Published in final edited form as:

Neurobiol Learn Mem. 2014 February ; 0: 22–27. doi:10.1016/j.nlm.2013.06.003.

Learning theory: a driving force in understanding orbitofrontal function

Michael A. McDannald¹, Joshua L. Jones², Yuji Takahashi¹, and Geoffrey Schoenbaum¹

¹ National Institute on Drug Abuse, Intramural Research Program. Baltimore MD.

² University of Maryland, School of Medicine. Baltimore, MD.

Abstract

Since it was demonstrated the orbitofrontal cortex (OFC) is critical to reversal learning, there has been considerable interest in specifying its role in flexible, outcome-guided behavior. Behavioral paradigms from the learning theory tradition, such as outcome devaluation, blocking, Pavlovian to instrumental transfer, and overexpectation have been a driving force in this research. The use of these procedures has revealed OFC's unique role in forming and integrating information about specific features of events and outcomes to drive behavior and learning. These studies highlight the power and importance of learning theory principles in guiding neuroscience research.

Perhaps the first clear demonstrations of the OFC's role in flexible, outcome-guided behavior and learning came from studies of reversal learning. In reversal studies, subjects must first discriminate a cue that predicts a rewarding outcome from another cue that predicts an aversive outcome or nothing. Discrimination is achieved when subjects respond to the rewarding cue, but withhold responding to the aversive cue. Next, the cue-outcome relationships are reversed. The previously rewarded cue now leads to the aversive outcome and the previously aversive cue now leads to the rewarding outcome. Reversal learning is achieved when behavioral responding reflects the new cue-outcome relationships. It has been demonstrated in a wide variety of species and settings that the OFC is critical for rapid reversal learning (Chudasama and Robbins, 2003; Dias, Robbins, and Roberts, 1996; Izquierdo, Suda, and Murray, 2004b; Jones and Mishkin, 1972; Schoenbaum, Nugent, Saddoris, and Setlow, 2002; Teitelbaum, 1964). This contribution is specific to the reversal, as impairing OFC function does not typically impair acquisition of the initial discrimination. The OFC is then not necessary for learning, per se, but is necessary for learning in the face of changing cue-outcome relationships.

At first glance reversal learning appears simple: subjects learn two cue-outcome relationships, these relationships are reversed, and subject's behavior reflects the reversal. A closer look reveals the simultaneous and concurrent changes taking place during reversal, which make pinpointing the contribution of the OFC difficult. In the initial discrimination subjects learn many associations between cues, responses and outcomes. These associations range from simple, habit-like associations to more cognitive associations by which cues contain information about predicted value and/or sensory features of outcomes, all of which may influence responding. These associations may be rewarding or aversive, adding the dimension of value. Reversal does not consist of changing only one of these associations,

Publisher's Disclaimer: This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final citable form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

but all of them... simultaneously. This requires learning, which invokes the need to signal accurate predictions, in order to engage error and attentional mechanisms to drive learning. Reversal also introduces the concurrent need for behavioral inhibition and engagement. Subjects must withhold a previously learned response while in the same setting engage a previously withheld response. Because reversal performance is assessed in a single format, wherein several processes are concurrently happening, it is difficult to distinguish deficits in these learning mechanisms from deficits in performance.

In some ways the difficulty in interpreting reversal data may account for its popularity. Because reversal learning consists of simultaneous changes in a variety of factors it can provide support for many potential theories of orbitofrontal and prefrontal function. For example, response theories have long claimed the role of the prefrontal cortex – and the OFC in particular - is to inhibit previously learned or prepotent responses (Eagle, Baunez, Hutcheson, Lehmann, Shah, and Robbins, 2008; Ferrier, 1876; Fineberg, Potenza, Chamberlain, Berlin, Menzies, Bechara, Sahakian, Robbins, Bullmore, and Hollander, 2010; Izquierdo and Jentsch, 2012; Jentsch and Taylor, 1999; Jones and Mishkin, 1972; Man, Clarke, and Roberts, 2009). Value theories generally claim the OFC is necessary for the use or assignment of value (Levy and Glimcher, 2011; 2012; Padoa-Schioppa, 2011; Padoa-Schioppa and Assad, 2006; Plassmann, O'Doherty, and Rangel, 2007). This is a common currency in which incommensurable goods (apples and oranges) may be compared. Specific-outcome theories state the OFC is necessary for predicting and integrating specific features of upcoming events, potentially independent of value (Clark, Hollon, and Phillips, 2012; Delamater, 2007; McDannald, Takahashi, Lopatina, Pietras, Jones, and Schoenbaum, 2012; Schoenbaum, Roesch, Stalnaker, and Takahashi, 2009). All of the above hypotheses claim reversal deficits as evidence. Yet while reversal learning was an excellent task to begin investigation of OFC function, this task would not be able to pinpoint a precise role for the OFC in flexible behavior.

Outcome devaluation

A giant leap in our understanding of OFC function has come from studies using outcome devaluation (Colwill and Rescorla, 1985a; b; Holland and Rescorla, 1975; Holland and Straub, 1979). Now known by behavioral neuroscientists, economists and clinicians alike, outcome devaluation has become the gold standard in assessing whether learned behavior is guided by the current value of the outcome. In a typical procedure, subjects first learn that a cue or action leads to a particular reward, for example, a piece of chocolate. Learning is demonstrated when subjects approach the location of chocolate during the predictive cue or perform actions leading to chocolate. Next, the value of the reward is reduced. In the case of chocolate this is achieved either by unlimited consumption, producing selective satiety, or through pairing it's consumption with nausea, forming a conditioned taste aversion. The critical test comes when the cue is presented or action made available in the absence of reward. During such tests, subjects spontaneously decrease cue-evoked responding or actions, reflecting the current value of the reward. Successful outcome devaluation requires two key capacities. First, in initial learning, cues or actions must come to represent detailed information about the rewards they predict. Second, this information must be immediately integrated with current information about that reward's value, so that responding can be adjusted before the reward is ever encountered.

To say this has been a transformative finding for the field of behavioral neuroscience would be a gross understatement. Outcome devaluation procedures have provided a concrete and reliable means of assessing outcome-guided behavior and its underlying neural circuitry. Such studies have repeatedly identified a critical role for the OFC. Lesions or transient inactivation of OFC in rodents (Gallagher, McMahan, and Schoenbaum, 1999; Pickens,

Saddoris, Gallagher, and Holland, 2005; Pickens, Saddoris, Setlow, Gallagher, Holland, and Schoenbaum, 2003) and macaques (Izquierdo and Murray, 2004; 2010; Izquierdo, Suda, and Murray, 2004a; Machado and Bachevalier, 2007; West, DesJardin, Gale, and Malkova, 2011) impair performance in outcome devaluation. Importantly, the deficits observed in these studies are not due to general impairments in learning. Interfering with OFC function typically does not impair acquisition of either Pavlovian or instrumental responding in acquisition. Further, the OFC is not necessary for choosing between valued and devalued rewards when they are present. The OFC is specifically necessary for integrating information about reward-predictive cues and the new value of those rewards. Consistent with work from rodents and macaques, human neuroimaging has found that BOLD signals in the OFC reflect both a cues relationship with an outcome and this BOLD response is decreased in outcome devaluation (Gottfried, O'Doherty, and Dolan, 2003). Thus, across a range of species the OFC is both necessary for and encodes information relevant to outcome devaluation procedures (Murray, O'Doherty, and Schoenbaum, 2007).

Despite the requirement that specific, cue-outcome associations be integrated to successfully decrease responding, the OFC's role in outcome devaluation has been cited as evidence for roles in both value and response inhibition. These descriptions have some merit. Successful demonstration of outcome devaluation involves knowledge of a reward's current value as well as the ability to decrease responding in the face of this knowledge. Fortunately, studies of learning theory have led to a wealth of behavioral procedures that isolate contributions of specific-outcome, value and response inhibition processes to learning and behavior. In the following sections we describe a handful of these procedures that have been especially useful.

Unblocking and transfer

Studies of unblocking and Pavlovian-to-instrumental transfer (PIT) have been highly valuable in isolating learning and behavior driven by general value versus specific outcomes. Blocking and unblocking are at the heart of modern learning theory. In a typical blocking study subjects first learn that a cue predicts reward ($A \rightarrow \text{Food}$). Next a novel cue is added and followed by the same reward ($AX \rightarrow \text{Food}$). The critical test comes when reward responding to the added cue X is tested in isolation. When this is done subjects show little or no responding to X, compared to another group that did not receive $A \rightarrow \text{Food}$ learning (Kamin, 1969). Prior training of A 'blocked' learning to X.

In a popular explanation, blocking of learning to the added cue occurs because the previously trained cue fully and accurately predicts all aspects of the reward. If blocking is about information then changing aspects of the reward should result in 'unblocking' or learning to the added cue. This is because the added cue is now in a position to provide new information about the reward. Studies by Dr. Peter Holland and colleagues have repeatedly demonstrated such unblocking effects (Holland, 1984; 1988; Holland and Gallagher, 1993; Holland and Kenmuir, 2005). In one form of unblocking, rats receive extensive training of a cue and a specific amount of reward. Once well-learned, a novel cue is added and the amount of reward is increased (Holland, 1984). Because the novel cue now predicts a larger reward, unblocking occurs. Importantly, this learning need not rely on specific-outcome properties of reward. This form of unblocking might be said to require information *about value*. Rescorla hypothesized that unblocking ought to occur even when there is no explicit change in value but only a change in the quality, or sensory-feature of rewards. To demonstrate this Rescorla first trained rats that two separate cues predicted two discriminable, yet equally valued rewards ($A \rightarrow O1$, $B \rightarrow O2$). Next novel cues were compounded with the initially trained cues and the reward changed in one condition ($AX \rightarrow O1$, $BY \rightarrow O1$). Using a transfer test he reported blocking of learning to the added cue

X, that provided no information about reward but substantial unblocking to cue Y, which selectively provided information about specific features of reward. Thus, learning may occur in situations in which value is held constant, but specific features of rewards are altered (Rescorla, 1999).

The unblocking data reported by Rescorla are consistent with results of selective Pavlovian-to-instrumental transfer (PIT) routinely reported by Dr. Bernard Balleine and colleagues (Corbit and Balleine, 2005; 2011; Corbit, Janak, and Balleine, 2007). In these studies subjects are taught that two separate instrumental responses lead to two discriminable yet equally preferred rewards ($R1 \rightarrow O1$, $R2 \rightarrow O2$). At the same time subjects also learn that two environmental cues lead to these different rewards ($A \rightarrow O1$, $B \rightarrow O2$). Once both instrumental and Pavlovian learning is established, a transfer test is given in which the cues are presented over the top of instrumental responding but no rewards are present. During presentation of a specific cue, subjects will typically perform actions leading to that same reward much more than action leading to the alternate reward. For example, cue A selectively augments responding for R1, while B selectively augments responding for R2. These results demonstrate that in initial learning both the cues and the actions came to predicts specific features of the rewards with which they were associated. Such specificity could not be achieved if subjects had simply come to make predictions of value over the course of learning, as there was no difference in the value of the two rewards.

Unblocking and selective PIT provide the ideal means of comparing OFC's contribution to signaling of value versus more specific information about outcomes in learning and behavior. Value theories predict the OFC should be critical to unblocking based on increases in reward value but not to unblocking driven by changes in specific-outcome properties of rewards. Similarly, OFC should not be necessary for selective transfer because differences in value are not driving the transfer effect. Specific-outcome theories predict just the opposite. OFC should only be necessary for unblocking when predictions of specific features of reward are necessary, but not when predictions of general value would suffice. OFC should also be necessary for the demonstration of specific transfer.

Recent findings from our lab support the specific-outcome theory (Burke, Franz, Miller, and Schoenbaum, 2008; McDannald, Lucantonio, Burke, Niv, and Schoenbaum, 2011). OFC lesions made prior to training procedures leave unblocking based on value fully intact but impair unblocking based on specific reward features (Figure 1). This was not due to general deficits in learning. OFC lesions rats had no problems in displaying overt conditioning and could fully discriminate between two differently-flavored rewards. Rather, this deficit was specific to new learning driven by changes in specific reward features.

Results from specific transfer studies are also consistent with specific-outcome theories of OFC function. OFC lesions have no impact on overt demonstration of instrumental or Pavlovian conditioning but impair the demonstration of specific transfer (Ostlund and Balleine, 2007; Scarlet, Delamater, Campese, Fein, and Wheeler, 2012). Critical roles for the OFC in unblocking and specific transfer are difficult to explain within value theories but are instead consistent with specific-outcome theories of OFC function.

Sensory preconditioning

In both blocking and transfer procedures aspects of rewards are controlled or manipulated to produce learning and behavior that could be driven by changes in specific-outcome properties. However, in both of these cases learning and behavior is still tied to properties of rewards. If the OFC is more broadly necessary for using specific information about upcoming events, this should extend to settings in which learning occurs in the absence of

intrinsically rewarding or valued events. Sensory preconditioning provides such a setting (Brogden, 1939).

In studies of sensory preconditioning animals first learn associations between neutral, auditory cues ($A \rightarrow B$ and $C \rightarrow D$). After this initial training, cue B is made to predict reward ($B \rightarrow \text{Food}$) while cue D predicts nothing ($D \rightarrow \emptyset$). In a subsequent probe test, responding spontaneously increases to the preconditioned cue A, via its association with B, while C responding remains low. In this instantiation, sensory preconditioning provides a direct means of assessing behavior driven by two distinct mechanisms: cached value and inferred value. Cached value representations are values that are pre-computed based on prior reinforcement history and stored in a cue. In contrast, inferred value reflects information that must be calculated on-the-fly, based upon that cue's relationship with other environmental cues and their outcomes. While both types of responding would appear to reflect value, only the inferred value requires a detailed knowledge of the causal structure of the environment. In sensory preconditioning, differential responding to cues A and C (but not B and D) must reflect access to such a model-based representation, because animals must correctly integrate the value-independent stimulus-stimulus association first learned with the subsequent association between stimulus and reward.

As such, sensory preconditioning provides an ideal means of testing value and specific-outcome theories of OFC function. Value theories would certainly predict that OFC inactivation during probe test should impair responding to cues B and D. Specific-outcome theories predict that OFC inactivation should have little or no impact on B & D responding but should significantly impair differential responding to A & C. Indeed, we have recently demonstrated this pattern of responding (Figure 2). Following training on the SPC paradigm, inactivation of the OFC during the probe test has no significant effect on responding to B & D, but significantly impairs differential responding to A & C. Thus, consistent with the specific-outcome theory, the OFC is necessary when responding must be supported by inferred value or model-based representations of the outcome, but not when responding can be based on a cached value.

Over-expectation

Until now we have discussed behavioral procedures that allow direct tests of specific-outcome and value theories of OFC function. Now we focus on procedures that allow for a direct test of specific-outcome and response inhibition theories: Pavlovian over-expectation and its comparison with extinction procedures. In studies of extinction learning, subjects first learn that a cue leads to an aversive or rewarding outcome. Learning is evident when the subjects display reward-related or aversive behavior to the cue. Extinction occurs when the cue is repeatedly presented but the outcome omitted. Over the course of extinction subjects will reduce the behavior that was initially acquired to the cue. It is generally recognized that simple extinction does not consist of unlearning the old cue-outcome association but instead of learning a new, typically inhibitory association between the cue and the absence of the outcome. Learning is thought to occur because there is discrepancy between the outcome that is predicted (the rewarding or aversive outcome) and the outcome that is received (nothing). In simple extinction the difference between these events is rather extreme: something vs. nothing. Thus, simple extinction may be the result of non-reinforcement. The outcome can be seen as propping up behavior and its removal results in the behavior's loss. Rescorla proposed that if extinction learning is driven by comparisons between predicted and received outcomes; it ought to occur in less extreme circumstances and when outcomes are actually presented during 'extinction'.

In his initial demonstration, Rescorla used aversive learning procedures to assess extinction through over-expectation (Rescorla, 1970; 2006; 2007). More recently we have implemented an appetitive version of this procedure (Haney, Calu, Takahashi, Hughes, and Schoenbaum, 2010; Takahashi, Roesch, Stalnaker, Haney, Calu, Taylor, Burke, and Schoenbaum, 2009c). In this task, rats are first trained that several cues independently predict reward. Subsequently, two cues are presented in compound and followed by the same reward. When responding to the individual cues is assessed in a probe test, rats exhibit reduced responding to the previously compounded cues. This reduced responding is thought to result from the violation of summed predictions for reward during compound training. In other words, in compound training, the rat predicts receipt of two rewards, one for each cue, but only obtains one. The resulting discrepancy between predicted and received outcomes drives inhibitory learning, leading to reduced responding when the cues are presented alone.

Comparing the role of the OFC in simple extinction versus extinction through over-expectation is an ideal means of testing response and specific outcome theories of OFC function. If the primary function of the OFC is response inhibition then OFC should be necessary for both forms of extinction, and it should be required at the time of the test, when responding must be inhibited. However, if the primary function of the OFC is to integrate predictions about outcomes from multiple sources and use this information to drive learning then OFC should only be necessary for extinction driven by over-expectation and its role should be restricted to the learning phase. We have recently showed just this pattern of effect (Figure 3) (Takahashi, Roesch, Stalnaker, Haney, Calu, Taylor, Burke, and Schoenbaum, 2009a). Rats were implanted with bilateral cannula aimed at their OFC. On the critical days of over-expectation, when the cues were presented in compound, neural activity in the OFC was disrupted by infusion of a cocktail of baclofen and muscimol (GABA agonists). This disruption of OFC function prevented rats from showing summation of responding to the two cues, an indicator of a lack of integration of the separate predictions. More importantly, OFC-inactivated rats also failed to show evidence of inhibitory learning in the probe test, when behavioral responding to one of the previously compounded cues was tested in extinction.

Subsequent work showed that rats with OFC lesions, in whom OFC function would be disrupted throughout, showed exactly the same pattern of impairment during over-expectation, including intact extinction learning during the probe test, when extinction was driven by reward omission (Takahashi et al., 2009a, supplemental data). Further, in a separate assessment, rats from this study were later tested in a Pavlovian reversal task. Discriminative responding was established to one cue signaling reward and another signaling nothing. Once established, the cue-reward associations were reversed while half of the rats received OFC inactivation with baclofen/muscimol while the remaining received control infusions leaving OFC intact (Burke, Takahashi, Correll, Leon Brown, and Schoenbaum, 2009). OFC inactivation impaired the reversal of responding, however, this impairment was selective to acquiring responding to the newly reward cue. OFC-inactivated rats showed normal extinction to the previously rewarded cue. The OFC is not necessary for the demonstration of simple extinction (Figure 3).

Thus OFC appears to be necessary for extinction by over-expectation but not for extinction by omission. These results are difficult to explain within response inhibition theories of OFC function but are consistent with specific-outcome theories.

Upward revaluation

So far we have discussed behavioral procedures from the learning theory tradition that have already been implemented by our lab and others to describe specific the contribution of the

OFC to learning and behavior. In addition to these procedures, many more have been developed by Holland, Rescorla and other learning theorists that have received little or no attention but will undoubtedly serve as valuable tools in future. One such procedure is upward revaluation (Rescorla, 1990). The premise of upward revaluation is very similar to outcome devaluation. A thirsty animal is first trained that a cue or response leads to flavored water. Once the cue-flavor association is well learned the value of the flavor is increased. This can be done by administering a conditioned flavor preference, in which consumption of the flavored water is followed by consumption of palatable sugar water. When this is done animals will come to prefer the flavor associated with sugar. Next responding to the cue or response is assessed in extinction. When this is done animals show much slower extinction to a flavor for which a conditioned flavor preference has been acquired, compared to another flavor for which no preference was acquired.

While similar to outcome devaluation in most respects the critical difference in upward revaluation is that the new value of the reward will *augment* responding, rather than reduce it. For those interested in comparing response and specific-outcome theories of the OFC it is difficult to think of a better experiment. If the OFC is necessary for using and integrating information about specific-outcomes then it would be absolutely necessary for the demonstration of upward revaluation. However, if the OFC is only necessary for response inhibition then interfering with OFC function would have absolutely no effect on upward revaluation. There is no behavior for the OFC to inhibit. Of course upward revaluation is just one of many behavioral procedures that have received little or no attention from neuroscientists.

Summary

Behavioral procedures such as reversal learning have been an excellent starting point for investigating the role of certain neural substrates in value-guided behavior. However, reversal-learning tasks, as we have shown, are not able to isolate the specific contributions of brain regions to very specific aspects of learning and behavior. It is for this reason that behavioral procedures from the learning theory tradition are highly valuable tools for neuroscientists, because they can allow the isolation of one or two variables of interest and allow for tests of highly specific theories. These tools have already greatly increased our understanding of orbitofrontal function. In moving forward this learning theory approach will be critical to further specifying the contribution of orbitofrontal and prefrontal cortices to flexible, outcome-directed behavior and learning.

Literature Cited

- Brogden WJ. Sensory Pre-Conditioning. *Journal of Experimental Psychology*. 1939; 25:323–332.
- Burke KA, Franz TM, Miller DN, Schoenbaum G. The role of the orbitofrontal cortex in the pursuit of happiness and more specific rewards. *Nature*. 2008; 454:340–344. [PubMed: 18563088]
- Burke KA, Takahashi YK, Correll J, Leon Brown P, Schoenbaum G. Orbitofrontal inactivation impairs reversal of Pavlovian learning by interfering with 'disinhibition' of responding for previously unrewarded cues. *European Journal of Neuroscience*. 2009; 30:1941–1946. [PubMed: 19912335]
- 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. *Journal of Neuroscience*. 2003; 23:8771–8780. [PubMed: 14507977]
- Clark JJ, Hollon NG, Phillips PEM. Pavlovian valuation systems in learning and decision making. *Current Opinion in Neurobiology*. 2012; 22:1054–1061. [PubMed: 22749132]

- Colwill RM, Rescorla RA. Instrumental Responding Remains Sensitive to Reinforcer Devaluation after Extensive Training. *Journal of Experimental Psychology-Animal Behavior Processes*. 1985a; 11:520–536.
- Colwill RM, Rescorla RA. Postconditioning Devaluation of a Reinforcer Affects Instrumental Responding. *Journal of Experimental Psychology-Animal Behavior Processes*. 1985b; 11:120–132.
- Corbit LH, Balleine BW. Double dissociation of basolateral and central amygdala lesions on the general and outcome-specific forms of pavlovian-instrumental transfer. *Journal of Neuroscience*. 2005; 25:962–970. [PubMed: 15673677]
- Corbit LH, Balleine BW. The general and outcome-specific forms of pavlovian-instrumental transfer are differentially mediated by the nucleus accumbens core and shell. *J Neurosci*. 2011; 31:11786–11794. [PubMed: 21849539]
- Corbit LH, Janak PH, Balleine BW. General and outcome-specific forms of Pavlovian-instrumental transfer: the effect of shifts in motivational state and inactivation of the ventral tegmental area. *European Journal of Neuroscience*. 2007; 26:3141–3149. [PubMed: 18005062]
- Delamater AR. The role of the orbitofrontal cortex in sensory-specific encoding of associations in pavlovian and instrumental conditioning. *Ann N Y Acad Sci*. 2007; 1121:152–173. [PubMed: 17872387]
- Dias R, Robbins TW, Roberts AC. Dissociation in prefrontal cortex of affective and attentional shifts. *Nature*. 1996; 380:69–72. [PubMed: 8598908]
- Eagle DM, Baunez C, Hutcheson DM, Lehmann O, Shah AP, Robbins TW. Stop-signal reaction-time task performance: role of prefrontal cortex and subthalamic nucleus. *Cerebral Cortex*. 2008; 18:178–188. [PubMed: 17517682]
- Ferrier, D. *The Functions of the Brain*. GP Putnam's Sons; New York: 1876.
- Fineberg NA, Potenza MN, Chamberlain SR, Berlin HA, Menzies L, Bechara A, Sahakian BJ, Robbins TW, Bullmore ET, Hollander E. Probing compulsive and impulsive behaviors, from animal models to endophenotypes: a narrative review. *Neuropsychopharmacology*. 2010; 35:591–604. [PubMed: 19940844]
- Gallagher M, McMahan RW, Schoenbaum G. Orbitofrontal cortex and representation of incentive value in associative learning. *J Neurosci*. 1999; 19:6610–6614. [PubMed: 10414988]
- Gottfried JA, O'Doherty J, Dolan RJ. Encoding predictive reward value in human amygdala and orbitofrontal cortex. *Science*. 2003; 301:1104–1107. [PubMed: 12934011]
- Haney RZ, Calu DJ, Takahashi YK, Hughes BW, Schoenbaum G. Inactivation of the central but not the basolateral nucleus of the amygdala disrupts learning in response to overexpectation of reward. *J Neurosci*. 2010; 30:2911–2917. [PubMed: 20181588]
- Holland PC. Unblocking in Pavlovian appetitive conditioning. *Journal of Experimental Psychology: Animal Behavior Processes*. 1984; 10:476–497. [PubMed: 6491608]
- Holland PC. Excitation and inhibition in unblocking. *Journal of Experimental Psychology: Animal Behavior Processes*. 1988; 14:261–279. [PubMed: 3404081]
- Holland PC, Gallagher M. Effects of amygdala central nucleus lesions on blocking and unblocking. *Behavioral Neuroscience*. 1993; 107:235–245. [PubMed: 8484889]
- Holland PC, Kenmuir C. Variations in unconditioned stimulus processing in unblocking. *J Exp Psychol Anim Behav Process*. 2005; 31:155–171. [PubMed: 15839773]
- Holland PC, Rescorla RA. The Effect of Two Ways of Devaluing the Unconditioned Stimulus after First- and Second-Order Appetitive Conditioning. *Journal of Experimental Psychology: Animal Behavior Processes*. 1975; 1:355–363. [PubMed: 1202141]
- Holland PC, Straub JJ. Differential effects of two ways of devaluing the unconditioned stimulus after Pavlovian appetitive conditioning. *Journal of Experimental Psychology: Animal Behavior Processes*. 1979; 5:65–78. [PubMed: 528879]
- Izquierdo A, Jentsch JD. Reversal learning as a measure of impulsive and compulsive behavior in addictions. *Psychopharmacology*. 2012; 219:607–620. [PubMed: 22134477]
- Izquierdo A, Murray EA. Combined unilateral lesions of the amygdala and orbital prefrontal cortex impair affective processing in rhesus monkeys. *J Neurophysiol*. 2004; 91:2023–2039. [PubMed: 14711973]

- Izquierdo A, Murray EA. Functional interaction of medial mediodorsal thalamic nucleus but not nucleus accumbens with amygdala and orbital prefrontal cortex is essential for adaptive response selection after reinforcer devaluation. *J Neurosci.* 2010; 30:661–669. [PubMed: 20071531]
- Izquierdo A, Suda RK, Murray EA. Bilateral orbital prefrontal cortex lesions in rhesus monkeys disrupt choices guided by both reward value and reward contingency. *J Neurosci.* 2004a; 24:7540–7548. [PubMed: 15329401]
- Izquierdo AD, Suda RK, Murray EA. Bilateral orbital prefrontal cortex lesions in rhesus monkeys disrupt choices guided by both reward value and reward contingency. *Journal of Neuroscience.* 2004b; 24:7540–7548. [PubMed: 15329401]
- Jentsch JD, Taylor JR. Impulsivity resulting from frontostriatal dysfunction in drug abuse: implications for the control of behavior by reward-related stimuli. *Psychopharmacology.* 1999; 146:373–390. [PubMed: 10550488]
- Jones B, Mishkin M. Limbic lesions and the problem of stimulus-reinforcement associations. *Experimental Neurology.* 1972; 36:362–377. [PubMed: 4626489]
- Jones JL, Esber GR, McDannald MA, Gruber AJ, Hernandez A, Mirens A, Schoenbaum G. Orbitofrontal Cortex Supports Behavior and Learning Using Inferred But Not Cached Values. *Science.* 2012; 338:953–956. [PubMed: 23162000]
- Kamin, LJ. Predictability, surprise, attention, and conditioning.. In: Campbell, BA.; Church, RM., editors. Punishment and aversive behavior. Appleton-Century-Crofts; New York: 1969.
- Levy DJ, Glimcher PW. Comparing apples and oranges: Using reward-specific and reward-general subjective value representation in the brain. *Journal of Neuroscience.* 2011; 31:14693–14707. [PubMed: 21994386]
- Levy DJ, Glimcher PW. The root of all value: a neural common currency for choice. *Current Opinion in Neurobiology.* 2012 epub ahead of print.
- Machado CJ, Bachevalier J. The effects of selective amygdala, orbital frontal cortex or hippocampal formation lesions on reward assessment in nonhuman primates. *European Journal of Neuroscience.* 2007; 25:2885–2904. [PubMed: 17561849]
- Man MS, Clarke HF, Roberts AC. The role of the orbitofrontal cortex and medial striatum in the regulation of prepotent responses to food rewards. *Cerebral Cortex.* 2009; 19:899–906. [PubMed: 18689858]
- McDannald MA, Lucantonio F, Burke KA, Niv Y, Schoenbaum G. Ventral Striatum and Orbitofrontal Cortex Are Both Required for Model-Based, But Not Model-Free, Reinforcement Learning. *Journal of Neuroscience.* 2011; 31:2700–2705. [PubMed: 21325538]
- McDannald MA, Takahashi YK, Lopatina N, Pietras BW, Jones JL, Schoenbaum G. Model-based learning and the contribution of the orbitofrontal cortex to the model-free world. *European Journal of Neuroscience.* 2012; 35:991–996. [PubMed: 22487030]
- Murray EA, O'Doherty JP, Schoenbaum G. What we know and do not know about the functions of the orbitofrontal cortex after 20 years of cross-species studies. *J Neurosci.* 2007; 27:8166–8169. [PubMed: 17670960]
- Ostlund SB, Balleine BW. Orbitofrontal cortex mediates outcome encoding in Pavlovian but not instrumental conditioning. *J Neurosci.* 2007; 27:4819–4825. [PubMed: 17475789]
- Padoa-Schioppa C. Neurobiology of economic choice: a goods-based model. *Annual Review of Neuroscience.* 2011; 34:333–359.
- Padoa-Schioppa C, Assad JA. Neurons in orbitofrontal cortex encode economic value. *Nature.* 2006; 441:223–226. [PubMed: 16633341]
- Pickens CL, Saddoris MP, Gallagher M, Holland PC. Orbitofrontal lesions impair use of cue-outcome associations in a devaluation task. *Behav Neurosci.* 2005; 119:317–322. [PubMed: 15727536]
- Pickens CL, Saddoris MP, Setlow B, Gallagher M, Holland PC, Schoenbaum G. Different roles for orbitofrontal cortex and basolateral amygdala in a reinforcer devaluation task. *J Neurosci.* 2003; 23:11078–11084. [PubMed: 14657165]
- Plassmann H, O'Doherty J, Rangel A. Orbitofrontal cortex encodes willingness to pay in everyday economic transactions. *Journal of Neuroscience.* 2007; 27:9984–9988. [PubMed: 17855612]
- Rescorla RA. Reduction in Effectiveness of Reinforcement after Prior Excitatory Conditioning. *Learning and Motivation.* 1970; 1:372–381.

- Rescorla RA. Instrumental Responses Become Associated with Reinforcers That Differ in One Feature. *Animal Learning & Behavior*. 1990; 18:206–211.
- Rescorla RA. Learning about qualitatively different outcomes during a blocking procedure. *Animal Learning & Behavior*. 1999; 27:140–151.
- Rescorla RA. Spontaneous recovery from overexpectation. *Learn Behav*. 2006; 34:13–20. [PubMed: 16786880]
- Rescorla RA. Renewal after overexpectation. *Learn Behav*. 2007; 35:19–26. [PubMed: 17557388]
- Scarlet J, Delamater AR, Campese V, Fein M, Wheeler DS. Differential involvement of the basolateral amygdala and orbitofrontal cortex in the formation of sensory-specific associations in conditioned flavor preference and magazine approach paradigms. *European Journal of Neuroscience*. 2012; 35:1799–1809. [PubMed: 22672133]
- Schoenbaum G, Nugent S, Saddoris MP, Setlow B. Orbitofrontal lesions in rats impair reversal but not acquisition of go, no-go odor discriminations. *Neuroreport*. 2002; 13:885–890. [PubMed: 11997707]
- Schoenbaum G, Roesch MR, Stalnaker TA, Takahashi YK. A new perspective on the role of the orbitofrontal cortex in adaptive behaviour. *Nature Reviews Neuroscience*. 2009; 10:885–892.
- Takahashi Y, Roesch MR, Stalnaker TA, Haney RZ, Calu DJ, Taylor AR, Burke KA, Schoenbaum G. The orbitofrontal cortex and ventral tegmental area are necessary for learning from unexpected outcomes. *Neuron*. 2009a; 62:269–280. [PubMed: 19409271]
- Takahashi YK, Roesch MR, Stalnaker TA, Haney RZ, Calu DJ, Taylor AR, Burke KA, Schoenbaum G. The Orbitofrontal Cortex and Ventral Tegmental Area Are Necessary for Learning from Unexpected Outcomes. *Neuron*. 2009b; 62:269–280. [PubMed: 19409271]
- Takahashi YK, Roesch MR, Stalnaker TA, Haney RZ, Calu DJ, Taylor AR, Burke KA, Schoenbaum G. The orbitofrontal cortex and ventral tegmental area are necessary for learning from unexpected outcomes. *Neuron*. 2009c; 62:269–280. [PubMed: 19409271]
- Teitelbaum H. A comparison of effects of orbitofrontal and hippocampal lesions upon discrimination learning and reversal in the cat. *Experimental Neurology*. 1964; 9:452–462. [PubMed: 14188532]
- West EA, DesJardin JT, Gale K, Malkova L. Transient inactivation of orbitofrontal cortex blocks reinforcer devaluation in macaques. *J Neurosci*. 2011; 31:15128–15135. [PubMed: 22016546]

Highlights

Orbitofrontal cortex (OFC) is critical to goal-directed behavior and learning.

Learning theory paradigms reveal a unique for OFC in signaling specific outcomes.

Learning theory a powerful tool in understanding prefrontal and neural function.

Figure 1. Contributions of the orbitofrontal cortex to blocking. The effects of pre-training OFC lesions on unblocking driven by changes in reward flavor or increases in reward value are summarized (McDannald et al., 2011). (A) Rats were trained to associate three visual cues with different flavors and quantities of reward. In an unblocking phases these visual cues were compounded with novel auditory and aspects of the reward selectively changed. A probe test was administered in which the novel auditory cues were presented in isolation and extinction. (B) Representative control (left) and neurotoxic lesion (right) are shown. (C) Food cup responding to cues X and Y are shown for control (left) and OFC-lesioned rats (right). Only control rats showed evidence of learning to cue Y, the cue signaling a change in reward flavor. (D) Food cup responding to cues X and Z are shown for control (left) and OFC-lesioned rats (right). Both control and OFC-lesioned rats showed evidence of learning to cue Z, the cue signaling an increase in reward number.

Figure 2.

Contributions of orbitofrontal cortex to sensory preconditioning. The effects of OFC inactivation on the expression of sensory preconditioning are summarized (Jones, Esber, McDannald, Gruber, Hernandez, Mirenzi, and Schoenbaum, 2012). (A) Rats were first trained to associate two pairs of neutral, auditory cues in a preconditioning phase. Next two of those cues were presented in conditioning phase in which one predicted reward while a second did not. Finally a probe test was administered in which all auditory cues were presented in isolation and extinction. (B) Cannula placement for control (black) and baclofen/muscimol (grey) infusions are shown. Infusions were given only on the probe test day. (C) Food cup responding to cues A and C are shown for control (left) and OFC-lesioned rats (right). Only control rats showed evidence of learning to cue A, the cue from which inferred value was calculated. (D) Food cup responding to cues B and D are shown for control (left) and OFC-lesioned rats (right). Both control and OFC-inactivated showed

significant differential responding to cues B and D, responding that could be supported by cached value.

Figure 3.

Contributions of orbitofrontal cortex to Pavlovian overexpectation. The effects of OFC inactivation on the expression of Pavlovian overexpectation and extinction are summarized (Burke et al., 2009; Takahashi, Roesch, Stainaker, Haney, Caiu, Taylor, Burke, and Schoenbaum, 2009b). (A) Rats were first trained to associate two auditory and one visual cue with differing reward identities. Next the visual and auditory cues predicting the same outcome were presented in compound while predicting the same outcome. A probe test was administered in the auditory cues were presented in isolation and extinction. (B) Cannula placement for control (grey) and baclofen/muscimol (black) infusions are shown. Infusions were made on the four days of compound training. (C) Food cup responding to cues A and C are shown for control (left) and OFC-lesioned rats (right). Only control rats showed evidence of inhibitory learning to cue A1, the cue presented in compound with V1, producing overexpectation. (D) Following overexpectation testing these same rats were tested in reversal learning. Following learning that one cue signaled reward a second did not,

these relationships were reversed. During reversal half of the rats received OFC inactivation with Baclofen/Muscimol while the remaining received control infusions leaving OFC intact. Food cup responding for the rewarded cue prior to reversal (pre) and on the final day of extinction (S5) for are both groups are shown. Control and OFC-inactivated rats showed equivalent levels of extinction.