Does the orbitofrontal cortex signal value?

Geoffrey Schoenbaum[1], **Yuji Takahashi**[1], **Tzu-Lan Liu**[2], and **Michael A. McDannald**[1]

[1]Department of Anatomy and Neurobiology, University of Maryland School of Medicine, Baltimore, Maryland [2]Department of Psychology, National Taiwan University, Taipei City, Taiwan

Abstract

The orbitofrontal cortex (OFC) has long been implicated in associative learning. Early work by Mishkin and Rolls showed that the OFC was critical for rapid changes in learned behavior, a role that was reflected in the encoding of associative information by orbitofrontal neurons. Over the years, new data—particularly neurophysiological data—have increasingly emphasized the OFC in signaling actual value. These signals have been reported to vary according to internal preferences and judgments and to even be completely independent of the sensory qualities of predictive cues, the actual rewards, and the responses required to obtain them. At the same time, increasingly sophisticated behavioral studies have shown that the OFC is often unnecessary for simple value-based behavior and instead seems critical when information about specific outcomes must be used to guide behavior and learning. Here, we review these data and suggest a theory that potentially reconciles these two ideas, value versus specific outcomes, and bodies of work on the OFC.

Keywords

orbital frontal cortex; overexpectation; reward; value-guided behavior

The orbitofrontal cortex, associative learning, and value-guided behavior

The first evidence implicating the orbitofrontal cortex (OFC) in associative learning clearly came from studies of reversal learning. *Reversal learning* refers to a variety of behavioral tasks in which an animal is first trained that something is good and predicts reward, while another thing is bad and predicts either nonreward or punishment. After the animal has learned to respond appropriately, the predictive associations are reversed, and the animal must switch or reverse their behavior. In a seminal report in 1972, Jones and Mishkin showed that monkeys with damage to the OFC were impaired in this behavioral function.[1] Although this was not the first study showing a role for OFC in associative learning, it stands out in retrospect because it also showed that this deficit was not due to impairments in initial learning and behavior but rather reflected a specific inability to change behavior after reversal. This basic result has since been replicated across an astonishing variety of species and reversal paradigms.[1–12] While the basis of this impairment remains a topic of debate,[13] it provides strong evidence that the OFC is critical for associative learning.

Accordingly, subsequent unit recording and functional magnetic resonance imaging (FMRI) studies have shown with increasing detail how the OFC represents associative information,

---

Address for correspondence: Geoffrey Schoenbaum, Department of Anatomy and Neurobiology, University of Maryland School of Medicine, 20 Penn St. HSF-2 S251, Baltimore, MD 21201. gscho002@umaryland.edu.

Conflicts of interest The authors declare no conflicts of interest.

particularly information about the value of expected outcomes. This began arguably with work by Thorpe and Rolls, recording from monkeys engaged in a reversal task.[14] They found that orbitofrontal neurons fired to the sight of the syringes used to deliver the juice rewards to the monkeys, and that this activity often declined or even reversed when the syringes were spiked with saline. Building on this report, numerous studies have since shown that neural activity in the OFC increases to cues and after responses that predict rewards, and that this activity can be remarkably indifferent to the actual features of these cues or even the outcomes they predict.[15–52] For example, in both rats and monkeys trained in a discrimination task involving multiple cues, orbitofrontal neurons tended to fire similarly to cues that predicted the same reward.[15,18,19] Similarly, orbitofrontal neurons have also been shown to change firing before delivery of a particular outcome depending on the outcome's current value[17] or even that of other outcomes that are available.[23,40,45,47] Similarly, the blood oxygen level–dependent (BOLD) signal in the OFC is also sensitive to the value of an expected outcome.[24,29] These data and other studies showing that orbitofrontal neurons can integrate information about probability, size, and time to reward[32,53,54] have been interpreted as showing that the specific features of the cues, responses, or outcomes defining the associations are not what is represented in the OFC; rather, the argument goes that what is being critically represented is the expected value or utility, and that the OFC might be the site where different items are converted into a common currency for the purposes of goods-based decision making.[55] Consistent with this view, there are neurons in monkey OFC whose activity reflects the value of available outcomes or goods, irrespective of a number of variables including the cue properties and location, response executed to obtain the reward, and the reward properties.[34]

Bolstering this view, behavioral work has shown increasingly that the OFC is important for value-guided behaviors. The most classic example comes from studies using reinforcer devaluation. In these tasks, animals are trained to respond to cues to obtain rewards. After the predictive relationships have been learned, the rewards are directly "devalued" by overfeeding, pairing with illness, or other means. In subsequent probe tests, monkeys and rats exhibit reduced responding to the cues predictive of the devalued rewards. Rats and monkeys with orbitofrontal lesions fail to show any effect of reinforcer devaluation on learned behavior, responding to cues signaling valued and devalued rewards alike.[8,56,57] This suggests that the OFC is critical when value must be used to guide a response. Interestingly, this deficit is observed despite apparently normal initial learning and correct selection of valued versus devalued rewards when the rewards themselves are present. We will return to these observations a bit later. As mentioned above, the importance of the OFC to value-based behavior might be suspected from reversal learning impairments in animals with OFC lesions. Yet reversal impairments are difficult to interpret. The tasks are typically complex, involving Pavlovian and instrumental contingencies and, in many cases, probability judgments, choice, and even appetitive and aversive learning. In addition, reversal tasks potentially confound impaired learning (i.e., the acquisition of the new information or extinction of the old) and performance (i.e., the use of the new information), since expectations are violated at the same time that animals are asked to use the new information to guide behavior. We have recently used a Pavlovian overexpectation task to provide a more direct test of the role of the OFC in *learning* (Fig. 1). In the first stage of overexpectation (Fig. 1A), several cues are separately trained to predict three sugar pellets. In the second stage (Fig. 1B), two of these cues are presented in compound, still followed by three sugar pellets. Subsequent probe testing (Fig. 1C), in which the cues are presented separately and without reward, typically reveals an immediate reduction in conditioned responding. The sudden decline in responding reflects learning that occurred during the compound training, learning thought to result from the receipt of less reward (three pellets) than would be expected based on the sum of the separate cues (six pellets). Bilateral inactivation of the OFC in rats during the compound training prevents the normal reduction

in conditioned responding during the later probe test (Fig. 1C), suggesting that the OFC is essential for *learning*.[58] Notably, summation of responding to the compound, which is evidence that the rat combines the predictive value of the cues in compound training, is also abolished by orbitofrontal inactivation. This suggests that the critical role the OFC plays is related to predicting the increased amount of reward in this setting. Interestingly, orbitofrontal inactivation in these same rats had no effect on extinction learning when reward was simply omitted,[59] a finding we also consider later.

The OFC is critical for outcome-guided, not value-guided, behavior

At first glance, the evidence presented above seems to provide overwhelming support for the idea that the OFC signals value, which can be used by the animal to either guide behavior or support new learning in the face of changing outcomes. Other findings suggest a closer look might be warranted. First, while neural correlates can be found to support the strong version of this hypothesis, these typically exist amidst a much more diverse population of neural correlates. Their presence may support the argument, but they are not necessarily representative. Further, these value correlates are by no means ubiquitous across studies; in many studies, including our own, activity in individual orbitofrontal neurons can be highly attuned to the specific features of cues (identity, match-nonmatch comparison, and even sequence),[15,21,25,41] outcomes (identity, size, location),[21,34] and even responses (left or right)[33,35,38,60] that are critical in defining the relevant associations. Second, although reversal learning, devaluation, and overexpectation studies described above seem to implicate the OFC in using information about value, in each case the involvement of this region was qualified. The OFC was necessary for rapid reversals but not initial discrimination learning, for changes in responding after devaluation but not simple conditioning, and for extinction by overexpectation but not by reward omission. Indeed, the OFC is not required for choosing between large and small rewards, and it also appears to be unnecessary for showing consistent preferences between different rewards (though perhaps it is for transitivity of preference—see Fellows, this volume[90]).

One solution to this conundrum is found in the concept of value. Although economic theory treats value as unitary, computational and learning theory accounts more directly relevant to brain function admit to multiple sources of value. In computational accounts, a distinction is drawn between model-free value, which is simply a unit value attached to or "cached" in cues or environmental states by virtue of prior training, versus model-based value, which is derived from knowledge of the structure of the task or environment.[61,62] Similarly, learning theory distinguishes between general value, or affective properties acquired by cues through repeated pairing with rewards, and a more specific value, which is accessed through a representation of a specific outcome and thus sensitive to that outcome's current desirability.[63–65] Thus, model-free value can be linked to general-affective value, and model-based value can be linked to specific-outcome value. Importantly, learning theory has operationalized these distinctions in a variety of tasks such as Pavlovian-to-instrumental transfer, conditioned reinforcement, unblocking, and it has been shown that different brain circuits appear to respect these distinctions.

Emerging evidence using these tasks suggests that the OFC may be critical to value-guided behavior when the values must be derived from model-based, outcome-specific representations and not when values derived from cached or general affective information are sufficient. For example, Ostlund and Balleine (Fig. 2) have shown that the OFC is necessary for Pavlovian-to-instrumental transfer.[66] *Pavlovian–instrumental transfer* refers to the increase in instrumental responding that occurs when a separately trained Pavlovian cue is presented with the instrumental response. The transfer effect—or the increased responding —is thought to reflect at least two processes: one being the general motivating properties of

the cue and the other being the ability of the cue to directly access a representation of the specific outcome. If there is only a single response—say a lever press—and one outcome—say a sucrose pellet—then these processes are indistinguishable; however, if two levers are available, one leading to the same outcome as the cue and a second leading to a different outcome, then they can be dissociated, since the general motivating properties will increase responding on both levers, and the outcome-specific mechanism only on the lever that leads to the same outcome. Remarkably, Balleine *et al.*[62] found that orbitofrontal lesions affected only the outcome-specific transfer effect (Fig. 2).

A second example comes from studies of conditioned reinforcement. Conditioned reinforcement refers to the ability of animals to acquire a new instrumental response to trigger presentation of Pavlovian cues that have been previously paired with rewards. Like transfer, this phenomenon reflects both the general value that the cue has acquired as well as its ability to trigger a representation of the specific reward and its current value.[67,68] In most settings, the influence of these cue-evoked properties on instrumental responding is confounded; indeed, conditioned reinforcement is typically reported to be insensitive to reinforcer devaluation,[67,68] suggesting that general affective properties are sufficient for entirely normal responding. However, it is possible to train the Pavlovian cue so that the general affective properties are blocked. This results in a cue that supports conditioned reinforcement primarily via its access to a representation of the reward and its current value. We have found that conditioned reinforcement tested in this manner is particularly sensitive to orbitofrontal lesions (Fig. 3), whereas conditioned reinforcement supported by a normally trained cue is not.[69,70]

A third example comes from studies of differential outcome expectancy. In these tasks, animals are presented with separate cues, which instruct the animals that one of two separate actions will produce a rewarding outcome. Learning these separate cue-action-outcome associations can be difficult. It has been demonstrated that when the outcomes produced by the two cue-action chains differ, discrimination proceeds significantly faster—that is, expecting different rewards can speed discrimination. As noted earlier, orbitofrontal damage typically has no effect on simple discrimination learning. This is true when the discrimination is a go/no-go task, and it is also true when the discrimination involves learning to select different responses to obtain reward. However, orbitofrontal-lesioned rats fail to show enhanced acquisition when different outcomes are employed, performing no better than when the outcomes are identical.[71] A modified version of the differential outcome expectancy effect also demonstrates that normal rats are more likely to make errors that share a common reinforcer than a different reinforcer. The outline of this procedure is shown in Figure 4A. Animals must learn three cue–action–outcome associations: cue 1 → action 1 → outcome 1; cue 2 → action 2 → outcome 2; and cue 3 → action 3 → outcome 1 (50% of trials) or outcome 2 (50% of trials). This discrimination is also difficult. Over the course of learning, many errors in action selection occur during the three cues. However, during cue 1, normal rats are significantly more likely to perform action 3 (leading to a shared outcome with action 1) than action 2 (leading to a different outcome than action 1; Fig. 4B). This pattern could only be observed if animals were using outcome expectancies to guide behavior. During cue 1, animals with OFC lesions are equally likely to perform action 3 and action 2, demonstrating that these rats are not using outcome expectancies to guide behavior.[72] Again, the OFC appears to be uniquely engaged when animals must use expectations of specific features of outcomes to guide discrimination learning.

Notably, this same distinction between behaviors that can be guided by general value versus those dependent on knowledge of the outcome may also hold to explain the role of the OFC in devaluation. As discussed earlier, the OFC is necessary for the normal decline in conditioned responding that occurs after devaluation of the predicted reward.[8,56,57] In an

apparent paradox, this effect occurs in the face of otherwise normal conditioning and the successful formation of conditioned taste aversions. Thus rats and monkeys seem to understand and respond appropriately to the cues, even using them to select among different rewards, yet they seem to be unable to change their behavior when the value of the reward changes. This is difficult to explain if the OFC is critical to value judgments generally; however, it would be expected if the OFC were required for accessing value through representations of the outcome. Consistent with this idea, the OFC plays a particularly critical role in the final test phase of reinforcer devaluation, when previously acquired information about the association between the cue and the food and the food and illness or satiety must be integrated to appropriately guide the response.[73,74]

Finally, the OFC's role in learning also appears to be restricted to situations in which expected value must be derived through a representation of the specific outcome. This can be shown using unblocking, in which learning driven by general value can be isolated from learning driven by specific outcome value. In a recent unblocking experiment, rats were trained that different cues predicted different amounts of differently flavored but equally palatable outcomes. Subsequently, these cues were presented in compound with new cues. For some compounds, a greater than expected amount of reward was delivered. This manipulation was meant to unblock learning by increasing the value of the reward without changing its sensory features. For other compounds, a differently flavored but equally palatable reward was delivered. This manipulation was meant to unblock learning by changing the identity of the reward without changing its general value. Accordingly, normal rats exhibited increased responding to the added cues that signaled either *more* or *different* rewards. Orbitofrontal-lesioned rats exhibited increased responding only to the cue that was followed by the larger reward (Fig. 5A); they did not appear to have learned in response to a change in reward identity (Fig. 5B).[75] Importantly, this deficit was observed despite entirely normal learning and normal responding in the compound training phase. Indeed, orbitofrontal-lesioned rats even exhibited similar differences in conditioned responding during initial training for cues predictive of large versus small rewards. Thus their ability to learn about differently valued rewards seemed unaffected.

Interestingly, although there were small differences in rats' responding to the cues predictive of the two rewards (or in preferences to the actual rewards), these differences failed to predict learning in response to a reward shift in controls (or lesioned rats; Fig. 5C). Thus learning in normal rats in response to the use of a different-flavored reward in the compound phase cannot be explained based on small differences in general value attributed to the cues predictive of the different rewards. Rather, it must be based on a recognition of the change in the identity of the expected outcome. The deficit observed in lesioned rats therefore is consistent with the idea that the OFC is critical for signaling specific information about outcomes, including perhaps specific value. Lacking this representation, orbitofrontal-lesioned rats failed to show identity unblocking, since the rats were essentially blind to these features.

Importantly, this perspective is not necessarily at odds with the role that the OFC plays in summation and the resultant extinction learning in the Pavlovian overexpectation task. As noted above, the OFC is necessary for summation—increased responding in response to the simultaneous or compound presentation of the two previously trained cues—and the learning that results from this effect.[58] Although this task does not formally involve the presentation of different outcomes, it is reasonable to assume that on a neural basis, summation reflects the extent to which the cues evoke dissociable or distinct representations of the expected outcomes. Consistent with this, explicitly different outcomes evoke larger summation effects than similar outcomes.[76] Thus the effect that depends critically on the OFC—summation—appears to rely on the ability to hold in mind (and add together) specific

information that differs between different outcomes, as opposed to general affective properties that would be the same. One might speculate that even with a single outcome, neural representations of that outcome in OFC differ in the context of one cue versus another. Thus the network representing impending sugar during a light would differ somewhat from that representing the same sugar during a tone. Having such nonoverlapping representations would allow the overall neural activity to be higher, either on a unit or at least a population basis, if the tone and light were later presented together. As it turns out, this is precisely what is observed if one records in the OFC during compound training in the overexpectation task.[77] As noted earlier, while inactivation of the OFC abolishes extinction driven by summation and overexpectation, inactivation in the same rats had no effect on extinction learning when reward was simply omitted.[59] Extinction does not require summation and thus can be accomplished by relying only on cached values. Data such as these suggest that the OFC will be involved in learning only when it is necessary to resort to model-based representations to recognize errors.

Viewing OFC function as constructing or implementing a model-based representation has so far accounted for a number of previous findings. A stronger test of this hypothesis would be to make predictions about other, untested behavior. Two procedures that would provide a strong test of our hypothesis are second-order conditioning and sensory preconditioning. In second-order conditioning, a primary cue is first trained to predict reward delivery. As a result of pairing with food, the primary cue will rapidly acquire conditioned responding. Next, a novel, secondary cue is paired with presentation to the primary cue alone. When this is done, the primary cue functions as a reward, and the secondary cue will come to elicit conditioned responding. Though both will drive behavior, primary and secondary cues differ in their sensitivities to the current value of the reward: if the reward is devalued, responding to the primary cue, but not the secondary cue, is reduced.[78] This suggests that value acquired by or cached in the primary cue is sufficient to support learning during second-order conditioning. The OFC should not be necessary for this. However, if the primary cue were trained so as to favor the formation of outcome-specific representations, as we have done to test the role of OFC in conditioned reinforcement,[69] the OFC should then be necessary for second-order conditioning to occur. The prediction that OFC is not required for the development of Pavlovian second-order responding may appear to be at odds with previous studies of its role in second-order schedules of reinforcement. However, the ability of a primary cue to bridge caps in reinforcement[79] or to prevent the extinction of instrumental response[80] may potentially differ from its ability to support new learning. Moreover, somewhat different neural circuits may be required for cues to support new Pavlovian[81] versus new instrumental learning.[82] Pavlovian second-order conditioning, as described above, solely assesses the ability of a primary cue to support new Pavlovian learning about a novel cue. Beyond this, it is also worth noting that in one of these studies,[79] the OFC-lesioned animals actually responded more rather than less on the second-order schedule, though their behavior was insensitive to omission of the conditioned stimulus. Although the overresponding was attributed to a role for OFC in response inhibition, it would also be consistent with the idea that these animals were still capable of attaching value to the second-cue by virtue of its repeated pairing with either the primary or the conditioned reinforcer.

The other, contrasting example we would highlight that would provide a test of this hypothesis would involve sensory preconditioning. In sensory preconditioning, animals are first trained that a neutral cue A predicts neutral cue B: A → B. In this phase, neither A nor B will elicit any food-related behavior. Next, cue B is directly paired with a food reward: B → reward. Finally, responding to cue A alone is tested. When this is done, normal rats show food-related behavior to cue A. This occurs even though cue A has never been paired directly with reward. Successful demonstration of sensory preconditioning requires

integration of two associations: A → B and B → reward. Such integration clearly requires a model-based representation. Unlike second-order conditioning, we expect the OFC will be necessary for the integrating function in sensory preconditioning. Finding dissociable contributions of the OFC to second-order conditioning and sensory preconditioning would greatly strengthen our hypothesis of its specific role in forming and using model-based representations to guide behavior.

Outcome-guided behavior and neural correlates of value

Overexpectation data notwithstanding, a reasonable question at this point is how one can reconcile the apparent critical involvement of OFC in outcome-guided as opposed to value-guided behavior with the wide-scale reporting of what seem to be simple value correlates in orbitofrontal neurons. Of course, the first answer to this criticism is to note that neural correlates are just that—correlates. Lesion and inactivation studies, such as those described above, directly assess whether those correlates are critical to supporting behavior, and to the best of our knowledge, there is no direct evidence to suggest that is the case. Another relatively straightforward answer is to propose that OFC performs multiple, diverse functions (either integrated or in different subregions[83]).

However, it is worth noting that in most of these single unit reports, the neural populations that are highlighted to support the general value coding hypothesis are only part of a much larger ensemble, the characteristics of which typically conform more closely to the idea that the OFC signals information about expected outcomes from which a value might be derived than with the idea that it directly signals general, cached, or economic value. For example, in initial and subsequent reports that orbitofrontal neurons reverse encoding during reversal learning,[14,18,19,22,84] the majority of the neurons failed to reverse and instead seemed to represent specific conjunctions between cues and rewards or punishments. Such activity seems more reminiscent of rules[84] or models of the task than pure value correlates. Likewise, while orbitofrontal neurons do fire similarly to cues that predict similar rewards, they also discriminate between different cues and even sequences of cues during different parts of trials, as though reflecting very specific information about the ongoing events.[15] Finally, even when pure economic value encoding is apparent, this correlate is part of a much larger population of neurons in which the firing activity does appear to be tied to the cues or responses or even the specific features of the expected outcomes.[34] Thus while some neurons may be found that represent economic value, it is not clear whether they define the specific function of the OFC or are something that emerges with sufficient training from a more general function of representing information relevant to defining the specific outcome and the circumstances surrounding (or state defining) its presentation.

By contrast, we have found that neurons in rat OFC generally do not integrate information to signal something approaching general or economic value. In tasks in which reward size is manipulated, we find that while many orbitofrontal neurons fire differentially in anticipation of different size rewards, the firing of these neurons is also typically dependent on other features of the trial, such as the preceding odor cue or the response necessary to obtain the reward.[35] Furthermore, there are typically equal representations of big and small rewards, rather than a dominant pattern related to encoding of size across the population, and neurons that fire based on size typically do not also fire based on delay of reward. Instead, delay is represented in different neurons. We have recently found a similar phenomenon in a task in which we independently manipulated reward size and identity; again, orbitofrontal neurons appear to create equivalent and largely independent representations of large and small reward and of each available identity.[85] Thus while orbitofrontal neurons seem to reflect value-relevant attributes of expected outcomes, they do so on a descriptive rather than an integrative fashion. Notably, this is different from what we have found downstream in

ventral striatum and midbrain; there, neurons active during predictive cues appear to integrate variables about value more readily.[86,87]

Of course, such a representation could obviously be used to derive values for use in comparisons among items. In some ways, the representation in OFC in a particular setting might be thought of as a goods space. However, the relevant variable would not be value on some linear or transitive scale. Rather, the space would be defined by the salient features of the outcomes in a particular setting. While it might be difficult to predict what features would be deemed salient by the animal, the experimenter could control this somewhat by what features are manipulated or made deterministic of reward. In this context, one would expect to see response-based correlates in tasks such as ours and others in which the response is held constant[33,35,38] and not in tasks typically run in primates where the response is made irrelevant.[32,34,88] Likewise with experience in a task, variables such as the cues preceding a reward or even the reward itself might become unimportant for some neurons if the game is to make rapid choices between different amounts of reward. This would yield neurons that would encode a relatively pure economic signal. However, we would not expect this to occur in most neurons, as in fact it does not, or in animals that receive little training or training, in which these features are not extraneous.

The OFC encodes states relevant to reward

According to this scheme, the OFC represents information about the specific features of expected outcomes. This representation may even include the current value of the specific outcome being represented; additionally, more general value could be derived in part from accessing this space. However, it does not seem to signal cached or general values, at least not in a way that is evident in orbitofrontal-dependent behaviors. As we have noted, a specific outcome representation is reminiscent of the concept behind model-based reinforcement learning, in which value is derived from knowledge of the states that define a task and the transition functions that link them together. Thus the OFC would be representing part of this state-based framework. In this regard, it is notable that the neural activity in the OFC seems to be broadly tuned to represent information about reward relevant events. Actual reward presentation is the most salient such event, and as a result many studies focus on firing during events immediately preceding reward. However, orbitofrontal neurons fire to nearly all of the events that define experimental trials. For example, in rats performing a simple odor discrimination task, phasic firing is seen to trial onset, approach to the odor port, odor sampling, odor port exit, approach to the reward well, during waiting in the well, during reward consumption, and to well exit.[15] In each case, this activity tends to anticipate the event instead of being triggered by that event, suggesting that the anticipatory activity before reward may simply be a special case of a more general predictive function. In addition, activity in these different periods often is influenced by features of the event relevant to reward.

The proposal that OFC signals reward-relevant states that define the task would be consistent with the results described above, including the role of the OFC in devaluation, overexpectation, transfer, unblocking, conditioned reinforcement, as well as a host of other less well-controlled tasks. It would also explain why the OFC is not necessary for simple discrimination learning or Pavlovian conditioning or extinction, since these can be accomplished without reference to a model-based representation. Importantly, this reconceptualization of the role of OFC in learning is also consistent with recent evidence from a three-choice reversal task showing that orbitofrontal-dependent deficits reflect an inability to accurately assign credit for errors to the cue that was selected.[83,89] In one of these studies, a careful analysis of the choice pattern of the animals showed that orbitofrontal-lesioned monkeys still recognized errors in reward prediction and changed

behavior as a result. However, they were unable to use their own choices to constrain the spread of effect of the errors. This result shows that orbitofrontal lesions leave intact some error signaling but that the residual capacity is not well constrained by information about what actions were recently taken. Information about actions and the contingencies between specific actions and outcomes is inherently part of a model-based system.

Acknowledgments

References

1. Jones B, Mishkin M. Limbic lesions and the problem of stimulus-reinforcement associations. Exp. Neurol. 1972; 36:362–377. [PubMed: 4626489]

2. Butter CM. Perseveration and extinction in discrimination reversal tasks following selective frontal ablations in Macaca mulatta. Physiol. Behav. 1969; 4:163–171.

3. Rolls ET, et al. Emotion-related learning in patients with social and emotional changes associated with frontal lobe damage. J. Neurol. Neurosurg. Psychiatr. 1994; 57:1518–1524. [PubMed: 7798983]

4. Meunier M, Bachevalier J, Mishkin M. Effects of orbital frontal and anterior cingulate lesions on object and spatial memory in rhesus monkeys. Neuropsychologia. 1997; 35:999–1015. [PubMed: 9226661]

5. McAlonan K, Brown VJ. Orbital prefrontal cortex mediates reversal learning and not attentional set shifting in the rat. Behav. Brain Res. 2003; 146:97–130. [PubMed: 14643463]

6. Bissonette GB, et al. Double dissociation of the effects of medial and orbital prefrontal cortical lesions on attentional and affective shifts in mice. J. Neurosci. 2008; 28:11124–11130. [PubMed: 18971455]

7. Bechara A, et al. Deciding advantageously before knowing the advantageous strategy. Science. 1997; 275:1293–1294. [PubMed: 9036851]

8. Izquierdo AD, Suda RK, Murray EA. Bilateral orbital prefrontal cortex lesions in rhesus monkeys disrupt choices guided by both reward value and reward contingency. J. Neurosci. 2004; 24:7540–7548. [PubMed: 15329401]

9. Chudasama Y, Robbins TW. 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. 2003; 23:8771–8780. [PubMed: 14507977]

10. Fellows LK, Farah MJ. Ventromedial frontal cortex mediates affective shifting in humans: evidence from a reversal learning paradigm. Brain. 2003; 126:1830–1837. [PubMed: 12821528]

11. Hornak J, et al. Reward-related reversal learning after surgical excisions in orbito-frontal or dorsolateral prefrontal cortex in humans. J. Cogn. Neurosci. 2004; 16:463–478. [PubMed: 15072681]

12. Schoenbaum G, et al. Lesions of orbitofrontal cortex and basolateral amygdala complex disrupt acquisition of odor-guided discriminations and reversals. Learn. Mem. 2003; 10:129–140. [PubMed: 12663751]

13. Schoenbaum G, et al. A new perspective on the role of the orbitofrontal cortex in adaptive behaviour. Nat. Rev. Neurosci. 2009; 10:885–892. [PubMed: 19904278]

14. Thorpe SJ, Rolls ET, Maddison S. The orbitofrontal cortex: neuronal activity in the behaving monkey. Exp. Brain Res. 1983; 49:93–115. [PubMed: 6861938]

15. Schoenbaum G, Eichenbaum H. Information coding in the rodent prefrontal cortex. I. Single-neuron activity in orbitofrontal cortex compared with that in pyriform cortex. J. Neurophysiol. 1995; 74:733–750. [PubMed: 7472378]

16. Schoenbaum G, Eichenbaum H. Information coding in the rodent prefrontal cortex. II. Ensemble activity in orbitofrontal cortex. J. Neurophysiol. 1995; 74:751–762. [PubMed: 7472379]

17. Critchley HD, Rolls ET. Hunger and satiety modify the responses of olfactory and visual neurons in the primate orbitofrontal cortex. J. Neurophysiol. 1996; 75:1673–1686. [PubMed: 8727405]

18. Critchley HD, Rolls ET. Olfactory neuronal responses in the primate orbitofrontal cortex: analysis in an olfactory discrimination task. J. Neurophysiol. 1996; 75:1659–1672. [PubMed: 8727404]

19. Rolls ET, et al. Orbitofrontal cortex neurons: role in olfactory and visual association learning. J. Neurophysiol. 1996; 75:1970–1981. [PubMed: 8734596]

20. Schoenbaum G, Chiba AA, Gallagher M. Orbitofrontal cortex and basolateral amygdala encode expected outcomes during learning. Nat. Neurosci. 1998; 1:155–159. [PubMed: 10195132]

21. Lipton PA, Alvarez P, Eichenbaum H. Crossmodal associative memory representations in rodent orbitofrontal cortex. Neuron. 1999; 22:349–359. [PubMed: 10069340]

22. Schoenbaum G, Chiba AA, Gallagher M. Neural encoding in orbitofrontal cortex and basolateral amygdala during olfactory discrimination learning. J. Neurosci. 1999; 19:1876–1884. [PubMed: 10024371]

23. Tremblay L, Schultz W. Relative reward preference in primate orbitofrontal cortex. Nature. 1999; 398:704–708. [PubMed: 10227292]

24. O'Doherty J, et al. Sensory-specific satiety-related olfactory activation of the human orbitofrontal cortex. Neuroreport. 2000; 11:893–897. [PubMed: 10757540]

25. Ramus SJ, Eichenbaum H. Neural correlates of olfactory recognition memory in the rat orbitofrontal cortex. J. Neurosci. 2000; 20:8199–8208. [PubMed: 11050143]

26. Tremblay L, Schultz W. Modifications of reward expectation-related neuronal activity during learning in primate orbitofrontal cortex. J. Neurophysiol. 2000; 83:1877–1885. [PubMed: 10758099]

27. O'Doherty J, et al. Abstract reward and punishment representations in the human orbitofrontal cortex. Nat. Neurosci. 2001; 4:95–102. [PubMed: 11135651]

28. Alvarez P, Eichenbaum H. Representations of odors in the rat orbitofrontal cortex change during and after learning. Behav. Neurosci. 2002; 116:421–433. [PubMed: 12049323]

29. Gottfried JA, O'Doherty J, Dolan RJ. Encoding predictive reward value in human amygdala and orbitofrontal cortex. Science. 2003; 301:1104–1107. [PubMed: 12934011]

30. O'Doherty J, et al. Temporal difference learning model accounts for responses in human ventral striatum and orbitofrontal cortex during Pavlovian appetitive learning. Neuron. 2003; 38:329–337. [PubMed: 12718865]

31. Hikosaka K, Watanabe M. Long- and short-range reward expectancy in the primate orbitofrontal cortex. Eur. J. Neurosci. 2004; 19:1046–1054. [PubMed: 15009152]

32. Roesch MR, Olson CR. Neuronal activity in primate orbitofrontal cortex reflects the value of time. J. Neurophysiol. 2005; 94:2457–2471. [PubMed: 15958600]

33. Feierstein CE, et al. Representation of spatial goals in rat orbitofrontal cortex. Neuron. 2006; 51:495–507. [PubMed: 16908414]

34. Padoa-Schioppa C, Assad JA. Neurons in orbitofrontal cortex encode economic value. Nature. 2006; 441:223–226. [PubMed: 16633341]

35. Roesch MR, Taylor AR, Schoenbaum G. Encoding of time-discounted rewards in orbitofrontal cortex is independent of value representation. Neuron. 2006; 51:509–520. [PubMed: 16908415]

36. Plassmann H, O'Doherty J, Rangel A. Orbitofrontal cortex encodes willingness to pay in everyday economic transactions. J. Neurosci. 2007; 27:9984–9988. [PubMed: 17855612]

37. van Duuren E, et al. Neural coding of reward magnitude in the orbitofrontal cortex during a five-odor discrimination task. Learn. Mem. 2007; 14:446–456. [PubMed: 17562896]

38. Furuyashiki T, Holland PC, Gallagher M. Rat orbitofrontal cortex separately encodes response and outcome information during performance of goal-directed behavior. J. Neurosci. 2008; 28:5127–5138. [PubMed: 18463266]

39. Hare TA, et al. Dissociating the role of the orbitofrontal cortex and the striatum in the computation of goal values and prediction errors. J. Neurosci. 2008; 28:5623–5630. [PubMed: 18509023]

40. Padoa-Schioppa C, Assad JA. The representation of economic value in the orbitofrontal cortex is invariant for changes in menu. Nat. Neurosci. 2008; 11:95–102. [PubMed: 18066060]

41. Simmons JM, Richmond BJ. Dynamic changes in representations of preceding and upcoming reward in monkey orbitofrontal cortex. Cereb. Cortex. 2008; 18:93–103. [PubMed: 17434918]

42. van Duuren E, Lankelma J, Pennartz CMA. Population coding of reward magnitude in the orbitofrontal cortex of the rat. J. Neurosci. 2008; 28:8590–8603. [PubMed: 18716218]

43. Fitzgerald TH, Seymour B, Dolan RJ. The role of human orbitofrontal cortex in value comparison for incommensurable objects. J. Neurosci. 2009; 29:8388–8395. [PubMed: 19571129]

44. Lara AH, Kennerly SW, Wallis JD. Encoding of gustatory working memory by orbitofrontal neurons. J. Neurosci. 2009; 29:765–774. [PubMed: 19158302]

45. Padoa-Schioppa C. Range-adapting representation of economic value in the orbitofrontal cortex. J. Neurosci. 2009; 29:14004–14014. [PubMed: 19890010]

46. vanDuuren E, et al. Single-cell and population coding of expected reward probability in the orbitofrontal cortex of the rat. J. Neurosci. 2009; 29:8965–8976. [PubMed: 19605634]

47. Kobayashi S, de Carvalho OP, Schultz W. Adaptation of reward sensitivity in orbitofrontal neurons. J. Neurosci. 2010; 30:534–544. [PubMed: 20071516]

48. O'Neill M, Schultz W. Coding of reward risk by orbitofrontal neurons is mostly distinct from coding of reward value. Neuron. 2010; 68:789–800. [PubMed: 21092866]

49. Plassman H, O'Doherty JP, Rangel A. Appetitive and aversive goal values are encoded in the medial orbitofrontal cortex at the time of decision making. J. Neurosci. 2010; 30:10799–10808. [PubMed: 20702709]

50. Kepecs A, et al. Neural correlates, computation and behavioural impact of decision confidence. Nature. 2008; 455:227–231. [PubMed: 18690210]

51. Arana FS, et al. Dissociable contributions of the human amygdala and orbitofrontal cortex to incentive motivation and goal selection. J. Neurosci. 2003; 23:9632–9638. [PubMed: 14573543]

52. Morrison SE, Salzman CD. The convergence of information about rewarding and aversive stimuli in single neurons. J. Neurosci. 2009; 29:11471–11483. [PubMed: 19759296]

53. Kennerley SW, et al. Neurons in the frontal lobe encode the value of multiple decision variables. J. Cogn. Neurosci. 2009; 21:1162–1178. [PubMed: 18752411]

54. Kennerley SW, Wallis JD. Evaluating choices by single neurons in the frontal lobe: outcome value encoded across multiple decision variables. Eur. J. Neurosci. 2009; 29:2061–2073. [PubMed: 19453638]

55. Montague PR, Berns GS. Neural economics and the biological substrates of valuation. Neuron. 2002; 36:265–284. [PubMed: 12383781]

56. Gallagher M, McMahan RW, Schoenbaum G. Orbitofrontal cortex and representation of incentive value in associative learning. J. Neurosci. 1999; 9:6610–6614. [PubMed: 10414988]

57. Machado CJ, Bachevalier J. The effects of selective amygdala, orbital frontal cortex or hippocampal formation lesions on reward assessment in nonhuman primates. Eur. J. Neurosci. 2007; 5:2885–2904. [PubMed: 17561849]

58. Takahashi YK, et al. The orbitofrontal cortex and ventral tegmental area are necessary for learning from unexpected outcomes. Neuron. 2009; 62:269–280. [PubMed: 19409271]

59. Burke KA, et al. Orbitofrontal inactivation impairs reversal of Pavlovian learning by interfering with `disinhibition' of responding for previously unrewarded cues. Eur. J. Neurosci. 2009; 30:1941–1946. [PubMed: 19912335]

60. Tsujimoto S, Genovesio A, Wise SP. Monkey orbitofrontal cortex encodes response choices near feedback time. J. Neurosci. 2009; 29:2569–2574. [PubMed: 19244532]

61. Daw ND, Niv Y, Dayan P. Uncertainty-based competition between prefrontal and dorsolateral striatal systems for behavioral control. Nat. Neurosci. 2005; 8:1704–1711. [PubMed: 16286932]

62. Balleine, BW.; Daw, ND.; O'Doherty, JP. Multiple forms of value learning and the function of dopamine. In: Glimcher, PW., et al., editors. Neuroeconomics: Decision Making and the Brain. Elsevier; Amsterdam: 2008. p. 367-385.

63. Dickinson A, Balleine BW. Motivational control of goal-directed action. Anim.Learn.Behav. 1994; 22:1–18.

64. Holland PC, Straub JJ. Differential effects of two ways of devaluing the unconditioned stimulus after Pavlovian appetitive conditioning. J. Exp. Psychol.: Anim. Behav. Process. 1979; 5:65–78. [PubMed: 528879]

65. Cardinal RN, et al. Emotion and motivation: the role of the amygdala, ventral striatum, and prefrontal cortex. Neurosci. Biobehav. Rev. 2002; 26:321–352. [PubMed: 12034134]

66. Ostlund SB, Balleine BW. Orbitofrontal cortex mediates outcome encoding in Pavlovian but not instrumental learning. J. Neurosci. 2007; 27:4819–4825. [PubMed: 17475789]

67. Parkinson JA, et al. Acquisition of instrumental conditioned reinforcement is resistant to the devaluation of the unconditioned stimulus. Q. J. Exp. Psychol. 2005; 58:19–30.

68. Burke KA, et al. Conditioned reinforcement can be mediated by either outcome-specific or general affective representations. Front. Integr. Neurosci. 2007; 1:2. doi:10.3389/neuro.07/002.2007. [PubMed: 18958230]

69. Burke KA, et al. The role of orbitofrontal cortex in the pursuit of happiness and more specific rewards. Nature. 2008; 454:340–344. [PubMed: 18563088]

70. Pears A, et al. Effects of orbitofrontal cortex lesions on responding with conditioned reinforcement. Brain Cogn. 2001; 47:44–46.

71. Ramirez DR, Savage LM. Differential involvement of the basolateral amygdala, orbitofrontal cortex, and nucleus accumbens core in the acquisition and use of reward expectancies. Behav. Neurosci. 2007; 121:896–906. [PubMed: 17907822]

72. McDannald MA, et al. Lesions of orbitofrontal cortex impair rats' differential outcome expectancy learning but not conditioned stimulus-potentiated feeding. J. Neurosci. 2005; 25:4626–4632. [PubMed: 15872110]

73. Pickens CL, et al. Orbitofrontal lesions impair use of cue-outcome associations in a devaluation task. Behav. Neurosci. 2005; 119:317–322. [PubMed: 15727536]

74. Pickens CL, et al. Different roles for orbitofrontal cortex and basolateral amygdala in a reinforcer devaluation task. J. Neurosci. 2003; 23:11078–11084. [PubMed: 14657165]

75. McDannald MA, et al. Ventral striatum and not orbitofrontal cortex are both required for model-based, but not model-free, reinforcement learning. J. Neurosci. 2011; 31:2700–2705. [PubMed: 21325538]

76. Watt A, Honey RC. Combining CSs associated with the same or different USs. Q. J. Exp. Psychol. 1997; 50B:350–367.

77. Schoenbaum G, et al. Neural activity in orbitofrontal cortex reflects a real-time estimate of rats' expectations for reward. Soc. Neurosci. Abstracts. 2010

78. Holland PC, Rescorla RA. The effects of two ways of devaluing the unconditioned stimulus after first and second-order appetitive conditioning. J. Exp. Psychol.: Anim. Behav. Process. 1975; 1:355–363. [PubMed: 1202141]

79. Pears A, et al. Lesions of the orbitofrontal but not medial prefrontal cortex disrupt conditioned reinforcement in primates. J. Neurosci. 2003; 23:11189–11201. [PubMed: 14657178]

80. Cousens GA, Otto T. Neural substrates of olfac-tory discrimination learning with auditory secondary reinforcement. I. Contributions of the basolateral amygdaloid complex and orbitofrontal cortex. Integr. Physiol. Behav. Sci. 2003; 38:272–294. [PubMed: 15119378]

81. Setlow B, Holland PC, Gallagher M. Disconnection of the basolateral amygdala complex and nucleus accumbens impairs appetitive pavlovian second-order conditioned responses. Behav. Neurosci. 2002; 116:267–275. [PubMed: 11996312]

82. Singh T, et al. Nucleus accumbens core and shell are necessary for reinforcer devaluation effects on pavlovian conditioned responding. Front. Integr. Neurosci. 2010; 4:126. [PubMed: 21088698]

83. Walton ME, et al. Separable learning systems in teh macaque brain and the role of the orbitofrontal cortex in contingent learning. Neuron. 2010; 65:927–939. [PubMed: 20346766]

84. Wallis JD, Anderson KC, Miller EK. Single neurons in prefrontal cortex encode abstract rules. Nature. 2001; 411:953–956. [PubMed: 11418860]

85. Liu T, Stalnaker TA, Schoenbaum G. Neural activity in the orbitofrontal cortex and ventral striatum in response to changes in value versus identity of expected rewards. Soc. Neurosci. Abstracts. 2011 Submitted.

86. Roesch MR, Calu DJ, Schoenbaum G. Dopamine neurons encode the better option in rats deciding between differently delayed or sized rewards. Nat. Neurosci. 2007; 10:1615–1624. [PubMed: 18026098]

87. Roesch MR, et al. Ventral striatal neurons encode the value of the chosen action in rats deciding between differently delayed or sized rewards. J. Neurosci. 2009; 29:13365–13376. [PubMed: 19846724]

88. Roesch MR, Olson CR. Neuronal activity related to reward value and motivation in primate frontal cortex. Science. 2004; 304:307–310. [PubMed: 15073380]

89. Tsuchida A, Doll BB, Fellows LK. Beyond reversal: a critical role for human orbitofrontal cortex in flexible learning from probabilistic feedback. J. Neurosci. 2010; 30:16868–16875. [PubMed: 21159958]

90. Fellows LK. Orbitofrontal contributions to value-based decision making: evidence from humans with frontal lobe damage. Ann. N.Y. Acad. Sci. 2011; 1239:51–58. This volume. [PubMed: 22145875]
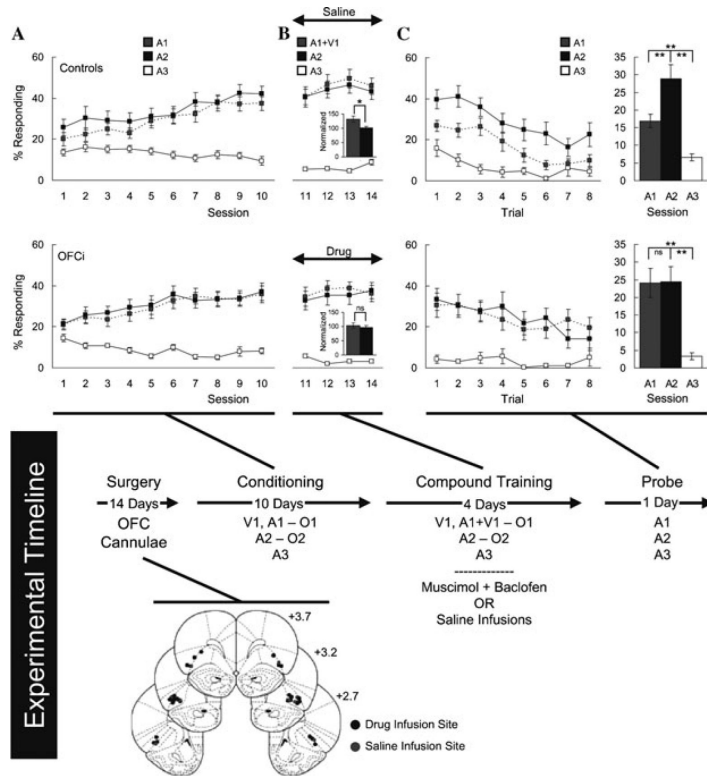
Figure 1.
OFC contribution to Pavlovian overexpectation, adapted from Takahashi *et al.*[58] Shown is the experimental timeline linking conditioning, compound conditioning, and probe phases to data from each phase. Top and bottom rows of plots indicate control (saline infusion into the OFC) and OFCi group (muscimol + baclofen infusion into the OFC), respectively. In timeline and plots, V1 is a visual cue (a cue light); A1, A2, and A3 are auditory cues (tone, white noise, and clicker, counterbalanced), and O1 and O2 are different flavored sucrose pellets (banana and grape, counterbalanced). Position of the cannula within OFC in saline controls (gray dot) and OFCi (black dot) rats are shown beneath the timeline. (A) Percentage of individuals responding to food cup during cue presentation across 10 days of conditioning. Gray, black, and white squares indicate A1, A2, and A3 cues, respectively. (B) Percentage of those responding to food cup during cue presentation across four days of compound training. Gray, black, and white squares indicate A1/V1, A2, and A3 cues, respectively. Gray and black bars in the insets indicate average normalized percentage responding to A1/V1 and A2, respectively. (C) Percentage of those responding to food cup during cue presentation in the probe test. Line graph shows those responding across the eight trials, and the bar graph shows average number responding in these eight trials. Gray, black, and white colors indicate A1, A2, and A3 cues, respectively. * $P < 0.05$ and ** $P < 0.01$ on *post hoc* contrast testing. NS, not significant. Error bars denote SEM.
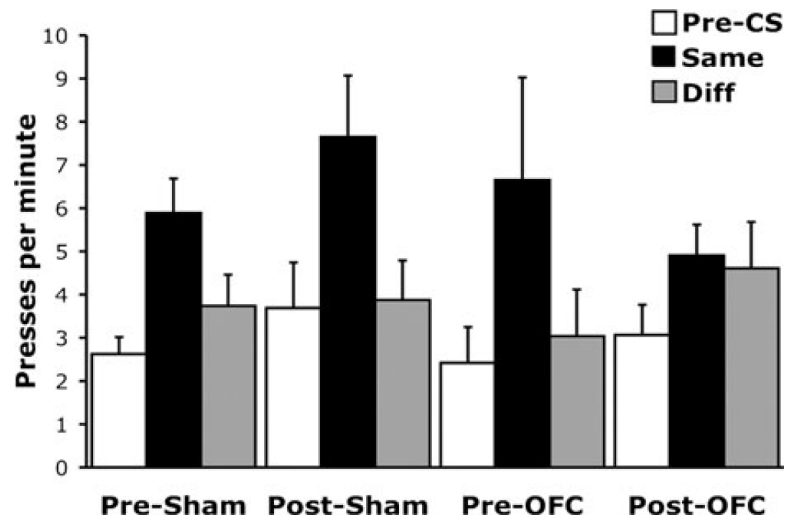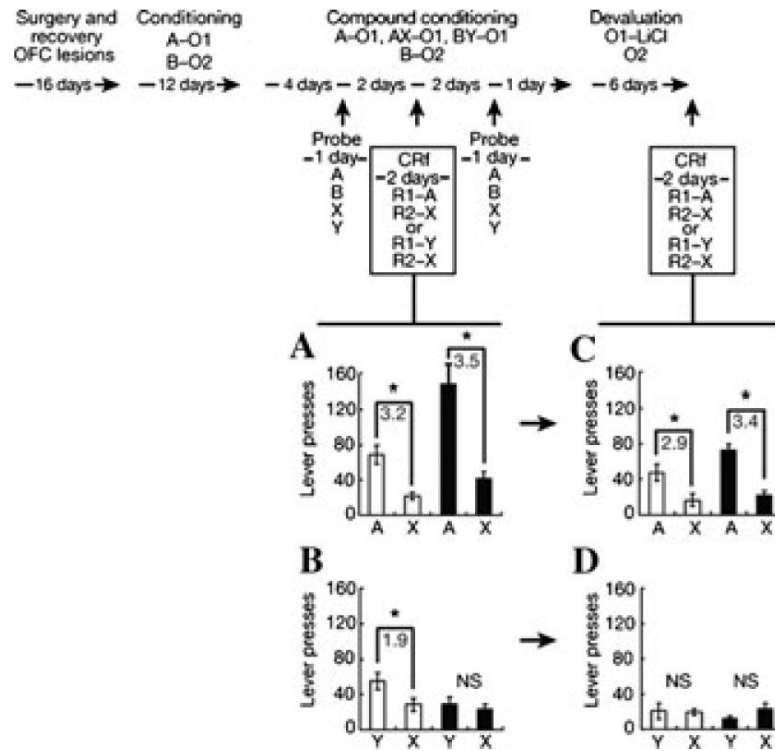
Figure 2.
OFC contribution to reinforcer-specific Pavlovian to instrumental transfer adapted from Ostlund *et al.*[67] Pavlovian training consisting of pairing two different auditory cues with either sucrose or pellet rewards. Sham or OFC lesions were given before or following Pavlovian training. Rats were trained that a right lever press produced pellets while a left lever press produced sucrose solution (counterbalanced). Each rat was tested twice, each with one lever present. Mean lever-press rates are shown during a cue-free baseline period (white bars), a cue signaling the same outcome as the lever press (black bar), and a cue signaling a different outcome (gray bar). *Pre* means Sham or OFC lesion was given before Pavlovian conditioning; *post* means Sham or OFC lesion was given following Pavlovian conditioning. Error bars denote SEM.

Figure 3.
OFC contribution to conditioned reinforcement adapted from Burke *et al.*[69] Shown is the experimental timeline linked to data from each conditioned reinforcement (CRf) test. In the timeline and figures, A, B, X, and Y are training cues; R1 and R2 are instrumental responses; and O1 and O2 are different flavored sucrose pellet reinforcers. Normal (open bars) and OFC-lesioned rats (black bars) were first trained to associate two different auditory cues (A and B) with two differently flavored sugar pellets (banana and grape). After these associations were learned, an "unblocking" phase occurred in which A was compounded with a novel cue, and X and B compounded with a novel cue Y. While AX predicted the original flavor, BY predicted another flavor. This arrangement has been shown to "block" learning to X but "unblock" learning to Y. (A–D) Lever pressing for A versus X, or Y versus X, in control (open bars) and lesioned (filled bars) rats before (A, B) and after (C, D) devaluation. Lesions diminished responding for Y before devaluation (A, B); controls diminished lever pressing for Y after devaluation (C, D). Lever pressing is averaged across two 30-min sessions in each figure. Asterisks indicate significance at $P < 0.05$ on *post hoc* contrast testing; the gray numbers indicate the ratio of responding on the two levers for each significant comparison. NS, not significant. Error bars denote SD.
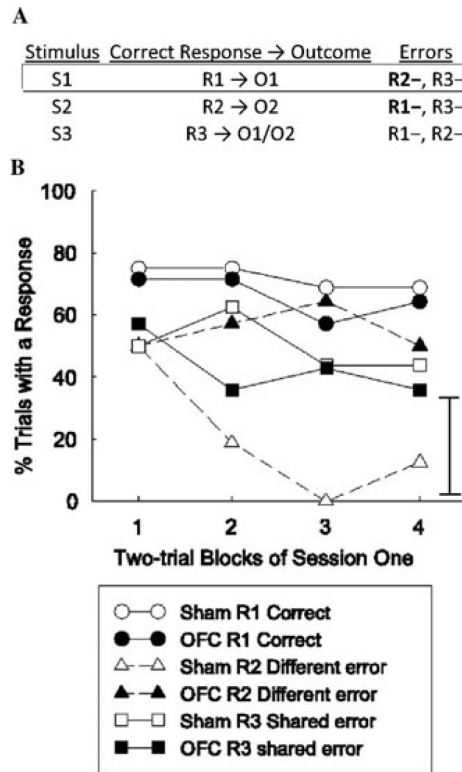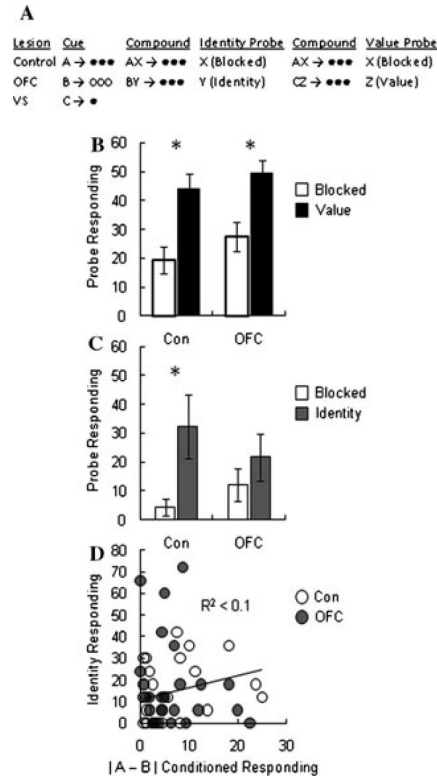
Figure 4.
OFC contribution to differential outcome expectancy; adapted from McDannald *et al.*[76] (A) Sham and OFC-lesioned rats were required to learn the following stimulus-action-outcome associations: S1 → R1 → O1, S2 → R2 → O2, and S3 → R3 → O1/O2. S1–3 were different auditory cues, R1–3 different operant responses, and O1,2 different flavors of sucrose. Performance during S1, in which R1 responses were always reinforced with O1, and errors, consisting of R2 or R3 responses, were not reinforced. R2 responses were always reinforced with a different outcome (O2) on S2 trials, whereas R3 was reinforced with the same outcome (O1) as R1 on one-half of the S3 presentations and with O2 on the other half of the S3 trials. (B) Responding during S1 trials is shown. Sham rats showed more shared-outcome errors (R3) than different-outcome errors (R2) during S1. This result can be attributed to the rats' use of specific outcome expectancies to guide responding: responses reinforced in the presence of different expectancies (e.g., R1 and R2) were differentiated more readily than responses that yielded the same outcome (e.g., R1 and R3). In contrast, OFC-lesioned rats showed equivalent levels of shared- and different-outcome errors. The absence of any difference in shared- and different-outcome errors in these rats is consistent with impairment in the associative basis for outcome expectancy learning. Error bars denote SEM.

A

Lesion	Cue	Compound	Identity Probe	Compound	Value Probe
Control	A → ●●●	AX → ●●●	X (Blocked)	AX → ●●●	X (Blocked)
OFC	B → ○○○	BY → ●●●	Y (Identity)	CZ → ●●●	Z (Value)
VS	C → ●				

Figure 5.

OFC contribution to value and identity unblocking; adapted from McDannald *et al.*[76] (A) Rats were first trained to associate three different visual cues (A, B, C) with three different amounts (3 or 1) and identities (O1 or O2; banana-flavored sugar pellets or grape-flavored sugar pellets) of reward: A → 3xO1, B → 3xO2, and C → O1. Following this phase an unblocking phase was given in which compounds of the original cues and three novel auditory cues, AX, BY and CZ, all predicted 3xO1. In this way X signaled the expected reward (Blocked), Y signaled a differently flavored but similarly valued reward (Identity) and Z signaled a differently valued but similarly flavored reward (Value). (B) Food cup responding to the Blocked cue (which signaled the expected number of pellets) and the Value cue (which signaled greater than expected reinforcement) for Control and OFC-lesioned rats. (C) Food cup responding to the Blocked cue (which signaled the expected flavor of pellets) and the Identity cue (which signaled an equally preferred but differently flavored pellet) for Control and OFC-lesioned rats. (D) The relationship between value sensitivity to cues A and B (which predicted equal amounts of different-flavored sugar pellets) in initial conditioning and the display of identity unblocking for Control and OFC-lesioned rats. Asterisks indicate significance at $P < 0.05$ on *post hoc* contrast testing. NS, not significant. Error bars denote SEM.