Journal Pre-proof

Rodent medial and lateral orbitofrontal cortices represent unique components of cognitive maps of task space

Laura A. Bradfield, Genevra Hart

PII: S0149-7634(19)30528-7

DOI: https://doi.org/10.1016/j.neubiorev.2019.11.009

Reference: NBR 3597

To appear in: *Neuroscience and Biobehavioral Reviews*

Received Date: 17 June 2019

Revised Date: 14 November 2019

Accepted Date: 15 November 2019

Please cite this article as: Bradfield LA, Hart G, Rodent medial and lateral orbitofrontal cortices represent unique components of cognitive maps of task space, *Neuroscience and Biobehavioral Reviews* (2019), doi: https://doi.org/10.1016/j.neubiorev.2019.11.009

Rodent medial and lateral orbitofrontal cortices represent unique components of cognitive maps of task space.

Laura A. Bradfield[1,2*] and Genevra Hart[3].

1. **Centre for Neuroscience and Regenerative Medicine, University of Technology Sydney, NSW, 2007, Australia.**
2. **St. Vincent's Centre for Applied Medical Research, St. Vincent's Hospital Sydney Limited, Sydney, NSW 2010, Australia.**
3. **Decision Neuroscience Laboratory, School of Psychology, University of New South Wales, NSW 2052, Australia.**

Corresponding Author:

Dr. Laura Bradfield at laura.bradfield@uts.edu.au

Centre for Neuroscience and Regenerative Medicine,
School of Life Sciences,
University of Technology Sydney (St. Vincent's Campus)
405 Liverpool St
Darlinghurst, NSW 2010.
PH: 02 83824950

Highlights

- First consideration of how the medial and lateral orbitofrontal cortices uniquely represent components of cognitive maps of task space
- Posits that medial orbitofrontal cortex represents terminal states (usually outcomes of actions)
- Posits that lateral orbitofrontal cortex represents initial states
- Medial and lateral orbitofrontal cortex achieve their respective functions somewhat independently, and somewhat interdependently

Abstract

The orbitofrontal cortex (OFC) has been proposed to function as a cognitive map of task space: a mental model of the steps involved in a task. This idea has proven popular because it

provides a cohesive explanation for a number of disparate findings regarding the OFC's role in a broad array of tasks. Concurrently, evidence has begun to reveal the functional heterogeneity of OFC subregions, particularly the medial and lateral OFC. How these subregions uniquely contribute to the OFC's role as a cognitive map of task space, however, has not been explored. Here we propose that, in rodents, the lateral OFC represents the agent's initial position within that task map (i.e. initial state), determining which actions are available as a consequence of that position, whereas the medial OFC represents the agent's desired future position within the task map (i.e. terminal state), influencing which actions are selected to achieve that position. We argue that these processes are achieved somewhat independently and somewhat interdependently, and are achieved through similar but non-identical circuitry.

Keywords/Phrases: Medial orbitofrontal cortex; Lateral orbitofrontal cortex; Decision-making; Cognitive map of task space; Goal-directed action; Model-based reinforcement learning

Introduction

The notion that goal-directed decision-making relies on a 'cognitive map' or 'mental model' of the different steps involved in a task has gained significant traction in recent years. A number of studies have begun to identify the putative neural structures and circuits supporting this function, with several lines of evidence suggesting the orbitofrontal cortex (OFC) plays a key role (*1–4*). Recent studies have also demonstrated a high degree of functional heterogeneity within the OFC (*5–11*), yet how these functionally distinct OFC subregions might differentially contribute to its function as a cognitive map of task space has never been considered.

In this review we argue that, in rodents at least, the majority of findings are consistent with the lateral OFC inferring the animal's *initial* position within a task and thus what actions are available for selection as a consequence (i.e. initial state), whereas the medial OFC provides information about what *future* position within task space is likely if the animal selects a particular action (i.e. the terminal state, represented by what goal will be achieved). To put it another way, the function of lateral OFC has been likened to the dropping of a google maps 'pin' to determine one's position within the task (*3*), information that can then be fed to downstream structures (notably the striatum, (*12*, *13*)) to provide contextual information for decision-making and action selection. Using this same analogy, we suggest that the medial OFC can also influence actions (routes), but that it does so by providing information about the desired final destination (i.e. by representing the final goal). In this way, the functions of lateral OFC and medial OFC appear to be highly related and even interdependent – it is difficult to determine which route is most desirable without knowing the final destination – however they are not interchangeable. Finally, we argue that both initial and terminal state information appears to be particularly OFC-dependent when it relies on partially unobservable information, because several studies have shown that actions can be selected accurately in spite of OFC inactivation (lateral and/or medial) when all of the necessary state information is observable. It should be noted that although this is primarily a review of rodent studies, where the evidence most clearly points to the account we have outlined here, there has been some evidence that human OFC represents cognitive maps in a similar manner (*2*, *14*).

The Role of States in Decision-Making

Within the laboratory, decision-making tends to be operationalized as the selection of goal-directed actions. Such actions are effortful, flexibly deployed, and must be motivated by both the value of the outcome or goal (goal criterion), and the contingency between the action and outcome (instrumental criterion (*15*)). Such actions are held in contrast to habitual actions that are elicited

automatically without regard for the outcome. Both goal-directed and habitual processes are integral to everyday functioning, yet are neurally separable as shown by experiments that have demonstrated deficits in one system but not the other as a result of damage to distinct brain structures (*16*–*18*).

Computationally, reinforcement learning (RL) models have posited that goal-directed actions are akin to model-based models (*19*). This means that in order to exert goal-directed control, animals build a mental model (i.e. a cognitive map of task space) of the various steps that are involved in a task. Animals can then simulate their prospective journey through the steps of the model to estimate the probability of obtaining their goal. Habits, on the other hand, are thought to be more akin to model-free models (*19*) (although this position is not without controversy (*20*), because they do not rely on cognitive maps in the same manner. Instead, they are elicited in response to various stimuli in the environment. Model-free RL posits that habits become valuable over the course of learning by being paired with reinforcers, causing value to be 'cached' at the level of the action itself. Unlike model-based RL, altered reinforcer value cannot immediately modify the cached value of actions until that reinforcer is experienced as a consequence of the action taken. Further, model-based but not model-free RL recognises the identity of the outcome; model-free RL recognises only its value.

Cognitive task maps consist of 'states', the concept of which has been useful for conceptualizing how action selection is influenced by internal and external cues. States are roughly akin to 'situations' determined by alterations in external stimuli, such as a light or a noise, or by internal information garnered from memory, or some combination of both. Within each state there are certain actions available and taking a particular action will transition the agent into the next state. Figure 1 shows an example of a state representation for the task of making a cup of tea. In this example, an individual might initiate making a cup of tea whilst in State 1, which they enter into upon walking into the kitchen and observing the kettle. For this task map then, State 1 is the initial

state. That individual then switches the kettle on, causing the person to transition into the 'kettle on' state (State 2) determined by the noise, sight, and feeling of the kettle switching on. The agent continues in this manner until the tea is finally being drunk, at which point the terminal 'goal' state is reached (Figure 1).

State representations are complex, and often require the integration of several sensory inputs with internal information. In our tea example above, the 'kettle on' state is determined by combining information from (at least) three sensory modalities (the noise, sight, and feeling of the kettle switching on) with the internal knowledge that 'tea' is the final goal. As has been noted elsewhere[1,18], the anatomical circuitry of the OFC is particularly well placed to integrate external and internal information in this way, as it receives inputs from many sensory areas (e.g. auditory cortex, olfactory tubercle), as well as areas central to learning and memory such as the hippocampus, amygdala, and ventral tegmental area (*21*–*23*). Moreover, the OFC has substantial outputs onto striatal action selection centres (*23*), providing a pathway via which OFC-dependent state representations could directly influence the selection of actions. Both medial and lateral subregions reflect this pattern of connectivity to some extent, with a significant degree of overlap, but there are also some differences. For example, although both medial and lateral OFC project to basolateral amygdala, only medial OFC appears to project to the central amygdala (*22*, *23*). Similar differences exist with regards to cortico-cortical, hippocampal, striatal and other thalamic connections with to medial versus lateral OFC (*21*–*23*). These differences are sometimes subtle (i.e. projections to different portions of the same structure), but are present nonetheless. Overall, these studies suggest the anatomical suitability of both medial and lateral OFC for representing state, but the differences in their connectivity suggest that their role in doing so is not identical.

Thus, there is anatomical support for our claim that the lateral OFC represents initial states for action selection (state S1 in Figure 1) whereas medial OFC represents terminal states (i.e. the

goal, state S7 in Figure 1). The empirical evidence for these claims is presented below, with a summary of the relevant findings presented in Table 1.

Instrumental Outcome Devaluation

Instrumental outcome devaluation is considered the gold standard protocol to assess whether actions are under goal-directed (model-based) control. The typical two-outcome version of this task is summarised in Table 2: In Stage 1, animals are trained to press two levers (A1 and A2) for two outcomes (O1 and O2). They are then given the opportunity to eat one of these outcomes to satiety to reduce its value relative to the alternate outcome[12]. Immediately following satiety treatment rats are placed back into the operant chambers and allowed to respond freely on each lever under extinction (i.e. with no outcomes delivered). According to model-based RL, if animals are selecting actions according to the value of prospective outcomes, then they should preferentially respond on the lever associated with the valued outcome relative to the lever associated with the devalued outcome (valued > devalued).

Critical to this task, animals must be able to accurately represent the outcomes of their actions (i.e. the terminal states). This is represented in Fig 2A. We have demonstrated (*4*, *6*) that disruption to medial OFC functioning (anterior medial OFC in particular) via either permanent excitotoxic lesions, or temporary inactivation across test using chemogenetics (hM4Di designer receptors exclusively activated by designer drugs: DREADDs), abolished rats' capacity to show goal-directed choice on an outcome devaluation task. Instead, rats performed the two available actions (devalued and non-devalued) at equal rates, and this result is represented in Fig 2B. Importantly, when an almost-identical devaluation test was conducted, except that each outcome was actually delivered as a result of each lever press on test (i.e. outcomes were observable), medial OFC inactivated animals showed intact devaluation (valued > devalued). This pattern of responding matched predictions from our computational modelling in which we removed the assumption of the

ability to infer terminal states defined by absent outcomes (the code for which can be found here: https://github.com/adezfouli/OFCSim, and the state space representation for which is in Figure 2A).

In contrast to this role for medial OFC, a number of experiments have demonstrated that inactivating the lateral OFC via either pre- or post-training lesions (*24*, *25*), chemogenetic inactivation of lateral OFC cell bodies (*26*) or its connections with the basolateral amygdala (*27*) leave instrumental outcome devaluation intact, represented in Fig 2B. This is consistent with the argument that lateral OFC represents unobservable initial states, because both actions (left and right lever press) are available simultaneously in a single initial state (state 1 in Figure 2A). Thus, the levers are observable such that available actions need not be inferred from memory. In two exceptions to these findings, Gremel and Costa (*28*, *29*) used a different kind of instrumental outcome devaluation task to infer that the lateral OFC might regulate instrumental outcome devaluation. The state space representation of this task is presented in Fig 2C. They trained animals to press a single lever in one context by using a random interval (RI) schedule to promote habits, and trained the same lever in another context using a random ratio (RR) schedule to promote goal-directed responding. On test, devaluation was demonstrated (Valued > Devalued) for lateral OFC-intact mice in the RR but not RI context, whereas lateral OFC inactivated mice (by way of lesion, chemogenetics, or optogenetics) did not show devaluation in either context (Valued = Devalued, Fig 2D). As a result, the authors concluded that the lateral OFC is necessary for instrumental outcome devaluation.

Although this remains a possibility, the involvement of lOFC in this task can equally be explained by its representation of initial state information. Specifically, it is likely that the alteration in reward schedules (RI vs. RR) between contexts required the formation of two different states to disambiguate these different contingencies (*30*, *31*) (e.g. State 1 in context 1, State 2 in context 2, Fig 2C). These latent states are internally generated, and function to retrieve the appropriate contingency and reduce interference by the competing contingency. As argued by Wilson et al, animals with whole/lateral OFC lesions are unable to partition information based on latent states

and so would attribute competing sets of contingencies to the same 'default' state such that they interfere with each other. Thus, we suggest that a similar process could be in effect here, such that intact animals partitioned information appropriately and demonstrated intact devaluation in the appropriate (goal-directed/RR) context, animals without a functional lateral OFC likely generalised between these states, or at the very least suffered interference between them.

Together, the results of these experiments neatly demonstrate an important difference between the role of initial and terminal state representation in action selection and the respective roles of lateral OFC and medial OFC in each. That is, as illustrated in Figure 2C, in the two action, two outcome devaluation task, initial state information is not necessary because both levers are observable, but animals must represent the different terminal states (O1 and O2, which are not presented on test) to select the appropriate action (valued > devalued). This capacity is profoundly disrupted by inactivations of medial but not lateral OFC (*4*, *6*, *24–26*) (Figure 2 A-B). By contrast, an intact lateral OFC is only necessary for instrumental devaluation when animals must use initial state information to partition learning about two distinct action-outcome contingencies in order to apply the appropriate response strategy (*28*, *29*) (Fig 2C-D).

Contingency Reversal

In a typical contingency reversal protocol, the animal is initially trained such that one action-outcome (A1-O1) contingency is rewarded but another is not (A2-), and this relationship is later reversed (A1-, A2-O1, see Table 2). Over time, intact animals will alter their response preference according to the currently rewarded contingency. Figure 3A-B shows a cognitive map of task space for reversal learning. According to this representation, when animals enter the reversal training

phase in Fig 3B, they partition the new contingency into a separate state (S2) from the original contingency (S1) to prevent interference. Consistent with the presumed role of lateral OFC in initial state representation, there are several demonstrations that lateral/ventrolateral OFC inactivations (mainly lesions) impair contingency reversal performance, generally being slower to acquire the reversed response and making more perseverative errors (*32*–*35*), shown in Figure 3C. Animals with lateral OFC lesions are presumed unable to partition information in this way, instead attributing both contingencies to the same 'default' state such that they suffer interference leading to impaired task performance. Likewise, lateral OFC lesions disrupt reversal learning when outcome identities are reversed (*26*). Rats were trained on two response-outcome pairings, A1-O1 and A2-O2, and then these contingencies were reversed; A1-O2, A2-O1, under chemogenetic inactivation of the lateral OFC. In between each phase of training, animals were tested for outcome devaluation to assess which response-outcome association was dominant. Whereas instrumental devaluation was intact following initial instrumental training (as mentioned above), when the lever press-outcome contingencies were reversed, lateral OFC inactivation did impair devaluation performance.

There are relatively fewer investigations of the role of the medial OFC in reversal learning, and several of the studies that have investigated it have done so in probabilistic tasks (*33*, *36*), confounding their interpretation. As we have argued previously (*4*) in tasks such as probabilistic learning, delay discounting, and even progressive ratio tasks, where medial OFC lesions are often found to impair performance (*33*, *36*–*38*), this is likely because each of these tasks require the estimation of likely outcomes when those outcomes are not directly observable. Thus it is straightforward to understand how removing the medial OFC would impair the ability to properly make such estimations and ultimately impair performance.

Gourley et al. (*38*) did, however, examine the effect of medial OFC lesions on a more straightforward reversal task in mice, and found them to impair reversal learning. The manner of this impairment was somewhat different, however, to that resulting from whole or ventrolateral OFC

lesions in the same study. Specifically, whereas ventrolateral OFC-lesioned mice were slower to learn the new (reversed) contingencies, medial OFC-lesioned mice were able to learn the new contingencies as readily as controls (represented in Figure 3C) but additionally performed more perseverative responses on the previously rewarded nose-poke. The fact that lesioned mice learned the new contingencies as quickly as controls suggests they were able to partition conflicting contingencies according to initial state (see[1] and Figure 3A-B). Instead, in line with our framework, the deficit may lie in their inability to appropriately extinguish the previously rewarded response, due to an inability to retrieve a representation of the expected outcome and appropriately assign an error-based learning mechanism to drive model-based extinction. Therefore, mice with medial OFC lesions would have had to rely on model-free processes alone to extinguish the initial contingency, causing slower extinction and more perseverative responses relative to controls.

More recently, Hervig, Robbins and colleagues (*39*) found that pharmacological inactivation of medial OFC *improved* reversal learning on a visual discrimination task, particularly in the early phase of learning. As they noted, this would be precisely the result if the medial OFC were responsible for mentally representing unobservable outcomes. This is because the animals with mOFC inactivation would be unable to recall the prior (pre-reversal) contingencies, and would thus experience less interference in acquiring the new (reversed) contingencies than controls. Indeed, interference between initial and reversed contingencies would be expected to be highest in controls during the early phases of reversal, and it is this phase in which they differed most from medial OFC-inactivated animals. These authors further found that lateral OFC inactivation impaired reversal performance relative to controls, in line with the other findings we have noted above. They also found, however, that lateral OFC inactivations appeared to improve the learning of a novel visual discrimination task, which also involves new contingency learning and thus would imply a shift to a new state. If replicable, therefore, this finding might suggest that a narrowing of the current account is necessary, in a manner suggesting that lateral OFC state partitioning is only necessary when there

is direct conflict between contingencies such as in reversal learning, but not when learning new associations.

Contingency Degradation

Like outcome devaluation, sensitivity to contingency degradation is considered characteristic of goal directed instrumental actions. A typical protocol is presented in Table 2. In Stage 1, animals are trained on two action-outcome contingencies, and then in Stage 2, animals receive additional, unsignalled deliveries of one of the outcomes (A1-O1/O1). This effectively reduces the contingency between that action and outcome (*4*, *40*). An animal using goal-directed control will reduce its responding on the degraded lever, whilst maintaining responding on the non-degraded lever (in accordance with the contingency criterion). We have previously argued that contingency degradation relies on the ability to separate the different contingencies learned in Stage 1 (original contingencies) and Stage 2 (degraded/nondegraded contingencies) into separate internal states (*40*). Importantly, however, because outcomes are present and observable throughout this task, an inability to infer terminal outcome states should not be necessary for intact degradation performance. Thus, our framework predicts that the lateral OFC but not the medial OFC should be critical for performance on this task.

There is some support for this prediction. Specifically, we have found that an intact medial OFC was not necessary for contingency degradation performance using this design (*4*). Although the role of the lateral OFC in this exact procedure has not been tested, the selective knockdown of brain derived neurotrophic factor (BDNF) in the lateral OFC in mice has been found to abolish nominally instrumental (nose-poke) contingency degradation performance (*41*). It should be noted however, that the same authors (*42*) used the same procedure and found that chemogenetic inactivation of ventrolateral OFC had no effect on degradation learning, but it did have an effect on sustained degradation performance. Although our framework predicts an effect of lateral OFC inactivation on

both learning and performance of degradation, it is possible that the ventrolateral target of these inactivations has a slightly different function to the more lateral portion. Even more recent findings from the same group (*43*) support this notion, because when the ventrolateral OFC was again targeted in the same procedure, but viral spread was increased (and thus likely to have spread to the more lateral OFC), degradation learning was once again impaired. Alternately, the reason for this discrepancy may be methodological, as these studies used a somewhat different procedure to that we have described (here and in Table 2), in which responding no longer produced any contingent outcomes (contingency of -1 rather than 0). Thus, it is possible that this more acute version of the task encouraged learning according to model-free processes that could proceed in the absence of state inference.

Taken together, therefore, these studies support the notion that lateral OFC regulates contingency degradation, but more studies are needed using a purely instrumental (e.g. lever press) version of the task that does not involve extinction. For this task, our framework makes the clear prediction that lateral but not medial OFC inactivation should impair performance.

Pavlovian-instrumental Transfer

Although the results of OFC dysfunction in each of the tasks reviewed above provides clear evidence of the functional heterogeneity of medial and lateral OFC, this is not always the case. In specific Pavlovian-instrumental transfer (PIT; Table 2), for example, lesions of both medial and lateral OFC have been found to impair performance (*4*, *6*, *24*). In this task, rats are first trained to associate two stimuli with two unique outcomes (S1-O1, S2-O2). In the second, instrumental, phase of training, they are trained to perform two distinct actions for those same outcomes (A1-O1, A2-O2). On test, when presented with the stimuli and levers together for the first time, but no food outcomes, animals will typically respond on the lever associated with the same outcome as the stimulus that is being presented (e.g. S1: A1 > A2 due to the association of both S1 and A1 with O1).

We suggest that specific PIT requires the inference of both initial and terminal state representations to guide performance, which is why lesions of both have been found to affect its performance. Specifically, initial state information is necessary on test because this is the first time that the Pavlovian and instrumental contingency information learned during training must be inferred from memory, and integrated to drive responding on the correct lever. In addition, unobservable terminal state information about which lever earned which outcome must be inferred to drive responding on the correct lever. Thus, consistent with the aforementioned findings, our framework predicts a role for both lateral and medial OFC in this task.

Do lateral and medial OFC represent state information in the same way for Pavlovian tasks?

All of the studies we have reviewed so far have been either entirely, or predominantly, instrumental in nature. However, a number of studies have shown that the lateral OFC also appears to regulate learning and/or expression in a number of Pavlovian tasks such as (Pavlovian) devaluation (*25*, *44–46*), overexpectation (*47*), identity unblocking (*48*), and preconditioning (*49*). Furthermore, if the same framework outlined above were applied to these tasks, then it would be difficult to argue that lateral OFC is representing initial but not terminal states. In Pavlovian devaluation, for example, this task is usually employed in the absence of any conflicting contingencies. Thus, if we apply the same logic outlined above for the instrumental version of this task, lateral OFC inactivations should leave it intact, but this is clearly not the case (*25*, *44–46*).

We believe that the reason for this might lie in the fundamental difference in how the cognitive map of a Pavlovian versus an instrumental devaluation task is mentally represented. Specifically, in an instrumental task an action must be performed in order to transition through states (referred to as the 'transition function' (*50*)). In our tea example above (Figure 1), the individual switches on the kettle to transition into the 'kettle on' state and the individual will not transition to that state if that action is not performed. This is not the case in Pavlovian studies in

which a cue (e.g. a tone) is typically presented, and food is delivered to the food receptacle regardless of the animal's actions. Thus, the animal transitions into a 'food delivered' state without performing any action (note that entry into the food receptacle could be seen as potentially instrumental, but could also be a Pavlovian conditioned response to the sight of the food). As such, it is not clear exactly what would lead to the state-to-state transition in a 'model-based' cognitive map of a Pavlovian task (time? Sequence of events?), and as has been noted in articles regarding the potential model-based nature of some Pavlovian responses (*51*, *52*) the algorithm – and hence the transition function – for model-based Pavlovian learning is still unknown.

This is relevant to the current argument, because the separation of initial and terminal states in instrumental studies is reliant on the action-based transitions between them. In the absence of such transitions, it is difficult to say what the potential role of medial versus lateral OFC in Pavlovian studies might be. Further, it has been argued (*51*, *53*) that in Pavlovian conditioning, the stimuli may take on properties of the outcome in a way that does not occur in instrumental conditioning (referred to as 'stimulus substitution'), and indeed phenomena such as conditioned reinforcement (*54*) and second-order conditioning (*55*) support this claim. If this is indeed the case, then this would suggest that it is difficult (if not impossible) to separate initial versus terminal state information in such studies, because the presentation of a cue itself has taken on properties of both the cue and outcome – i.e. the initial and terminal states. As such, there may be no role for the medial OFC in Pavlovian learning, as this is fulfilled by the Pavlovian cue which is fully observable. Instead, such learning relies on the lateral OFC to retrieve a representation of task structure, but only when Pavlovian cues are ambiguous (*3*). Alternatively, it is possible that the framework we have outlined here applies solely to instrumental studies, and that medial versus lateral OFC might have very different roles in Pavlovian studies that may or may not be related to state representation. More research is clearly needed to tease apart these many possibilities.

Medial versus lateral orbitofrontal cortical function across species

Although the studies reviewed here have primarily been conducted in rodents, we believe it is possible that these relative roles for medial and lateral OFC could apply across species. The question of degree of OFC homology between rodents and primates (including humans) in particular is still under debate, although it is generally accepted the OFC is much more developed in primates (*56*). Most notable are differences with regards to cytoarchitecture: rodents differ from primates in that they possess only agranular but not dysgranular or granular cortex (*57*) (although its composition is similar to that of primates (*58*)). Anatomical circuitry does provide some homological support as OFC-striatal (*13*) and OFC-thalamic (*59*) connections appear to be topographically similar across species. Most notably for current purposes, there is abundant evidence that the medial and lateral OFC share somewhat overlapping but distinct projection networks in both rodents (*23*, *60*) and primates (*58*, *59*).

There are also similarities with regards to the functions ascribed to human/primate and rodent OFC and their consequences for associative learning mechanisms. For instance, there is substantial evidence that the primate OFC is involved in encoding expected or abstract reinforcer value (*61*), choice alternatives (*62*, *63*), or the sensory properties of reinforcers (*56*, *64*), but not the specific motor response chosen (*65*). These findings are consistent with our claims because the inference of a terminal state by medial OFC involves integrating expected and abstract reinforcer value, as well as the sensory properties of reinforcers. Likewise, if lateral OFC provides information about initial task state, as we have also claimed, this would be reflected in neuronal responses relating to choice alternatives. Critically, as has been suggested in the primate literature (*66*), we propose that the actual choices and therefore activity related to motor responses, occurs downstream of the OFC, likely in striatal regions (*12*).

The notion that medial and lateral OFC are functionally separable in primates also appears to be borne out by data from primate lesion studies. For example, there are several findings suggesting

that primates with lateral OFC damage are, like rodents, particularly impaired on tasks that require altering behaviour when reinforcement contingencies change, such as contingency reversal and goGo/nogo NoGo tasks (64, 65). Medial OFC lesions on the other hand impair devaluation performance in macaques (66) just as they do in rodents (4). Moreover, there have been at least two studies conducted in humans that have specifically provided evidence for cognitive mapping and its dependence on OFC (2, 12). Taken together, therefore, we suggest that at the level of associative learning and the expression of learned associations, there are significant parallels between the primate and rodent OFC.

Whether our account of medial versus lateral OFC function applies across species or not, it is undeniable that any OFC-based representations of an individual's goals or location in task space would be substantially more complex in primates. For example, for rodents and other lower-order (i.e. non-primate) mammals, effective reinforcers are almost exclusively those related to immediate biological requirements for survival (i.e. food, shelter, reproduction), or stimuli that have become associated with those outcomes. In contrast, humans and to some extent other primates, exist in rich and complex social worlds, where types of reinforcement extend well beyond the boundaries of biological requirement to meet higher-order emotional needs. Indeed, it has been argued that emotions are states elicited by goals or reinforcers (*67*) and the role of the human (and primate) OFC in this context is to decode and represent reinforcers and the reward value of stimuli that predict them. It is along these lines that we propose that primate and rodent OFC are most likely to differ; humans are capable of selecting goals that will provide complex emotional reinforcement associated with, for example, altruism, trust, humour, and pride. It remains unclear what role, if any, rodent OFC plays in complex emotion, and along these lines significant differences may emerge.

Heterogeneity within the Orbitofrontal Cortex

Here we have reviewed studies that investigate the distinct functions of medial and lateral subregions of OFC. However it is worth noting that even within different subregions of OFC, subpopulations of neurons have shown a remarkable degree of heterogeneity with regards to the types of functions they subserve. Electrophysiological studies in rats and monkeys have discovered, for example, that a subpopulation of neurons located within the lateral OFC preferentially codes for value (*65*), but others have different lateral OFC populations that respond to outcome identities, outcome location, or even responses, independently of value (*68–70*). Moreover, these differing responses are not always static across the changing conditions of the task, with some OFC neurons observed to respond to value prior to reversal and then stop responding to value after reversal, in a seemingly context (or state) dependent manner (*70*). An even more recent study has managed to optogenetically manipulate distinct populations of cells within lateral OFC that either selectively responded to feeding or social interaction experiences, and found differing effects on both types of behaviour (*71*). Studies such as these complement the notion that the OFC integrates the rich and varied range of information that is necessary to form representations of state, and further suggest that intercellular processes may contribute to such representations.

Finally, we wish to conclude by re-iterating that although we have here proposed different roles for medial and lateral OFC in representing terminal versus initial states, respectively, these roles are likely achieved interdependently rather than independently. For example, although we have argued that medial OFC does not infer initial states directly, medial OFC-dependent retrieval of particular outcomes (terminal states) may sometimes be necessary for the accurate inference of the initial state. A potential example can be drawn from Pavlovian-instrumental transfer, in which the presentation of a stimulus may retrieve a representation of its associated outcome (terminal state), and infer the availability of that outcome in the initial state, promoting selection of the response that will deliver it. An alternative real-world example might be if you visit your regular café under the understanding it is Saturday (initial state) and you therefore infer from your knowledge that the daily special is steak (outcome state). The converse can also true; when you see that the daily special

offered is steak (outcome state), you may infer that today must be Saturday (initial state). In this way, it is possible to imagine many scenarios in which the medial and lateral OFC regions might work in concert to achieve a cohesive cognitive map of task space.

Acknowledgements: We thank Bernard Balleine, Melissa Sharpe, Amir Dezfouli, and Marios Panayi for helpful discussions and comments regarding this manuscript.

This work was supported by the National Health and Medical Research Council of Australia Grant (GNT 1148244) awarded to L.A.B.

References

1. R. C. Wilson, Y. K. Takahashi, G. Schoenbaum, Y. Niv, Orbitofrontal Cortex as a Cognitive Map of Task Space. *Neuron*. **81**, 267–279 (2014).

2. N. W. Schuck, M. B. Cai, R. C. Wilson, Y. Niv, Human Orbitofrontal Cortex Represents a Cognitive Map of State Space. *Neuron*. **91**, 1402–1412 (2016).

3. M. J. Sharpe, T. Stalnaker, N. W. Schuck, S. Killcross, G. Schoenbaum, Y. Niv, An Integrated Model of Action Selection: Distinct Modes of Cortical Control of Striatal Decision Making. *Annu. Rev. Psychol.* **70**, 53–76 (2019).

4. L. A. Bradfield, A. Dezfouli, M. Van Holstein, B. Chieng, B. W. Balleine, Medial Orbitofrontal Cortex Mediates Outcome Retrieval in Partially Observable Task Situations. *Neuron*. **88**, 1268–1280 (2015).

5. A. Izquierdo, Functional heterogeneity within rat orbitofrontal cortex in reward learning and decision making. *J. Neurosci.* **37**, 10529–10540 (2017).

6. L. A. Bradfield, G. Hart, B. W. Balleine, Inferring action-dependent outcome representations depends on anterior but not posterior medial orbitofrontal cortex. *Neurobiol. Learn. Mem.* **155**, 463–473 (2018).

7. M. P. H. Gardner, J. C. Conroy, C. V. Styer, T. Huynh, L. R. Whitaker, G. Schoenbaum, Medial orbitofrontal inactivation does not affect economic choice. *Elife*, e38963 (2018).

8. S. Bouret, B. J. Richmond, Ventromedial and Orbital Prefrontal Neurons Differentially Encode Internally and Externally Driven Motivational Values in Monkeys. *J. Neurosci.* **30**, 8591–8601 (2010).

9. M. P. Noonan, J. Sallet, P. H. Rudebeck, M. J. Buckley, M. F. Rushworth, Does the medial orbitofrontal cortex have a role in social valuation? *Eur. J. Neurosci.* **31**, 2341–2351 (2010).

10. M. Malvaez, C. Shieh, M. D. Murphy, V. Y. Greenfield, K. M. Wassum, Distinct cortical–amygdala projections drive reward value encoding and retrieval. *Nat. Neurosci.* **22**, 762–769 (2019).

11. P. H. Rudebeck, E. A. Murray, Dissociable Effects of Subtotal Lesions within the Macaque Orbital Prefrontal Cortex on Reward-Guided Behavior. *J. Neurosci.* **31**, 10569–10578 (2011).

12. S. Nonomura, K. Nishizawa, Y. Sakai, Y. Kawaguchi, S. Kato, M. Uchigashima, M. Watanabe, K. Yamanaka, K. Enomoto, S. Chiken, H. Sano, S. Soma, J. Yoshida, K. Samejima, M. Ogawa, K. Kobayashi, A. Nambu, Y. Isomura, M. Kimura, Monitoring and Updating of Action Selection for Goal-Directed Behavior through the Striatal Direct and Indirect Pathways. *Neuron*. **99**, 1302-1314.e5 (2018).

13. S. R. Heilbronner, J. Rodriguez-Romaguera, G. J. Quirk, H. J. Groenewegen, S. N. Haber, Circuit-Based Corticostriatal Homologies Between Rat and Primate. *Biol. Psychiatry*. **80**, 509–521 (2016).

14. N. W. Schuck, Y. Niv, Sequential replay of nonspatial task states in the human hippocampus. *Science (80-.).* **364**, eaaw5181 (2019).

15. A. Dickinson, B. Balleine, Motivational control of goal-directed action. *Anim. Learn. Behav.* **22**,

1–18 (1994).

16. H. H. Yin, B. J. Knowlton, B. W. Balleine, Lesions of dorsolateral striatum preserve outcome expectancy but disrupt habit formation in instrumental learning. *Eur. J. Neurosci.* **19**, 181–189 (2004).

17. H. H. Yin, B. J. Knowlton, B. W. Balleine, Blockade of NMDA receptors in the dorsomedial striatum prevents action-outcome learning in instrumental conditioning. *Eur. J. Neurosci.* **22**, 505–512 (2005).

18. H. H. Yin Ostlund, S.B., Knowlton, B.J. and Balleine, B.W., The role of the dorsomedial striatum in instrumental conditioning. *Eur. J. Neurosci.* **22**, 513–23 (2005).

19. N. D. Daw, Y. Niv, P. Dayan, Uncertainty-based competition between prefrontal and dorsolateral striatal systems for behavioral control. *Nat. Neurosci.* **8**, 1704–1711 (2005).

20. A. Dezfouli, B. W. Balleine, Habits, action sequences and reinforcement learning. *Eur. J. Neurosci.* **35**, 1036–1051 (2012).

21. R. L. Reep, J. V. Corwin, V. King, Neuronal connections of orbital cortex in rats: Topography of cortical and thalamic afferents. *Exp. Brain Res.* **111**, 215–232 (1996).

22. A. J. McDonald, F. Mascagni, L. Guo, Projections of the medial and lateral prefrontal cortices to the amygdala: a Phaseolus vulgaris leucoagglutinin study in the rat. *Neuroscience*. **71**, 55–75 (1996).

23. W. B. Hoover, R. P. Vertes, Projections of the medial orbital and ventral orbital cortex in the rat. *J. Comp. Neurol.* **519**, 3766–3801 (2011).

24. S. B. Ostlund, B. W. Balleine, Orbitofrontal Cortex Mediates Outcome Encoding in Pavlovian But Not Instrumental Conditioning. *J. Neurosci.* **27**, 4819–4825 (2007).

25. M. C. Panayi, S. Killcross, Functional heterogeneity within the rodent lateral orbitofrontal cortex dissociates outcome devaluation and reversal learning deficits. *Elife*, e37357 (2018).

26. S. L. Parkes, P. M. Ravassard, J.-C. Cerpa, M. Wolff, G. Ferreira, E. Coutureau, Insular and Ventrolateral Orbitofrontal Cortices Differentially Contribute to Goal-Directed Behavior in Rodents. *Cereb. Cortex*. **28**, 2313–2325 (2018).

27. N. T. Lichtenberg, Z. T. Pennington, S. M. Holley, V. Y. Greenfield, C. Cepeda, M. S. Levine, K. M. Wassum, Basolateral Amygdala to Orbitofrontal Cortex Projections Enable Cue-Triggered Reward Expectations. *J. Neurosci.* **37**, 8374–8384 (2017).

28. C. M. Gremel, R. M. Costa, Orbitofrontal and striatal circuits dynamically encode the shift between goal-directed and habitual actions. *Nat. Commun.* **4**, 2264 (2013).

29. C. M. Gremel, J. H. Chancey, B. K. Atwood, G. Luo, R. Neve, C. Ramakrishnan, K. Deisseroth, D. M. Lovinger, R. M. Costa, Endocannabinoid Modulation of Orbitostriatal Circuits Gates Habit Formation. *Neuron*. **90**, 1312–1324 (2016).

30. S. J. Gershman, D. M. Blei, Y. Niv, Context, Learning, and Extinction. *Psychol. Rev.* **117**, 197–209 (2010).

31. S. J. Gershman, Y. Niv, Exploring a latent cause theory of classical conditioning. *Learn. Behav.* **40**, 255–268 (2012).

32. Y. Chudasama, T. W. Robbins, 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 (2003).

33. G. L. Dalton, N. Y. Wang, A. G. Phillips, S. B. Floresco, Multifaceted Contributions by Different Regions of the Orbitofrontal and Medial Prefrontal Cortex to Probabilistic Reversal Learning. *J. Neurosci.* **36**, 1996–2006 (2016).

34. G. Schoenbaum, S. L. Nugent, M. P. Saddoris, B. Setlow, Orbitofrontal lesions in rats impair reversal but not acquisition of go, no-go odor discriminations. *Neuroreport*. **13**, 885–890 (2002).

35. G. Schoenbaum, B. Setlow, M. P. Saddoris, M. Gallagher, Encoding Predicted Outcome and Acquired Value in Orbitofrontal Cortex during Cue Sampling Depends upon Input from Basolateral Amygdala. *Neuron*. **39**, 855–867 (2003).

36. A. C. Mar, A. L. J. Walker, D. E. Theobald, D. M. Eagle, T. W. Robbins, Dissociable Effects of Lesions to Orbitofrontal Cortex Subregions on Impulsive Choice in the Rat. *J. Neurosci.* **31**, 6398–6404 (2011).

37. A. Münster, W. Hauber, Medial Orbitofrontal Cortex Mediates Effort-related Responding in Rats. *Cereb. Cortex*. **28**, 4379–4389 (2018).

38. S. L. Gourley, A. S. Lee, J. L. Howell, C. Pittenger, J. R. Taylor, Dissociable regulation of instrumental action within mouse prefrontal cortex. *Eur. J. Neurosci.* **32**, 1726–1734 (2010).

39. M. E. Hervig, L. Fiddian, L. Piilgaard, T. Božič, M. Blanco-Pozo, C. Knudsen, S. F. Olesen, J. Alsiö, T. W. Robbins, Dissociable and Paradoxical Roles of Rat Medial and Lateral Orbitofrontal Cortex in Visual Serial Reversal Learning. *Cereb. Cortex*, 1–14 (2019).

40. L. Bradfield, J. Bertran-Gonzalez, B. Chieng, B. W. Balleine, The Thalamostriatal Pathway and Cholinergic Control of Goal-Directed Action: Interlacing New with Existing Learning in the Striatum. *Neuron*. **79**, 153–166 (2013).

41. K. S. Zimmermann, J. A. Yamin, D. G. Rainnie, K. J. Ressler, S. L. Gourley, Connections of the Mouse Orbitofrontal Cortex and Regulation of Goal-Directed Action Selection by Brain-Derived Neurotrophic Factor. *Biol. Psychiatry*. **81**, 366–377 (2017).

42. K. S. Zimmermann, C. Li, D. G. Rainnie, K. J. Ressler, S. L. Gourley, Memory Retention Involves the Ventrolateral Orbitofrontal Cortex: Comparison with the Basolateral Amygdala. *Neuropsychopharmacology*. **43**, 373–383 (2018).

43. A. J. Whyte, H. W. Kietzman, A. M. Swanson, L. M. Butkovich, B. R. Barbee, G. J. Bassell, C. Gross, S. L. Gourley, Reward-Related Expectations Trigger Dendritic Spine Plasticity in the Mouse Ventrolateral Orbitofrontal Cortex. *J. Neurosci.* **39**, 4595–4605 (2019).

44. C. L. Pickens, M. P. Saddoris, B. Setlow, M. Gallagher, P. C. Holland, G. Schoenbaum, Different Roles for Orbitofrontal Cortex and Basolateral Amygdala in a Reinforcer Devaluation Task. *J. Neurosci.* **23**, 11078–11084 (2003).

45. M. Gallagher, R. W. McMahan, G. Schoenbaum, Orbitofrontal Cortex and Representation of Incentive Value in Associative Learning. *J. Neurosci.* **19**, 6610–6614 (1999).

46. A. Izquierdo, E. A. Murray, Combined Unilateral Lesions of the Amygdala and Orbital Prefrontal Cortex Impair Affective Processing in Rhesus Monkeys. *J. Neurophysiol.* **91**, 2023–2039 (2004).

47. Y. K. Takahashi, M. R. Roesch, T. A. Stalnaker, R. Z. Haney, D. J. Calu, A. R. Taylor, K. A. Burke, G. Schoenbaum, The Orbitofrontal Cortex and Ventral Tegmental Area Are Necessary for

Learning from Unexpected Outcomes. *Neuron*. **62**, 269–280 (2009).

48. M. A. McDannald, G. R. Esber, M. A. Wegener, H. M. Wied, T.-L. Liu, T. A. Stalnaker, J. L. Jones, J. Trageser, G. Schoenbaum, Orbitofrontal neurons acquire responses to 'valueless' Pavlovian cues during unblocking. *Elife*. **3** (2014), doi:10.7554/eLife.02653.

49. J. L. Jones, G. R. Esber, M. A. McDannald, A. J. Gruber, A. Hernandez, A. Mirenzi, G. Schoenbaum, Orbitofrontal Cortex Supports Behavior and Learning Using Inferred But Not Cached Values. *Science (80-.).* **338**, 953–956 (2012).

50. N. D. Daw, Model-based reinforcement learning as cognitive search: Neurocomputational theories. *Cogn. Search Evol. Algorithms, Brain* (2012).

51. P. Dayan, K. C. Berridge, Model-based and model-free Pavlovian reward learning: Revaluation, revision, and revelation. *Cogn. Affect. Behav. Neurosci.* **14**, 473–492 (2014).

52. W. M. Pauli, G. Gentile, S. Collette, J. M. Tyszka, J. P. O'Doherty, Evidence for model-based encoding of Pavlovian contingencies in the human brain. *Nat. Commun.* **10**, 1099 (2019).

53. I. P. Pavlov, Conditioned reflexes. An investigation of the physiological activity of the cerebral cortex. Translated and edited by E. V. ANREP. *Cond. reflexes. An Investig. Physiol. Act. Cereb. cortex. Transl. Ed. by E. V. ANREP* (1927).

54. T. W. Robbins, The acquisition of responding with conditioned reinforcement: Effects of pipradrol, methylphenidate, d-amphetamine, and nomifensine. *Psychopharmacology (Berl).* **58**, 79–87 (1978).

55. P. C. Holland, R. A. Rescorla, The effect of two ways of devaluing the unconditioned stimulus after first- and second-order appetitive conditioning. *J. Exp. Psychol. Anim. Behav. Process.* **1**, 355–363 (1975).

56. E. T. Rolls, The orbitofrontal cortex and emotion in health and disease, including depression. *Neuropsychologia*. **128**, 14–43 (2019).

57. M. Laubach, L. M. Amarante, K. Swanson, S. R. White, What, if anything, is rodent prefrontal cortex? *eNeuro* (2018), 5, doi:10.1523/ENEURO.0315-18.2018.

58. J. D. Wallis, Cross-species studies of orbitofrontal cortex and value-based decision-making. *Nat. Neurosci.* **15**, 13–19 (2012).

59. J. L. Price, Definition of the orbital cortex in relation to specific connections with limbic and visceral structures and other cortical regions. *Ann. N. Y. Acad. Sci.* **1121**, 54–71 (2007).

60. R. L. Reep, J. V. Corwin, V. King, Neuronal connections of orbital cortex in rats: topography of cortical and thalamic afferents. *Exp. Brain Res.* **111**, 215–232 (1996).

61. J. O'Doherty, M. L. Kringelbach, E. T. Rolls, J. Hornak, C. Andrews, Abstract reward and punishment representations in the human orbitofrontal cortex. *Nat. Neurosci.* **4**, 95–102 (2001).

62. E. L. Rich, J. D. Wallis, Decoding subjective decisions from orbitofrontal cortex. *Nat. Neurosci.* **19**, 973–980 (2016).

63. C. Padoa-Schioppa, J. A. Assad, Neurons in the orbitofrontal cortex encode economic value. *Nature*. **441**, 223–226 (2006).

64. J. D. Wallis, E. K. Miller, Neuronal activity in primate dorsolateral and orbital prefrontal cortex during performance of a reward preference task. *Eur. J. Neurosci.* **18**, 2069–2081 (2003).

65. L. E. Grattan, P. W. Glimcher, Absence of Spatial Tuning in the Orbitofrontal Cortex. *PLoS One*. **9**, e112750 (2014).

66. B. Lau, P. W. Glimcher, Value Representations in the Primate Striatum during Matching Behavior. *Neuron*. **58**, 451–463 (2008).

67. E. T. Rolls, Emotion and decision-making explained: A précis. *Cortex*. **59**, 185–193 (2014).

68. L. Tremblay, W. Schultz, Relative reward preference in primate orbitofrontal cortex. *Nature*. **398**, 704–708 (1999).

69. M. R. Roesch, A. R. Taylor, G. Schoenbaum, Encoding of Time-Discounted Rewards in Orbitofrontal Cortex Is Independent of Value Representation. *Neuron*. **51**, 509–520 (2006).

70. G. Schoenbaum, A. A. Chiba, M. Gallagher, Neural Encoding in Orbitofrontal Cortex and Basolateral Amygdala during Olfactory Discrimination Learning. *J. Neurosci.* **19**, 1876–1884 (1999).

71. J. H. Jennings, C. K. Kim, J. H. Marshel, M. Raffiee, L. Ye, S. Quirin, S. Pak, C. Ramakrishnan, K. Deisseroth, Interacting neural ensembles in orbitofrontal cortex for social and feeding behaviour. *Nature*. **565**, 645–649 (2019).

Figure Captions

Figure 1: Simplified state space representation of making a cup of tea by the agent according to model-based RL. S1, the initial state is determined by being in the kitchen, taking 'A1' action of switching on the kettle transitions the agent into state S2, determined by the kettle being on. State-to-state transitions are achieved between each state from S1-S7 in this manner by taking actions A1-A6, until the terminal state 'S7' determined by the tea being drunk (goal), is reached. S – state, A – action.

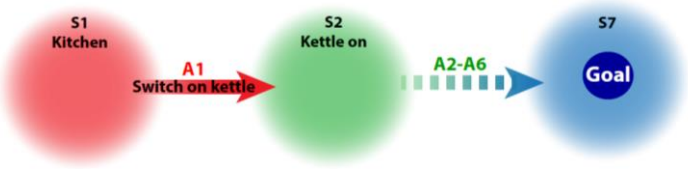

Figure 2. State space and data representations for the role of medial versus lateral OFC in instrumental outcome devaluation tasks. A, State space representation of a two-lever two-outcome instrumental outcome devaluation task. Rats in the initial state (S1) can choose to make one of two actions (A1 or A2) which will transition them into one of two terminal states (O1 and O2). **B,** When tested under extinction following outcome devaluation, intact goal-directed rats and rats with lateral OFC disruption preferentially choose the action that had given the still-valued outcome (A1, blue), whereas rats with mOFC disruption choose A1 and A2 equally[4-5, 24-26]. **C,** State space representation of a single lever contextual discrimination task; rats need to use state information (S1 or S2) to determine the reinforcement schedule that delivers O1. **D,** When tested under extinction in the goal-directed context, intact rats show outcome devaluation but rats with lOFC disruption don't[28-29]. S – State, A – Action, O – Outcome, mOFC – medial orbitofrontal cortex, lOFC – lateral orbitofrontal cortex.

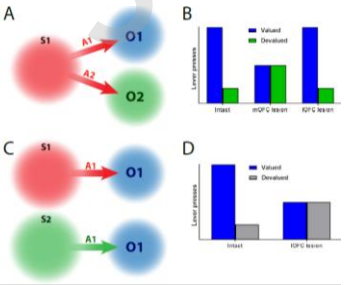

Figure 3. State space and data representation for single-outcome contingency reversal.

A, Rats in the initial state (S1) are trained such that one action is reinforced (A1-O) and the other isn't (A2-no O). **B,** When these contingencies are reversed (A1-no O, A2-O) rats form a new state (S2) within which to encode the new learning. **C,** Following reversal, intact rats rapidly bias responding to the newly reinforced action (accurate responses) and cease responding on the non-reinforced action (perseverative errors), whereas rats with mOFC disruption show increased perseverative errors, and rats with lOFC disruption show both increased perseverative errors and decreased accurate responses[1,32,35]. S – State, A – Action, O – Outcome.

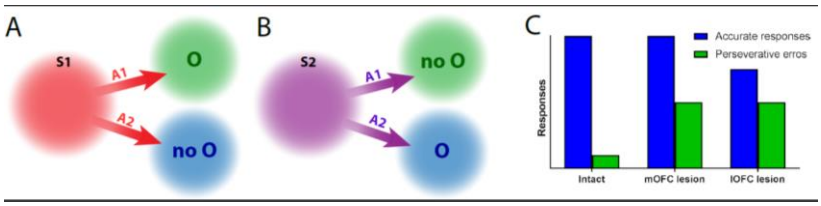

Table 1. Summary of behavioral findings with regards to orbitofrontal cortex subregions.

OFC – Orbitofrontal cortex.

Task	Structure inactivated/ lesioned	Finding	Reference
Instrumental Devaluation	Anterior medial OFC	Impaired	Bradfield et al., (2015; 2018)
Instrumental Devaluation	Posterior medial OFC	Intact	Munster & Hauber (2017); Bradfield et al., (2018).
Instrumental Devaluation	Lateral OFC	Intact	Ostlund & Balleine, (2007); Lichtenberg et al., (2017); Panayi & Killcross, 2018a; Parkes et al., (2018).
Instrumental Devaluation	Lateral OFC	Impaired	Gremel & Costa (2013; 2016)
Contingency reversal (value-based)	Lateral OFC	Impaired (slower reversal)	Chudsama & Robbins (2003); Dalton et al., 2016; Schoenbaum et al. (2003); Wilson et al. (2014)
Contingency reversal (identity-based)	Lateral OFC	Impaired	Parkes et al. (2018)
Contingency reversal (value-based)	Medial OFC	Higher perseveration during reversal	Gourley et al. (2010)
Contingency reversal (probabilistic task)	lateral OFC	Impairment	Dalton et al., (2016)
Contingency reversal (delay discounting)	Medial OFC	Facilitation	Mar et al. (2011)
Contingency Degradation	Lateral OFC	Impaired	Zimmerman et al., (2017)
Contingency Degradation	Lateral OFC	Encoding intact, sustained performance impaired	Zimmerman et al, (2018)
Contingency Degradation	Medial OFC	Intact	Bradfield et al. (2015)
Pavlovian instrumental transfer	Lateral OFC	Impaired	Ostlund & Balleine, (2007)
Pavlovian instrumental transfer	Medial OFC	Impaired	Bradfield et al., (2015; 2018)

Table 2. Summary of relevant behavioural paradigms

A – action, O – outcome, S – stimulus.

Paradigm	Stage 1	Stage 2	Test
Instrumental Devaluation	A1-O1 A2-O2	O1: 1hr	A1 < A2
Contingency Reversal (valued-based)	A1-O1 A2-	A1- A2-O1	
Contingency Reversal (identity-based)	A1-O1 A2-O2	A1-O2 A2-O1	O1: 1hr Original: A1 < A2 Reversed: A2 < A1
Contingency Degradation	A1-O1 A2-O2	A1-O1/O1 A2-O2	A1 < A2
Pavlovian-instrumental transfer	S1-O1 S2-O2	A1-O1 A2-O2	S1: A1 > A2 S2: A2 > A1