions, they should only change the value of the action taken in the terminal state, and choose the first stage options based on the probabilities of transitioning to the second stage state that lead to greater reward. Contrary to the predictions of the cognitive map hypothesis, OFC inactivation did not disrupt the ability of rats to use models to prospectively make decisions, and instead selectively affected their ability to use adjust the state-space model to incorporate the new values (Miller et al., 2018). This resulted in a trial-to-trial failure to update decisions after incorporating new information about the model. Interestingly this surprising result replicates effects of OFC manipulations in a probabilistic economic choice task in which similar learning-specific effects of OFC inactivation have been reported (Constantinople et al., 2019). Another study indicated similar importance of OFC in learning about hierarchical states, but not for use of the hidden states – rats showed no deficits in reversals at the time of OFC inhibition, but were impaired on reversals the subsequent day when OFC was back online (Keiflin et al., 2013).

Similarly, we have found that OFC is not required for economic choices that are established, however it is required for economic choice when goods pairs are encountered together for the first time, even if the subjects are familiar with each of the options independently (Gardner et al., 2020). The study suggested that when confronted with options they have never chosen between, subjects adjust the feature space of the options as they become more familiar with the decisions. Inactivation of OFC disrupted the ability of rats to make refinements to the space and quickly reach a stable preference for the novel choices. Evidence of OFC encoding a putative feature space has been shown in humans (Constantinescu et al., 2016), and this encoding appears even for novel pairings of things that have been experienced separately before, providing a potential neural substrate for the

importance of OFC to this process (Barron et al., 2013). Although this concept of a feature space lies outside of the state space framework of reinforcement learning, the essence of it can be captured by the discrete state space framework of the cognitive map hypothesis. This feature-based space is consistent with other observations that OFC is important for keeping track of sensory-sensory associations.

Current OFC theories do not fully encapsulate all findings from causal studies

Each of the formal theories of orbitofrontal function considered within this review explain a considerable portion of the work on OFC function. And even when OFC might not be explicitly necessary for behavior, large proportions of individual neurons strongly encode information about task variables within many behavioral paradigms, information that would be relevant to each of the proposed accounts. Yet at the same time, each of the theories falls short of explaining all phenomena for which OFC is known to be necessary. In general, theories either ascribe more function to OFC than causal studies indicate, or they fail to capture evidence from other settings in which OFC is necessary. For instance, the theory that OFC encodes the economic value of goods to drive choice behavior suggests that OFC is required whenever a computation of expected utility is needed in order to compare options. It is clear from the evidence that this is not always true, particularly when decisions have become well-practiced, even though those decisions do seem to involve a process in which individual values are being integrated in real-time (Gardner et al., 2019). Further, it is not clear from this theory why OFC would be required for tasks that require tracking of sensory-sensory associations such as sensory preconditioning, where OFC is required in the initial phase of sensory learning (Hart et al., 2020; Jones et al., 2012; Wang et al., 2020), identity unblocking (McDannald et al., 2014), or specific Pavlovian-to-instrumental transfer (PIT) (Lichtenberg et al., 2017; Ostlund and Balleine, 2007), where OFC seems to be necessary to learn or use specific outcome associations independent of value. If OFC is simply determining the value of options relevant at any given moment, why would it be needed for ascribing value to a stimulus when sensory features have been swapped with no changes in value?

Similar problems arise in evaluating the theory that OFC is required for credit assignment or that it represents the cognitive map used to guide behavior. Although credit assignment requires a well-defined state or action representation to which value can be attached, it is not clear why identity unblocking would occur in the RL framework in which this hypothesis is framed. The credit assignment hypothesis also fails to explain the many studies implicating OFC in inference in the absence of clear learning requirements, such as its importance in revaluation tasks, specific transfer, or for responding to preconditioned cues. By contrast, the cognitive map hypothesis does a much better job of explaining these results, while having difficulty explaining why OFC does not appear to be required for multiple reversals, decisions within the Daw Two-Step Task, or for acting on established preferences in economic choice tasks. If OFC is necessary for maintaining a pointer to the current latent state in the cognitive map underlying Pavlovian revaluation, why is it not similarly necessary for this operation in these other tasks?

Orbitofrontal cortex and cognitive re-mapping

While it is not clear how to expand the economic choice or credit assignment hypotheses to fully account for these diverse data, it is possible to tweak the cognitive map hypothesis to explain the lack of necessity of OFC for well-practiced behaviors that nevertheless seem to require the use of internal models. This tweak involves the proposal that OFC is required not to support the use of cognitive maps but rather only for their formation and updating. According to this modest adjustment, different brain regions encode different maps or associations, and the OFC is primarily important for helping to adjust, accommodate, or make inferences within or between these extant maps (Figure 1). This change in the hypothesis could be considered similar to ideas that OFC is important for behavioral flexibility (Rolls, 1996), but incorporates this flexibility into a more formalized and established theory; it harks back to older ideas about prefrontal subdivisions being critical for domain-specific working or scratchpad memory (Goldman-Rakic, 1987), but with the domain-specific function of the OFC being tied not to particular sensory domains or even affect but to the integration of the disparate associative information comprising distinct mapped learning episodes. In this regard, OFC would be crucial for appropriately combining such episodes to provide a scaffold for other brain regions to incorporate this information into a region-centric map, information perhaps not typically available to each individual area. In the framework of RL, this would consist of the formation of a new state, including hierarchical latent states proposed as OFC-critical in the cognitive map hypothesis, or a change in the state transition structure specific to a particular brain region.

To understand how this change to the cognitive map hypothesis provides a stronger fit of the full array of experimental results, we can revisit the situations in which the cognitive map hypothesis, as it stands, has difficulty explaining observations. As mentioned above, the OFC is most important for the first reversal, if at all (Kazama and Bachevalier, 2009; Keiflin et al., 2013; Panayi and Killcross, 2018; Rudebeck and Murray, 2011b). Because the cognitive map hypothesis suggests that OFC keeps track of the different latent states that facilitate reversal learning (Wilson et al., 2014), it should be required to continually identify switches between states, thus it should be necessary for the first and subsequent reversals. If instead OFC is required for developing a new hierarchical state within another brain region or network, once the state has been fully formed, additional reversals would no longer require OFC. This would explain why first reversals would be most sensitive to OFC manipulations, as the new state would take longer to develop (or might not develop properly at all in some cases) with impairments to OFC. This would also explain why performance that appears to be model-based, such as in economic choice tasks and the two-step, does not depend on OFC at choice in well-practiced animals, while at the same time predicting that similar behaviors that require establishing or updating the model at the time of inactivation are dependent on OFC.

Of course, this modified proposal is also still consistent with the broad evidence that the OFC is important for learning and behavior in a host of paradigms that emphasize the integration of associative information acquired in different settings, such as specific PIT, unblocking, sensory preconditioning, and outcome revaluation (Hart et al., 2020; Howard et al., 2020; Izquierdo et al., 2004; Jones et al., 2012; Lichtenberg et al., 2017; McDannald et

al., 2014; McDannald et al., 2005; Ostlund and Balleine, 2007; Panayi and Killcross, 2018; Parkes et al., 2018; Wang et al., 2020). Importantly however, its involvement in these settings, according to this new model, differs from earlier theories in a way that gets at the heart of what this theory proposes as the function of OFC. Specifically, OFC's role in these settings is not to support navigation within the current model (Figure 1, left), but rather it is to establish that model in the first place - or more typically to modify a existing models (Figure 1, right).

While generally not much emphasized, the need to modify an existing model is actually inherent in each of the iconic OFC-dependent settings enumerated above. This is because in each case, the critical OFC-dependent tests involve a new use of disparate models, formed in each prior stage of training and potentially within different networks. To take one example, in a Pavlovian conditioning paradigm, the network representing the learning of stimulus-outcome associations is likely somewhat different from the network storing the outcome-revaluation association. OFC would be required for making an inference during the extinction test because it allows these disparate episodes to be connected and formed into a single cohesive model. A similar logic can be applied for why the OFC is critical to responding to preconditioned cues in the final test phase of sensory preconditioning or to drive specific transfer in PIT. In fact, there already exists some evidence for this as OFC appears to be not necessary for learning in initial acquisition of either Pavlovian or instrumental behaviors, but is required for subsequent learning of associations, or updates, to an already existing model (Panayi and Killcross, 2020; Parkes et al., 2018).

Here it is worth noting that this reconceptualization (Figure 1, right) makes several novel predictions that differ from the current model (Figure 1, left), and which can be tested in these simple (and perhaps even in more complex) tasks. First, learning without the OFC likely would not be completely normal. Specifically such learning would presumably not be "mapped" appropriately and might not be capable of supporting model-based behavior later, even if the OFC is back online. Some evidence supporting this idea already exists in the finding that OFC inactivation during the initial sensory-sensory learning phase of preconditioning abolishes subsequent probe test behavior (Hart et al., 2020). Second, behavior that remains after OFC inactivation in probe tests in settings like revaluation or preconditioning would not be habitual but instead would simply be using an outdated model. While a subtle distinction from current thinking, this predicts that other brain regions involved in storing the models, such as DMS and DLPFC (Sharpe et al., 2019), would still be controlling these behaviors, something that could be shown experimentally. Third, if OFC is online during the probe tests in these simple tasks, then the new models should be established, at least partially, as they are deployed; this predicts that the OFC might not be necessary for model-based performance in subsequent test sessions.

This ability to connect similar representations also hints at OFC's potential importance for developing learning sets, or representations of the generalized structures of similar tasks, to support faster learning of new situations compatible with these learned task structures (Behrens et al., 2018). In fact, OFC encodes cognitive maps more readily when confronted with new stimuli arranged in an already learned state space (Zhou et al., 2020), and it is necessary for use of a learned space when confronted with a new situation (Gardner et al.,

2020). OFC's importance for learning these generalized task structures would endow OFC with the ability to make inferential steps giving it a 'leg up' over other brain regions that putatively can encode a state space with hidden states.

Of course, we do not mean to draw an ironclad distinction here between situations where OFC is necessary and situations where it is not. It is likely that many situations do not clearly fall into one category or the other; in certain situations or tasks, perhaps where changes in the underlying cognitive map are especially dynamic or unstable, the contribution of the OFC would be continue to be required for what appears to be well-practiced behavior. In these settings, the novel proposed function would be indistinguishable from the existing cognitive map hypothesis. This would explain why in a limited number of cases, the OFC does appear to play an ongoing role when a behavior, arguably, should be well-learned (Noonan et al., 2010; Vertechi et al., 2019; Walton et al., 2010).

Conclusions

Here we have reviewed a number of currently dominant proposals to explain OFC function. Each can draw substantial support from correlative and, more importantly, causal studies; however each also fails to explain critical data from relatively recent studies. While the economic choice and credit assignment hypotheses seem hard to modify to fully capture these diverse data, the idea that the OFC supports the use of cognitive maps appears more robust. Specifically we have proposed that if the role of OFC is modified to focus on a role in supporting the formation or modification of existing maps, then this account largely explains the current evidence concerning the key role of OFC in behavior. Critically this account also makes a number of predictions that could be readily tested.

In addition to these testable direct predictions, indirect support could also come from examining how the resultant cognitive maps are relayed to other brain regions. Such data would help confirm that OFC modifies the maps and it could also clarify details of how that happens. For instance, modification may occur secondary to OFC gaining behavioral control when inferences are required, thereby allowing state spaces to be learned about independently of anatomical connections, or it may occur indirectly, through a role for OFC in providing key information regarding the current state space to broadly signaling teaching mechanisms in the brain (Gardner et al., 2018b; Namboodiri et al., 2019; Takahashi et al., 2011), or the OFC may provide direct input to downstream areas, including sensory regions, to reshape the maps (Banerjee et al., 2020). Such learning signals, whether direct or indirect, might be used as a template for learning of sensory-sensory associations, or state transition probabilities in other regions, including the possibility of sensory areas (Banerjee et al., 2020; Bao et al., 2019; Liu et al., 2020). Further work discriminating more directly between these possibilities in experiments tracking state-specific neural representations would be of great interest.

Acknowledgements

The authors would like to thank Yael Niv, Jingfeng Zhou, and Kaue Costa for their feedback and insights. This work was supported by the Intramural Research Program at the National Institute on Drug Abuse (ZIA-DA000587). The opinions expressed in this article are the authors' own and do not reflect the view of the NIH/DHHS.

References

- Averbeck BB (2015). Theory of choice in bandit, information sampling and foraging tasks. *PLoS Comput Biol* 11, e1004164.
- Ballesta S, Shi W, Conen KE, and Padoa-Schioppa C. (2020). Values Encoded in Orbitofrontal Cortex Are Causally Related to Economic Choices. *bioRxiv*, 2020.2003.2010.984021.
- Banerjee A, Parente G, Teutsch J, Lewis C, Voigt FF, and Helmchen F. (2020). Value-guided remapping of sensory cortex by lateral orbitofrontal cortex. *Nature* 585, 245–250. [PubMed: 32884146]
- Bao X, Gjorgieva E, Shanahan LK, Howard JD, Kahnt T, and Gottfried JA (2019). Grid-like Neural Representations Support Olfactory Navigation of a Two-Dimensional Odor Space. *Neuron* 102, 1066–1075 e1065. [PubMed: 31023509]
- Barron HC, Dolan RJ, and Behrens TE (2013). Online evaluation of novel choices by simultaneous representation of multiple memories. *Nature Neuroscience* 16, 1492–1498. [PubMed: 24013592]
- Bechara A, Damasio H, Tranel D, and Damasio AR (1997). Deciding advantageously before knowing the advantageous strategy. *Science* 275, 1293–1295. [PubMed: 9036851]
- Behrens TEJ, Muller TH, Whittington JCR, Mark S, Baram AB, Stachenfeld KL, and Kurth-Nelson Z. (2018). What Is a Cognitive Map? Organizing Knowledge for Flexible Behavior. *Neuron* 100, 490–509. [PubMed: 30359611]
- Blanchard TC, Hayden BY, and Bromberg-Martin ES (2015). Orbitofrontal cortex uses distinct codes for different choice attributes in decisions motivated by curiosity. *Neuron* 85, 602–614. [PubMed: 25619657]
- Botvinick M, and An J. (2009). Goal-directed decision making in prefrontal cortex: A computational framework. *Adv Neural Inf Process Syst* 21, 169–176. [PubMed: 25258502]
- Bradfield LA, and Hart G. (2020). Rodent medial and lateral orbitofrontal cortices represent unique components of cognitive maps of task space. *Neurosci Biobehav Rev* 108, 287–294. [PubMed: 31743727]
- Breiter HC, Aharon I, Kahneman D, Dale A, and Shizgal P. (2001). Functional imaging of neural responses to expectancy and experience of monetary gains and losses. *Neuron* 30, 619–639. [PubMed: 11395019]
- Butter CM (1969). Perseveration in Extinction and in Discrimination Reversal Tasks Following Selective Frontal Ablations in Macaca Mulatta. *Physiol Behav* 4, 163-&.
- Cardinal RN, Parkinson JA, Hall J, and Everitt BJ (2002). Emotion and motivation: the role of the amygdala, ventral striatum, and prefrontal cortex. *Neurosci Biobehav Rev* 26, 321–352. [PubMed: 12034134]
- Chan SC, Niv Y, and Norman KA (2016). A Probability Distribution over Latent Causes, in the Orbitofrontal Cortex. *J Neurosci* 36, 7817–7828. [PubMed: 27466328]
- Chudasama Y, and Robbins TW (2003). Dissociable contributions of the orbitofrontal and infralimbic cortex to pavlovian autoshaping and discrimination reversal learning: further evidence for the functional heterogeneity of the rodent frontal cortex. *J Neurosci* 23, 8771–8780. [PubMed: 14507977]
- Constantinescu AO, O'Reilly JX, and Behrens TEJ (2016). Organizing conceptual knowledge in humans with a gridlike code. *Science* 352, 1464–1468. [PubMed: 27313047]
- Constantinople CM, Piet AT, Bibawi P, Akrami A, Kopec C, and Brody CD (2019). Lateral orbitofrontal cortex promotes trial-by-trial learning of risky, but not spatial, biases. *Elife* 8.
- Costa VD, and Averbeck BB (2020). Primate Orbitofrontal Cortex Codes Information Relevant for Managing Explore-Exploit Tradeoffs. *J Neurosci* 40, 2553–2561. [PubMed: 32060169]
- Dalton GL, Wang NY, Phillips AG, and Floresco SB (2016). Multifaceted Contributions by Different Regions of the Orbitofrontal and Medial Prefrontal Cortex to Probabilistic Reversal Learning. *J Neurosci* 36, 1996–2006. [PubMed: 26865622]
- Daw ND, Gershman SJ, Seymour B, Dayan P, and Dolan RJ (2011). Model-based influences on humans' choices and striatal prediction errors. *Neuron* 69, 1204–1215. [PubMed: 21435563]

- Delamater AR (2007). The role of the orbitofrontal cortex in sensory-specific encoding of associations in pavlovian and instrumental conditioning. *Ann N Y Acad Sci* 1121, 152–173. [PubMed: 17872387]
- Fellows LK (2007). The role of orbitofrontal cortex in decision making: a component process account. *Ann N Y Acad Sci* 1121, 421–430. [PubMed: 17846161]
- Fellows LK (2011). Orbitofrontal contributions to value-based decision making: evidence from humans with frontal lobe damage. *Ann N Y Acad Sci* 1239, 51–58. [PubMed: 22145875]
- Fellows LK, and Farah MJ (2003). Ventromedial frontal cortex mediates affective shifting in humans: evidence from a reversal learning paradigm. *Brain* 126, 1830–1837. [PubMed: 12821528]
- FitzGerald TH, Seymour B, and Dolan RJ (2009). The role of human orbitofrontal cortex in value comparison for incommensurable objects. *J Neurosci* 29, 8388–8395. [PubMed: 19571129]
- Gallagher M, McMahan RW, and Schoenbaum G. (1999). Orbitofrontal cortex and representation of incentive value in associative learning. *Journal of Neuroscience* 19, 6610–6614. [PubMed: 10414988]
- Gardner MPH, Conroy JC, Sanchez DC, Zhou J, and Schoenbaum G. (2019). Real-Time Value Integration during Economic Choice Is Regulated by Orbitofrontal Cortex. *Curr Biol* 29, 4315–4322 e4314. [PubMed: 31813612]
- Gardner MPH, Conroy JC, Styer CV, Huynh T, Whitaker LR, and Schoenbaum G. (2018a). Medial orbitofrontal inactivation does not affect economic choice. *eLIFE* 7, e38963.
- Gardner MPH, Conroy JS, Shaham MH, Styer CV, and Schoenbaum G. (2017). Lateral orbitofrontal inactivation dissociates devaluation-sensitive behavior and economic choice. *Neuron* 96, 1192–1203. [PubMed: 29154127]
- Gardner MPH, Sanchez D, Conroy JC, Wikenheiser AM, Zhou J, and Schoenbaum G. (2020). Processing in Lateral Orbitofrontal Cortex Is Required to Estimate Subjective Preference during Initial, but Not Established, Economic Choice. *Neuron*.
- Gardner MPH, Schoenbaum G, and Gershman SJ (2018b). Rethinking dopamine as generalized prediction error. *Proc Biol Sci* 285.
- Gershman SJ, and Niv Y. (2013). Perceptual estimation obeys Occam’s razor. *Front Psychol* 4, 623. [PubMed: 24137136]
- Goldman-Rakic PS (1987). Circuitry of primate prefrontal cortex and regulation of behavior by representational memory. In *Handbook of Physiology: The Nervous System*, Mountcastle VB, Plum F, and Geiger SR, eds. (Bethesda, MD: American Physiology Society), pp. 373–417.
- Gottfried JA, O’Doherty J, and Dolan RJ (2003). Encoding predictive reward value in human amygdala and orbitofrontal cortex. *Science* 301, 1104–1107. [PubMed: 12934011]
- Gremel CM, and Costa RM (2013). Orbitofrontal and striatal circuits dynamically encode the shift between goal-directed and habitual actions. *Nat Commun* 4, 2264. [PubMed: 23921250]
- Groman SM, Keistler C, Keip AJ, Hammarlund E, DiLeone RJ, Pittenger C, Lee D, and Taylor JR (2019). Orbitofrontal Circuits Control Multiple Reinforcement-Learning Processes. *Neuron* 103, 734–746 e733. [PubMed: 31253468]
- Hart EE, Sharpe MJ, Gardner MPH, and Schoenbaum G. (2020). Responding to preconditioned cues is devaluation sensitive and requires orbitofrontal cortex during cue-cue learning. *eLIFE* 9, e59998.
- Holland PC, and Gallagher M. (2004). Amygdala-frontal interactions and reward expectancy. *Curr Opin Neurobiol* 14, 148–155. [PubMed: 15082318]
- Howard JD, Reynolds R, Smith DE, Voss JL, Schoenbaum G, and Kahnt T. (2020). Targeted Stimulation of Human Orbitofrontal Networks Disrupts Outcome-Guided Behavior. *Curr Biol* 30, 490–498 e494. [PubMed: 31956033]
- Iversen SD, and Mishkin M. (1970). Perseverative interference in monkeys following selective lesions of the inferior prefrontal convexity. *Exp Brain Res* 11, 376–386. [PubMed: 4993199]
- Izquierdo A. (2017). Functional heterogeneity within rat orbitofrontal cortex in reward learning and decision making. *Journal of Neuroscience* 37, 10529–10540. [PubMed: 29093055]
- Izquierdo A, Suda RK, and Murray EA (2004). Bilateral orbital prefrontal cortex lesions in rhesus monkeys disrupt choices guided by both reward value and reward contingency. *J Neurosci* 24, 7540–7548. [PubMed: 15329401]

- Jang AI, Costa VD, Rudebeck PH, Chudasama Y, Murray EA, and Averbeck BB (2015). The Role of Frontal Cortical and Medial-Temporal Lobe Brain Areas in Learning a Bayesian Prior Belief on Reversals. *J Neurosci* 35, 11751–11760. [PubMed: 26290251]
- Jones B, and Mishkin M. (1972). Limbic lesions and the problem of stimulus-reinforcement associations. *Experimental Neurology* 36, 362–377. [PubMed: 4626489]
- Jones JL, Esber GR, McDannald MA, Gruber AJ, Hernandez A, Mirenski A, and Schoenbaum G. (2012). Orbitofrontal cortex supports behavior and learning using inferred but not cached values. *Science* 338, 953–956. [PubMed: 23162000]
- Kazama A, and Bachevalier J. (2009). Selective aspiration or neurotoxic lesions of orbital frontal areas 11 and 13 spared monkeys' performance on the object discrimination reversal task. *Journal of Neuroscience* 29, 2794–2804. [PubMed: 19261875]
- Keiflin R, Reese RM, Woods CA, and Janak PH (2013). The orbitofrontal cortex as part of a hierarchical neural system mediating choice between two good options. *Journal of Neuroscience* 33, 15989–15998. [PubMed: 24089503]
- Kuwabara M, Kang N, Holy TE, and Padoa-Schioppa C. (2020). Neural mechanisms of economic choices in mice. *Elife* 9.
- Levy DJ, and Glimcher PW (2012). The root of all value: a neural common currency for choice. *Curr Opin Neurobiol* 22, 1027–1038. [PubMed: 22766486]
- Lichtenberg NT, Pennington ZT, Holley SM, Greenfield VY, Cepeda C, Levine MS, and Wassum KM (2017). Basolateral Amygdala to Orbitofrontal Cortex Projections Enable Cue-Triggered Reward Expectations. *J Neurosci* 37, 8374–8384. [PubMed: 28743727]
- Lin Z, Nie C, Zhang Y, Chen Y, and Yang T. (2020). Evidence accumulation for value computation in the prefrontal cortex during decision making. *Proc Natl Acad Sci U S A* 117, 30728–30737. [PubMed: 33199637]
- Liu D, Deng J, Zhang Z, Zhang ZY, Sun YG, Yang T, and Yao H. (2020). Orbitofrontal control of visual cortex gain promotes visual associative learning. *Nat Commun* 11, 2784. [PubMed: 32493971]
- McDannald MA, Esber GR, Wegener MA, Wied HM, Liu TL, Stalnaker TA, Jones JL, Trageser J, and Schoenbaum G. (2014). Orbitofrontal neurons acquire responses to 'valueless' Pavlovian cues during unblocking. *Elife* 3, e02653.
- McDannald MA, Lucantonio F, Burke KA, Niv Y, and Schoenbaum G. (2011). Ventral striatum and orbitofrontal cortex are both required for model-based, but not model-free, reinforcement learning. *J Neurosci* 31, 2700–2705. [PubMed: 21325538]
- McDannald MA, Saddoris MP, Gallagher M, and Holland PC (2005). Lesions of orbitofrontal cortex impair rats' differential outcome expectancy learning but not conditioned stimulus-potentiated feeding. *J Neurosci* 25, 4626–4632. [PubMed: 15872110]
- McDannald MA, Takahashi YK, Lopatina N, Pietras BW, Jones JL, and Schoenbaum G. (2012). Model-based learning and the contribution of the orbitofrontal cortex to the model-free world. *Eur J Neurosci* 35, 991–996. [PubMed: 22487030]
- Miller KJ, Botvinick MM, and Brody CD (2018). Value Representations in Orbitofrontal Cortex Drive Learning, not Choice. *bioRxiv*, 245720.
- Montague PR, and Berns GS (2002). Neural economics and the biological substrates of valuation. *Neuron* 36, 265–284. [PubMed: 12383781]
- Murray EA, O'Doherty JP, and Schoenbaum G. (2007). What we know and do not know about the functions of the orbitofrontal cortex after 20 years of cross-species studies. *J Neurosci* 27, 8166–8169. [PubMed: 17670960]
- Namoodiri VMK, Otis JM, van Heeswijk K, Voets ES, Alghorazi RA, Rodriguez-Romaguera J, Mihalas S, and Stuber GD (2019). Single-cell activity tracking reveals that orbitofrontal neurons acquire and maintain a long-term memory to guide behavioral adaptation. *Nat Neurosci* 22, 1110–1121. [PubMed: 31160741]
- Niv Y. (2019). Learning task-state representations. *Nat Neurosci* 22, 1544–1553. [PubMed: 31551597]
- Noonan MP, Chau BKH, Rushworth MFS, and Fellows LK (2017). Contrasting Effects of Medial and Lateral Orbitofrontal Cortex Lesions on Credit Assignment and Decision-Making in Humans. *J Neurosci* 37, 7023–7035. [PubMed: 28630257]

- Noonan MP, Walton ME, Behrens TE, Sallet J, Buckley MJ, and Rushworth MF (2010). Separate value comparison and learning mechanisms in macaque medial and lateral orbitofrontal cortex. *Proc Natl Acad Sci U S A* 107, 20547–20552. [PubMed: 21059901]
- O’Doherty J, Kringelbach ML, Rolls ET, Hornak J, and Andrews C. (2001). Abstract reward and punishment representations in the human orbitofrontal cortex. *Nat Neurosci* 4, 95–102. [PubMed: 11135651]
- O’Doherty J, Rolls ET, Francis S, Bowtell R, McGlone F, Kobal G, Renner B, and Ahne G. (2000). Sensory-specific satiety-related olfactory activation of the human orbitofrontal cortex. *Neuroreport* 11, 893–897. [PubMed: 10757540]
- Ostlund SB, and Balleine BW (2007). Orbitofrontal cortex mediates outcome encoding in Pavlovian but not instrumental learning. *Journal of Neuroscience* 27, 4819–4825. [PubMed: 17475789]
- Padoa-Schioppa C. (2011). Neurobiology of economic choice: a goods-based model. *Annual Review of Neuroscience* 34, 333–359.
- Padoa-Schioppa C, and Assad JA (2006). Neurons in orbitofrontal cortex encode economic value. *Nature* 441, 223–226. [PubMed: 16633341]
- Padoa-Schioppa C, and Assad JA (2008). The representation of economic value in the orbitofrontal cortex is invariant for changes in menu. *Nature Neuroscience* 11, 95–102. [PubMed: 18066060]
- Padoa-Schioppa C, and Conen KE (2017). Orbitofrontal cortex: a neural circuit for economic decisions. *Neuron* 96, 736–754. [PubMed: 29144973]
- Padoa-Schioppa C, Jandolo L, and Visalberghi E. (2006). Multi-stage mental process for economic choice in capuchins. *Cognition* 99, B1–B13. [PubMed: 16043168]
- Panayi M,C, Khamassi M, and Killcross S. (2021). The rodent lateral orbitofrontal cortex as an arbitrator selecting between model-based and model-free learning systems. *Behavioral Neuroscience*.
- Panayi MC, and Killcross S. (2018). Functional heterogeneity within the rodent lateral orbitofrontal cortex dissociates outcome devaluation and reversal learning deficits. *Elife* 7.
- Panayi MC, and Killcross S. (2020). The role of the rodent lateral orbitofrontal cortex in simple Pavlovian cue-outcome learning depends on training experience. *bioRxiv*.
- Parkes SL, Ravassard PM, Cerpa JC, Wolff M, Ferreira G, and Coutureau E. (2018). Insular and Ventrolateral Orbitofrontal Cortices Differentially Contribute to Goal-Directed Behavior in Rodents. *Cereb Cortex* 28, 2313–2325. [PubMed: 28541407]
- Pickens CL, Saddoris MP, Gallagher M, and Holland PC (2005). Orbitofrontal lesions impair use of cue-outcome associations in a devaluation task. *Behav Neurosci* 119, 317–322. [PubMed: 15727536]
- Rao RP (2010). Decision making under uncertainty: a neural model based on partially observable markov decision processes. *Front Comput Neurosci* 4, 146. [PubMed: 21152255]
- Riceberg JS, and Shapiro ML (2012). Reward stability determines the contribution of orbitofrontal cortex to adaptive behavior. *J Neurosci* 32, 16402–16409. [PubMed: 23152622]
- Riceberg JS, and Shapiro ML (2017). Orbitofrontal Cortex Signals Expected Outcomes with Predictive Codes When Stable Contingencies Promote the Integration of Reward History. *J Neurosci* 37, 2010–2021. [PubMed: 28115481]
- Rich EL, and Wallis JD (2016). Decoding subjective decisions from orbitofrontal cortex. *Nat Neurosci* 19, 973–980. [PubMed: 27273768]
- Roesch MR, Taylor AR, and Schoenbaum G. (2006). Encoding of time-discounted rewards in orbitofrontal cortex is independent of value representation. *Neuron* 51, 509–520. [PubMed: 16908415]
- Rolls ET (1996). The orbitofrontal cortex. *Philos Trans R Soc Lond B Biol Sci* 351, 1433–1443; discussion 1443–1434. [PubMed: 8941955]
- Rudebeck PH, and Murray EA (2011a). Balkanizing the primate orbitofrontal cortex: distinct subregions for comparing and contrasting values. *Annals of the New York Academy of Sciences* 1239, 1–13. [PubMed: 22145870]
- Rudebeck PH, and Murray EA (2011b). Dissociable effects of subtotal lesions within the macaque orbital prefrontal cortex on reward-guided behavior. *J Neurosci* 31, 10569–10578. [PubMed: 21775601]

- Rustichini A, Conen KE, Cai X, and Padoa-Schioppa C. (2017). Optimal coding and neuronal adaptation in economic decisions. *Nat Commun* 8, 1208. [PubMed: 29084949]
- Schoenbaum G, and Eichenbaum H. (1995). Information coding in the rodent prefrontal cortex. I. Single-neuron activity in orbitofrontal cortex compared with that in pyriform cortex. *J Neurophysiol* 74, 733–750. [PubMed: 7472378]
- Schoenbaum G, Nugent S, Saddoris MP, and Setlow B. (2002). Orbitofrontal lesions in rats impair reversal but not acquisition of go, no-go odor discriminations. *Neuroreport* 13, 885–890. [PubMed: 11997707]
- Schoenbaum G, Roesch MR, Stalnaker TA, and Takahashi YK (2009). A new perspective on the role of the orbitofrontal cortex in adaptive behaviour. *Nat Rev Neurosci* 10, 885–892. [PubMed: 19904278]
- Schoenbaum G, and Setlow B. (2001). Integrating orbitofrontal cortex into prefrontal theory: common processing themes across species and subdivisions. *Learn Mem* 8, 134–147. [PubMed: 11390633]
- Schuck NW, Cai MB, Wilson RC, and Niv Y. (2016). Human Orbitofrontal Cortex Represents a Cognitive Map of State Space. *Neuron* 91, 1402–1412. [PubMed: 27657452]
- Sharpe MJ, Stalnaker T, Schuck NW, Killcross S, Schoenbaum G, and Niv Y. (2019). An Integrated Model of Action Selection: Distinct Modes of Cortical Control of Striatal Decision Making. *Annu Rev Psychol* 70, 53–76. [PubMed: 30260745]
- Stalnaker TA, Cooch NK, McDannald MA, Liu TL, Wied H, and Schoenbaum G. (2014). Orbitofrontal neurons infer the value and identity of predicted outcomes. *Nat Commun* 5, 3926. [PubMed: 24894805]
- Stolyarova A, and Izquierdo A. (2017). Complementary contributions of basolateral amygdala and orbitofrontal cortex to value learning under uncertainty. *Elife* 6.
- Takahashi YK, Roesch MR, Wilson RC, Toreson K, O'Donnell P, Niv Y, and Schoenbaum G. (2011). Expectancy-related changes in firing of dopamine neurons depend on orbitofrontal cortex. *Nat Neurosci* 14, 1590–1597. [PubMed: 22037501]
- Teitelbaum H. (1964). A Comparison of Effects of Orbitofrontal and Hippocampal Lesions Upon Discrimination Learning and Reversal in the Cat. *Exp Neurol* 9, 452–462. [PubMed: 14188532]
- Thorpe SJ, Rolls ET, and Maddison S. (1983). The orbitofrontal cortex: neuronal activity in the behaving monkey. *Exp Brain Res* 49, 93–115. [PubMed: 6861938]
- Ting CC, Yu CC, Maloney LT, and Wu SW (2015). Neural mechanisms for integrating prior knowledge and likelihood in value-based probabilistic inference. *J Neurosci* 35, 1792–1805. [PubMed: 25632152]
- Vertechi P, Lottem E, Sarra D, Godinho B, Treves I, Quendera T, Oude Lohuis MN, and Mainen ZF (2019). Inference based decisions in a hidden state foraging task: differential contributions of prefrontal cortical areas. *bioRxiv*, 679142.
- Vilares I, Howard JD, Fernandes HL, Gottfried JA, and Kording KP (2012). Differential representations of prior and likelihood uncertainty in the human brain. *Curr Biol* 22, 1641–1648. [PubMed: 22840519]
- Wallis JD (2007). Orbitofrontal cortex and its contribution to decision-making. *Annu Rev Neurosci* 30, 31–56. [PubMed: 17417936]
- Wallis JD, and Miller EK (2003). Neuronal activity in primate dorsolateral and orbital prefrontal cortex during performance of a reward preference task. *Eur J Neurosci* 18, 2069–2081. [PubMed: 14622240]
- Walton ME, Behrens TE, Buckley MJ, Rudebeck PH, and Rushworth MF (2010). Separable learning systems in the macaque brain and the role of orbitofrontal cortex in contingent learning. *Neuron* 65, 927–939. [PubMed: 20346766]
- Walton ME, Behrens TE, Noonan MP, and Rushworth MF (2011). Giving credit where credit is due: orbitofrontal cortex and valuation in an uncertain world. *Ann N Y Acad Sci* 1239, 14–24. [PubMed: 22145871]
- Wang F, Howard JD, Voss JL, Schoenbaum G, and Kahnt T. (2020). Targeted Stimulation of an Orbitofrontal Network Disrupts Decisions Based on Inferred, Not Experienced Outcomes. *J Neurosci*.

- West EA, DesJardin JT, Gale K, and Malkova L. (2011). Transient inactivation of orbitofrontal cortex blocks reinforcer devaluation in macaques. *J Neurosci* 31, 15128–15135. [PubMed: 22016546]
- Wilson RC, Takahashi YK, Schoenbaum G, and Niv Y. (2014). Orbitofrontal cortex as a cognitive map of task space. *Neuron* 81, 267–279. [PubMed: 24462094]
- Zhou J, Gardner MPH, Stalnaker TA, Ramus SJ, Wikenheiser AM, Niv Y, and Schoenbaum G. (2019a). Rat Orbitofrontal Ensemble Activity Contains Multiplexed but Dissociable Representations of Value and Task Structure in an Odor Sequence Task. *Curr Biol* 29, 897–907 e893. [PubMed: 30827919]
- Zhou J, Jia C, Montesinos-Cartegen M, Gardner MPH, Zong W, and Schoenbaum G. (2020). Evolving schema representations in orbitofrontal ensembles during learning. In review.
- Zhou J, Montesinos-Cartegen M, Wikenheiser AM, Gardner MPH, Niv Y, and Schoenbaum G. (2019b). Complementary Task Structure Representations in Hippocampus and Orbitofrontal Cortex during an Odor Sequence Task. *Curr Biol* 29, 3402–3409 e3403. [PubMed: 31588004]

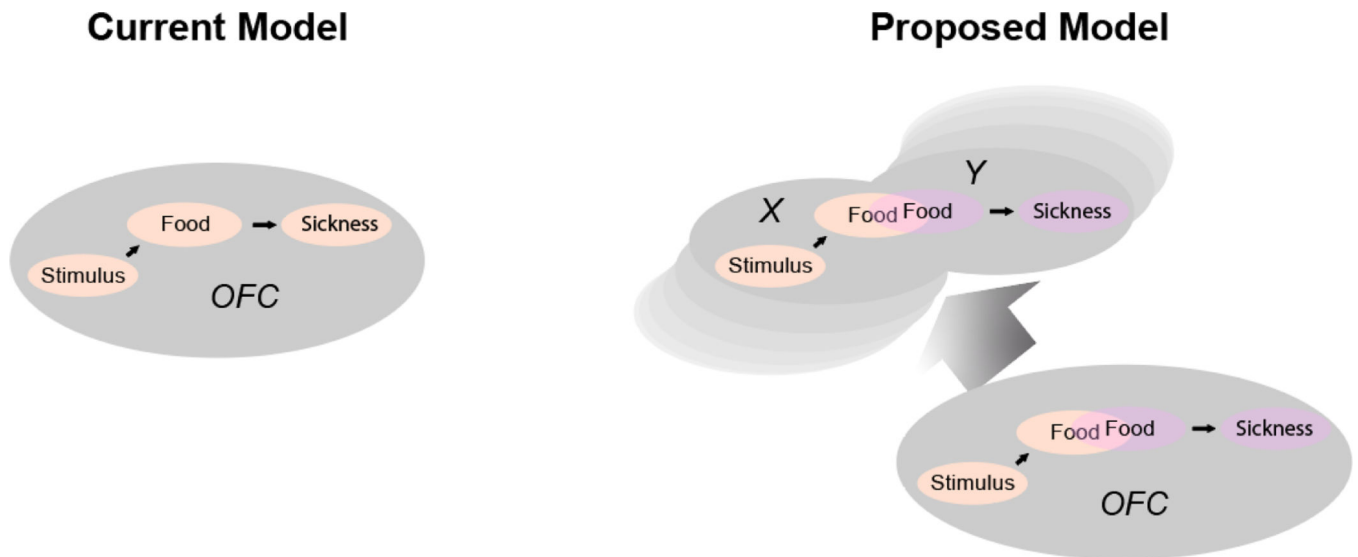


Figure 1: Illustration contrasting current (left) and proposed (right) model of OFC function in the probe test of the classic illness-based Pavlovian devaluation paradigm.

In the probe test, a conditioned stimulus is presented in isolation, without reward, following devaluation of its predicted reward by pairing with illness. Rats, monkeys, and humans reduce responding to the conditioned stimulus compared to control cues, and manipulations that affect OFC function in the probe test selectively disrupt this normal reduction in responding. The current model holds that this occurs because the OFC is necessary for representing the associative structure or cognitive map of the task, or for maintaining a pointer to the current location in this map; the proposed model is similar in that it implicates the OFC in mapping, however it holds that OFC allows maps already established of the disparate learning episodes to be knitted together, integrated, which is a prerequisite for their use to guide behavior in the probe test. Critically, however, the maps are located in networks (X, Y) that are functionally separate, at least in part, from OFC. Thus once knitted together, perhaps through formation of a new assimilated state ($S_{food}^X, S_{food}^Y \rightarrow S_{food}^Z$), static use of the integrated map does not depend on OFC. This predicts that subsequent performance after this probe test would be insensitive or less sensitive to OFC manipulations. It also predicts that OFC would potentially play a role in mapping episodes in earlier learning.