

Review

Thinking Outside the Box: Orbitofrontal Cortex, Imagination, and How We Can Treat Addiction

Geoffrey Schoenbaum^{*,1}, Chun-Yun Chang¹, Federica Lucantonio¹ and Yuji K Takahashi¹

¹NIDA Intramural Research Program, Baltimore, MD, USA

Addiction involves an inability to control drug-seeking behavior. While this may be thought of as secondary to an overwhelming desire for drugs, it could equally well reflect a failure of the brain mechanisms that allow addicts to learn about and mentally simulate non-drug consequences. Importantly, this process of mental simulation draws upon, but is not normally bound by, our past experiences. Rather we have the ability to think outside the box of our past, integrating knowledge gained from a variety of similar and not-so-similar life experiences to derive estimates or imagine what might happen next. These estimates influence our current behavior directly and also affect future behavior by serving as the background against which outcomes are evaluated to support learning. Here we will review evidence, from our own work using a Pavlovian over-expectation task as well as from other sources, that the orbitofrontal cortex is a critical node in the neural circuit that generates these estimates. Further we will offer the specific hypothesis that degradation of this function secondary to drug-induced changes is a critical and likely addressable part of addiction.

Neuropsychopharmacology (2016) **41**, 2966–2976; doi:10.1038/npp.2016.147; published online 14 September 2016

The future exists; it is in our imagination. There, we consider options, evaluate potential consequences, and make our plans. This process of mental simulation draws upon, but is not typically bound by, our past experiences. This is because most situations are a bit unlike prior experiences. Further, because life is complex and sparsely sampled, our knowledge of cause and effect is rarely complete, so prior experiences often do not provide concrete answers for what to do in a given situation. To circumvent this limitation, we can think outside the box of our past, integrating knowledge gained from a variety of similar and not-so-similar experiences to derive estimates about—imagine—what might happen in a new situation. These estimates—containing both cognitive and emotional content—influence our current behavior directly. However, they can also affect future behavior by serving as the background against which outcomes are evaluated to support learning. Here we will review evidence, from our own work using a Pavlovian over-expectation task as well as from other sources, that the orbitofrontal cortex (OFC) is a critical node in the neural circuit that generates these estimates. Further we will offer the specific hypothesis that degradation of this function secondary to drug-induced changes in the OFC is a critical and possibly reversible part of addiction. Please note it is not our intention to review OFC function broadly or the general role of OFC in addiction; we would refer the reader to other recent reviews

for more information on those topics (Lucantonio *et al*, 2012; Schoenbaum and Shaham, 2007; Stalnaker *et al*, 2015).

THE OFC IS CRITICAL FOR BEHAVIOR AND LEARNING BASED ON IMAGINED OUTCOMES

Much of the work we will describe in detail here involves a task called Pavlovian over-expectation (Rescorla, 1970). This task consists of three phases: conditioning, compound training, and probe testing (Figure 1). In the conditioning phase, rats are trained that several cues predict reward. Subsequently, in the compound training phase, two of the cues are presented together, still followed by the same reward. Typically this results in increased responding at the food cup during the compound cue. This increased responding—termed summation—is thought to reflect a heightened expectation for reward. Importantly this heightened expectation represents a novel prediction. The rats have never before experienced the cues compounded and have never received a double reward, and yet even on the very first exposure to the compound cue, the rats respond more. This behavior is particularly counterintuitive since the compounded cues each predict the same food pellets, in the same number, delivered in the same location. Thus, it is not immediately apparent, based on past experience, that the food pellets should be larger or more plentiful when both cues are presented. Indeed to the extent the compound cue is perceived as a new thing, one would predict less rather than more responding. Yet behavioral summation often does occur, suggesting that the rats jump to the conclusion that the compound cue will be followed by more reward. Furthermore, not only is this novel estimate evident in their

*Correspondence: Dr G Schoenbaum, NIDA Intramural Research Program, 251 Bayview Dr, Baltimore, MD 21224, USA, Tel: +1 443 722 6746, E-mail: geoffrey.schoenbaum@nih.gov

Received 9 May 2016; revised 11 July 2016; accepted 2 August 2016; accepted article preview online 11 August 2016

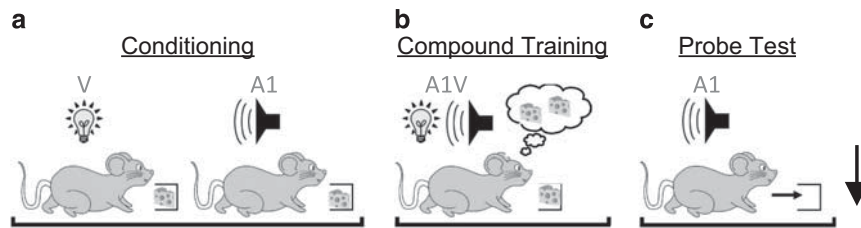


Figure 1 Cartoon illustration of the critical elements in the Pavlovian over-expectation task. The first phase consists of conditioning in which two distinct cues, such as a light, labeled V, and a tone, labeled A1, are paired with reward (represented in the image by a piece of cheese) (a). Subsequently, these cues are presented together in compound training, still followed by the standard reward (b). Finally, the tone is presented alone again in an unrewarded probe test (c).

behavior, it also supports error-based, extinction learning when it goes unmet. This learning is evident in the probe test phase, when the previously compounded cues are presented separately and without reward. Rats suddenly respond less to the cues when they are separated. This decline in responding is formally a form of extinction learning. It shows effects like spontaneous recovery and renewal (Rescorla, 2006, 2007). However, the extinction occurs not due to a change in the amount of reward but rather due to a change in expectation.

While this task involves many functions, it contains within its design a fundamental building block of imagination, namely the need to predict something—in this case a reward—that has never been received. The operation of this imagined prediction is evident in current behavior in the form of summation (ie, increased responding to the compound cue) and future behavior in the form of extinction learning (ie, reduced responding to the compounded cue on the first trial of the probe test). The task design also incorporates several important control behaviors to distinguish general deficits in learning from more specific deficits in generating novel reward predictions. For example, we can assess the acquisition and maintenance of conditioned responding to the individual cues and the extinction learning that occurs because of reward omission during the probe test (not to be confused with the extinction occurring during the compound phase, which is assessed at the start of the probe test). Unlike summation and the resultant learning, these control behaviors can be mediated through direct past experience with reward delivery or omission. Viewed in this manner, Pavlovian over-expectation provides a near ideal vehicle with which to identify neural substrates of imagination, at least with regard to the prediction of rewarding outcomes.

And what better place to start looking than in the OFC? The OFC has long been associated with learning. Indeed reversal deficits have been so long and closely associated with orbitofrontal damage that they have become almost the *sine qua non* of the ‘orbitofrontal syndrome’ (Bechara *et al*, 1997; Fellows and Farah, 2003; Hornak *et al*, 2004; Izquierdo *et al*, 2004; Jones and Mishkin, 1972; Tsuchida *et al*, 2010; Walton *et al*, 2010). At the same time, parallel research over the past 20 years has defined the OFC as critical for signaling associative information and predicting outcomes (Murray *et al*, 2007; Rudebeck and Murray, 2014; Stalnaker *et al*, 2015).

Might the involvement of the OFC in these two functions—outcome prediction and learning—be related to a critical role in imagining novel outcomes? Such a role would explain why the OFC is necessary for changes in conditioned

responding after changes in the value of the predicted reward (Gallagher *et al*, 1999; Gremel and Costa, 2013; Izquierdo *et al*, 2004; Machado and Bachevalier, 2007; Pickens *et al*, 2003, 2005; West *et al*, 2011), while it is not required for changes in behavior initially or for maintaining responding in these same settings. Similarly a role in learning, while multifaceted (Walton *et al*, 2011), might in part reflect the importance of orbitofrontal-dependent predictions in supporting error signaling functions in complicated settings (Schoenbaum *et al*, 2009; Takahashi *et al*, 2011).

To test this hypothesis, we began by inactivating the OFC during the compound training phase of the Pavlovian over-expectation task (Takahashi *et al*, 2009). We reasoned that if the OFC was necessary for generating novel outcome predictions, then if we inactivated it during compound training in the Pavlovian over-expectation task, we should prevent any modest increases in responding to the compounded cues and also abolish any evidence of learning revealed later in the probe test. And in fact this was exactly what we observed (Figure 2). While controls showed no impact of vehicle infusions, rats that received infusions of GABA agonists during compound training did not show any evidence of summation during compound training (Figure 2, middle column), and they failed to reduce responding to the compounded cues when they were presented separately in the probe test (Figure 2, right column). Importantly these effects were observed on a background of completely normal initial conditioning (Figure 3, left column) and normal responding to the individual cues in the other two phases.

Further, we piloted the study using lesions. In this group, orbitofrontal function was presumably abolished throughout training, yet we saw exactly the same effects—normal conditioning, loss of summation in compound training, and no evidence of extinction as a result of over-expectation at the start of the probe test (Takahashi *et al*, 2009 and Supplementary Data). This correspondence between lesion and inactivation rules out any acute off-target effects that can occur with rapid changes in neural processing in a circuit (Otchy *et al*, 2015). It is also particularly interesting to note that the orbitofrontal lesions abolished extinction as a result of over-expectation without affecting extinction occurring as a result of reward omission *within* the probe test. This dissociation demonstrates that the role of the OFC was not generally related to extinction, flexible responding, or response inhibition. Rather the OFC was necessary when the reinforcement histories of individual elements alone were insufficient to support normal behavior or extinction learning.

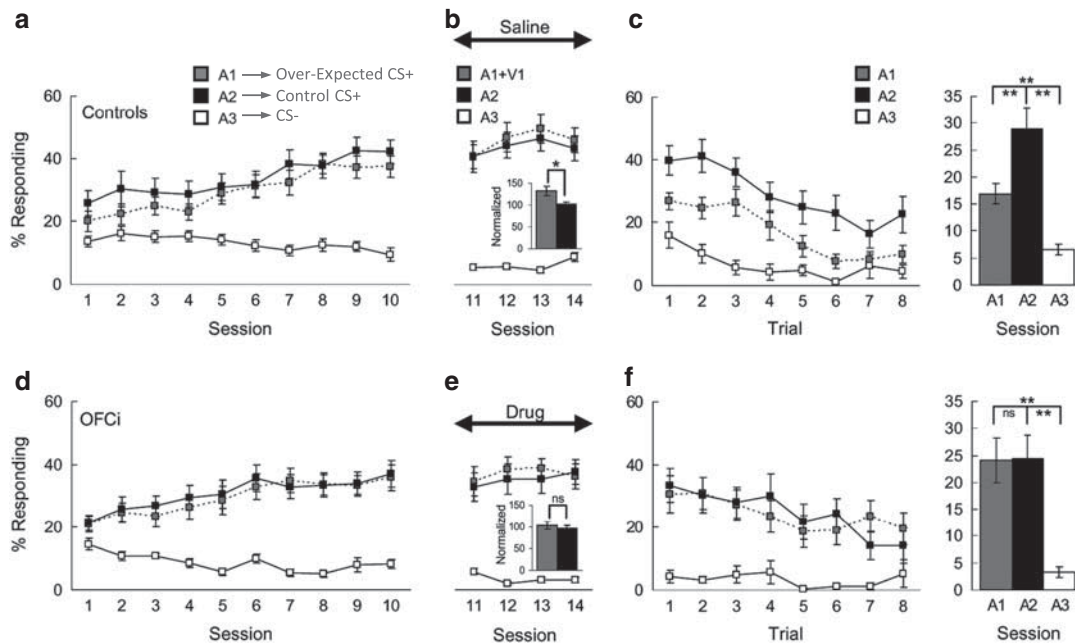


Figure 2 Effect of orbitofrontal inactivation on Pavlovian over-expectation. The task used was similar to that in Figure 1, using V1 (visual stimulus, light) and A1 (auditory stimulus, tone). A1 and V1 both underwent conditioning in which they were independently paired with reward; data from V1 is not shown since it serves only as a tool to induce over-expectation. For direct comparison to A1, we included two additional auditory cues as controls: A2, which was paired with reward but never compounded (Control CS+), and A3, which was presented without reward (CS-). Upper panels show data from the control group, which received vehicle infusions into OFC prior to each compound session; bottom panels show data from the experimental group, which received infusions of GABA agonists into the OFC during compound sessions. (a, d) Conditioned responding to the three auditory cues across 10 days of conditioning in control (a) and experimental (d) groups. There were no differences between groups during conditioning. (b, e) Conditioning responding to cue presentation across 4 days of compound training in control (b) and experimental (e) groups. Bar graphs inset show responding to the compound cue (A1/V1) and A2 normalized to the last day of conditioning. Controls showed an ~30% increase in responding to A1 when it was compounded; inactivated rats showed no change. (c, f) Trial-by-trial (left) and average conditioned responding (right) to the three auditory cues in the unrewarded probe test in control (c) and experimental (f) groups. Controls show a significant reduction in responding to A1 on the first trial; inactivated rats show no change. Gray, black, and white colors indicate responding to A1 or A1/V1, A2, and A3 cues, respectively (* $p < 0.05$, ** $p < 0.01$). Error bars = SEM. NS, nonsignificant; OFC, orbitofrontal cortex; OFCi, OFC inactivation group. Adapted from Takahashi et al (2009).

Of course, while these data confirm the importance of the OFC to behavior and learning based on novel outcome predictions, they do not require that the OFC be critical to integrating predictions, at least not to the complete exclusion of a number of other reasonable competing alternatives. If one looks at even a very simple learning rule, like that proposed by Rescorla and Wagner (1972), there are a number of functions necessary for extinction in this setting, interfering with any one of which might lead to the behavioral effect caused by orbitofrontal lesions or inactivation in our study. For example, although we hypothesized that this effect reflected the loss of the integrative function—the ability to combine predictions of several cues to generate the optimistic estimate that extra reward would follow the compound cue—another equally plausible alternative explanation is that the OFC is the storage site of the critical reinforcement histories of the individual cues, which it then sends to some other area for integration. This would be in accord with the historical emphasis on the OFC as a look-up table of sorts (Rolls, 1996) and perhaps with more recent ideas that the OFC is critical for signaling value (Padoa-Schioppa, 2011; Padoa-Schioppa and Assad, 2006). Or the effect of orbitofrontal inactivation on extinction due to over-expectation might reflect a critical role for this area in signaling prediction errors. This idea would be consistent with claims made in fMRI and a few single-unit recording studies that neural activity in OFC actually directly signals such

teaching signals OFC (Nobre et al, 1999; Schultz and Dickinson, 2000; Sul et al, 2010; Thorpe et al, 1983; Tobler et al, 2006). In the next section, we will describe our efforts to distinguish between these competing hypotheses.

NEURAL ACTIVITY IN THE OFC REFLECTS IMAGINED OUTCOMES WHEN BEHAVIOR AND LEARNING IS DRIVEN BY THESE ESTIMATES

We used single-unit recording to address the question of what function the OFC was contributing to behavior and learning during compound training. In this study (Takahashi et al, 2013), the rats were trained on the same task used in our earlier experiment with only minor changes to allow recording. One change was to compress the transition points between conditioning and compound training into a single session, so that we could track the activity of individual single units across this transition. We reasoned that if orbitofrontal neurons were signaling negative reward prediction errors to drive learning, then this would be evident as differences in firing to the reward following A1 at the end of conditioning, when it was fully expected, vs firing to the same reward following A1 in compound training, when it was over-expected. Similarly, if the orbitofrontal neurons were simply signaling the reinforcement histories of

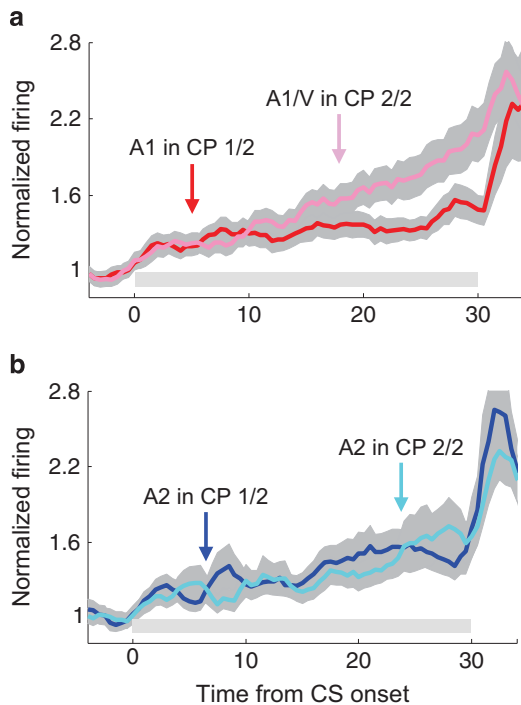


Figure 3 Orbitofrontal neurons signal integrated reward predictions when cues are compounded. Line plots show population responses to A1 (a) and A2 (b) across the transition between conditioning and compound training. There is a sudden increase in neural activity to the compound cue (A1/V) but no change for A2. Dark and light red indicate population response to A1 in the first half of the session and population response to A1/V in the second half, respectively. Dark and light blue indicate population responses to A2 in the first half and second half of the session, respectively. Gray shadings indicate SEM. Gray bars indicate a period of cue presentation. Reward was presented at the end of the cue period. Adapted from Takahashi *et al* (2013).

the individual cues, then there should be relatively little change in firing to A1 itself across this transition, whereas if the neurons were signaling the novel prediction derived from integrating multiple cues, then firing to A1 should increase substantially when it was paired with another reward predicting cue.

Electrode arrays were implanted in the OFC of a group of rats and, after recovery from surgery, they underwent training. We isolated 130 neurons across the transition from conditioning to compound training in sessions that ultimately resulted in evidence of extinction by over-expectation. These neurons showed no evidence of error signaling. Instead they fired at similar rates to reward after A1 alone *vs* reward after A1 and V together. Indeed only 6/130 neurons showed a statistically significant difference in firing to reward in those two conditions, a population that was not above that expected by chance given our analysis. This result is consistent with a number of prior reports showing that while orbitofrontal neurons fire to reward, this firing is generally not affected by whether or not reward was expected (Kennerley *et al*, 2011; McDannald *et al*, 2014; Takahashi *et al*, 2009, 2013).

We also analyzed activity in these 130 neurons during presentation of the cues in order to distinguish activity related to the reinforcement histories *vs* the integration of these predictive histories. Consistent with predictions of the

latter hypothesis, firing to A1 and A2 was similar in this population at the end of conditioning and then increased significantly to A1 when it was compounded without changing to A2. This pattern of firing was evident in the average baseline-normalized firing of these neurons to the cues, which, as illustrated here by the population responses (Figure 3), increased to A1 when it was compounded with V but did not change to A2. Notably this increase to A1V was maximal at the start of the compound training and then declined, both during this session and across subsequent compound training sessions.

To directly test whether the increase in firing in the OFC is in fact necessary for extinction in the over-expectation task, as it should be if it is responsible for the novel reward prediction, we conducted a final experiment in which we used optogenetics to manipulate the activity of neurons in the OFC in a temporally specific manner. Rats received infusions into OFC of a viral construct containing either eYFP or NpHR linked to a CAMKII promoter. Fiber probes were implanted so that we could deliver light over the infusion sites. After 2 weeks recovery time to allow viral expression, these rats were then trained in the Pavlovian over-expectation task. Green light was delivered into OFC in order to inhibit pyramidal neurons during the compound cue, when our prior recording results suggested activity would increase. The same rats were subsequently retrained and light was delivered during the inter-trial interval as a control for the specificity of the effect. Consistent with the proposal that increased neural activity in the OFC in response to the compound cue is in fact necessary for learning, NpHR, but not eYFP, rats failed to learn when light was delivered during the compound cue, but both learned normally when light was delivered in the inter-trial interval (Figure 4).

Overall these data are in accord with the emerging consensus that the OFC is critical for behavior that goes beyond simple reinforcement history (Stalnaker *et al*, 2015; Wilson *et al*, 2014). Such behavior is variously described as goal-directed or model-based (Daw *et al*, 2005; Dickinson and Balleine, 1994), although these terms are typically used to describe instrumental behavior, whereas orbitofrontal function does not seem to be constrained to instrumental activities. Indeed if anything, it is more closely associated with Pavlovian behavior. Ignoring this distinction, the behaviors that fall into this category have in common that they require the mental simulation of possible outcomes given an associative structure reflecting the current cause-and-effect relationships within the environment. When value-based behavior depends on this sort of simulative process, it seems to require the OFC across species. This explanation applies in a variety of situations in which specific behaviors seem to be sensitive to OFC manipulations, including summation during over-expectation as shown here (Takahashi *et al*, 2009, 2013), Pavlovian and in some cases instrumental devaluation (Gallagher *et al*, 1999; Gremel and Costa, 2013; Izquierdo *et al*, 2004; Machado and Bachevalier, 2007; Pickens *et al*, 2003, 2005; West *et al*, 2011), sensory preconditioning (Jones *et al*, 2012), as well as less well-controlled behaviors such as delayed discounting (Mobini *et al*, 2002; Winstanley *et al*, 2004) and even economic choice and regret (Camille *et al*, 2004, 2011). The same process can also explain the involvement of the OFC in

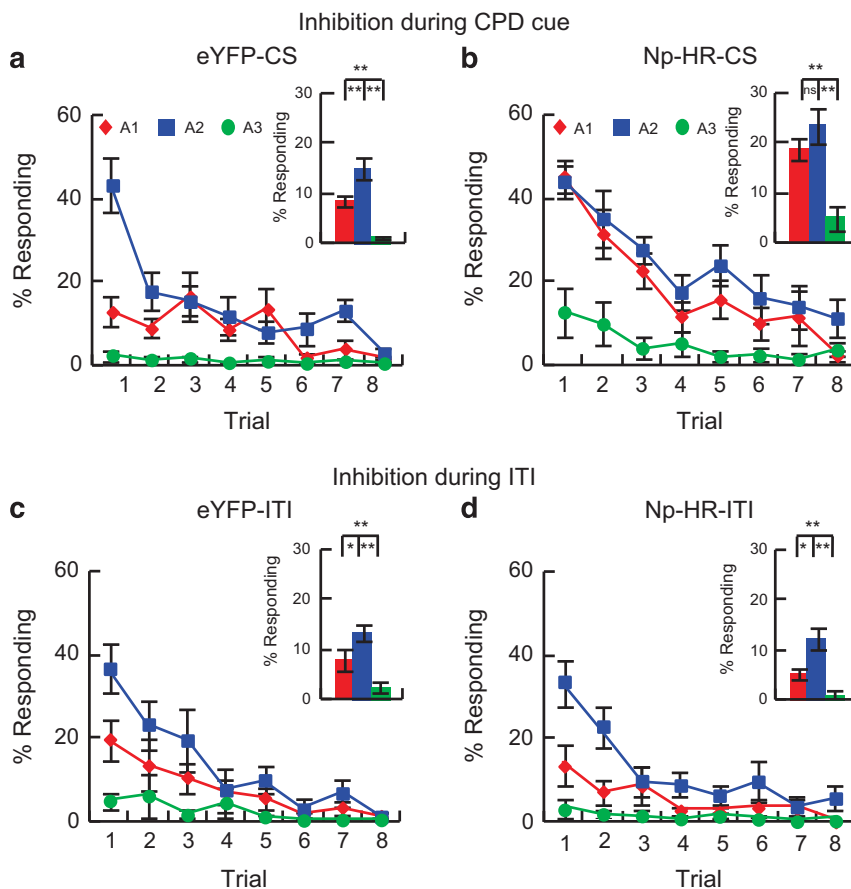


Figure 4 Inhibition of the OFC during presentation of the compound cues prevents learning as a result of Pavlovian over-expectation. (a–d) Conditioned responding to the three auditory cues (A1, A2, and A3) during the probe test at the end of the over-expectation in control (eYFP) (a, c) and experimental (NpHR) (b, d) groups. Top panels show responding after inhibition during the compound (CPD) cue. Bottom panels show responding after retraining and inhibition during the inter-trial intervals (ITI). The line plots show responding across eight trials, and bar graphs show average responding of eight trials. Red, blue, and green indicate A1, A2, and A3, respectively. Controls showed evidence of learning from over-expectation on the first trial in both cases (CS and ITI), whereas the experimental group showed evidence of learning only after ITI inhibition. Inhibition of OFC during the compound cue prevented the learning induced by over-expectation. * $p < 0.05$, ** $p < 0.01$. Error bars = SEM. eYFP, enhanced yellow fluorescent protein; Np-HR, halorhodopsin; OFC, orbitofrontal cortex. Adapted from Takahashi *et al* (2013).

behaviors that seem to require the ability to predict information about outcomes aside from value (McDannald *et al*, 2011; Ostlund and Balleine, 2007).

The data described above add to this literature a clear demonstration that this same function also contributes to the predictions that are used to modulate learning. Further they show that neural activity in the OFC, at the time that behavior and learning requires the ability to make such novel outcome predictions, reflects this function rather than several alternatives.

WHAT DOES THIS HAVE TO DO WITH ADDICTION?

Addiction is fundamentally a disorder of behavioral control. This is evident from a consideration of the diagnostic criteria for substance dependence in the DSM (APA, 2000). While historically these have included phenomena like tolerance and withdrawal, which are largely related to physiological effects of the drug, increasingly these criteria are based on an inability to control drug use in the face of a variety of adverse

consequences. While these symptoms can be conceptualized as an overpowering desire for the drug, they might equally well be explained as an inability to apply knowledge of (or perhaps even learn about to begin with) the consequences that often follow drug use. While these two factors are not mutually exclusive, framing addiction as involving the latter is important because a failure to use—or even learn about appropriately—non-drug consequences is often not considered as the core problem. If it is, then the observation that addictive can cause long-lasting changes in orbitofrontal structure and function raises the possibility that orbitofrontal dysfunction may play a critical role in addiction (Jentsch and Taylor, 1999; Lucantonio *et al*, 2012; Volkow and Fowler, 2000).

DRUG USE DISRUPTS ORBITOFONTAL-DEPENDENT BEHAVIOR AND LEARNING BASED ON IMAGINED OUTCOMES

If orbitofrontal function is compromised by exposure to addictive drugs, then this dysfunction should be evident in

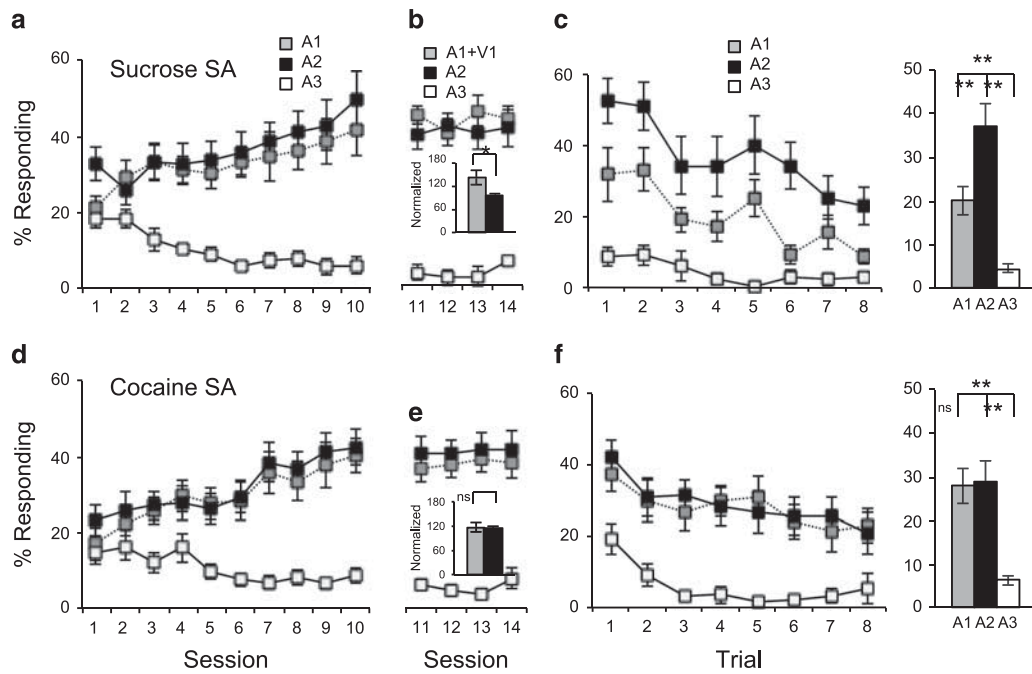


Figure 5 Effect of cocaine self-administration on Pavlovian over-expectation. The task used was identical to that used in Figure 2. Upper panels show data from the control group (Sucrose SA), which self-administered sucrose until approximately 3 weeks prior to the start of over-expectation training; bottom panels show data from the experimental group (Cocaine SA), which self-administered cocaine. (a, d) Conditioned responding to the three auditory cues across 10 days of conditioning. There were no differences between groups. (b, e) Conditioned responding to 4 days of compound training. Bar graphs inset show responding to A1/V1 and A2 normalized to the last day of conditioning. Controls showed an ~30% increase in responding to A1 when it was compounded; cocaine rats showed no change. (c, f) Trial-by-trial (left panel) and average conditioned responding (right panel) to the three auditory cues in the unrewarded probe test. Controls show a significant reduction in responding to A1 on the first trial; cocaine rats show no change. Gray, black, and white colors indicate responding to A1 or A1/V1, A2, and A3 cues, respectively (* $p < 0.05$, ** $p < 0.01$). Error bars = SEM. NS, nonsignificant; OFC, orbitofrontal cortex; SA, self-administration. Adapted from Lucantonio *et al* (2014a).

experimental settings that assess orbitofrontal function, even when no drug is present. Indeed to the extent that orbitofrontal dysfunction plays a special role in any of the more pernicious, long-term aspects of addiction, such as craving and relapse, these effects should be evident in non-drug settings long after cessation of drug use. There is now substantial evidence in both clinical populations and experimental models that this is the case.

For example, many labs have shown that reversal learning, long associated with orbitofrontal function, is impaired in addicts and in monkeys and rats that have experience with addictive drugs (Bechara and Damasio, 2002; Bechara *et al*, 2001; Calu *et al*, 2007; Ersche *et al*, 2008; Izquierdo *et al*, 2009; Krueger *et al*, 2009; Porter *et al*, 2011; Schoenbaum *et al*, 2004; Taylor and Jentsch, 2001). The precise etiology is difficult to disentangle in patients, but in the animal studies it is clear that these deficits are secondary to drug. Deficits in reversal learning are present whether these drugs are administered passively or taken actively, and they do not appear to require extended or special patterns of access. As we will describe, this is also generally the case for deficits in more specific orbitofrontal-dependent tasks. This ubiquity suggests that, while not a marker of addiction, orbitofrontal dysfunction may be a fundamental and unavoidable effect of drug use, present early and throughout what is a multifactorial and evolving process. In this regard, it is worth noting that the reversal deficits and associated changes in information coding in the orbitofrontal cortex have been shown to persist for months after the end of drug use,

potentially implicating orbitofrontal dysfunction in difficult to treat enduring phenomena such as craving and relapse.

A more restricted set of studies has shown deficits in much better controlled orbitofrontal-dependent behavioral tasks. For example, passive exposure to cocaine or active use of amphetamine impairs changes in learned behavior that normally occur as a result of devaluation of the reinforcer (LeBlanc *et al*, 2013; Nelson and Killcross, 2006; Schoenbaum and Setlow, 2005). These deficits are similar to those observed after damage or manipulations of a circuit that typically includes the OFC across species. Deficits have also been observed in other orbitofrontal-dependent behaviors in rats with a history of drug use, such as sensory preconditioning (Wied *et al*, 2013) and Pavlovian-to-instrumental transfer (LeBlanc *et al*, 2012; Wyvell and Berridge, 2001). Notably these tasks have in common that they require information about outcomes to be inferred or simulated as in the Pavlovian over-expectation task described above.

Accordingly, we have found recently (Lucantonio *et al*, 2014b) that rats trained to self-administer cocaine (2 weeks, 3h per day, no cues) subsequently show deficits in Pavlovian over-expectation (Figure 5). Like rats with orbitofrontal lesions or in whom the OFC was inactivated pharmacologically or optogenetically, cocaine-experienced rats failed to show any evidence of summation when the cues were compounded (Figure 5, middle column) or any sign of extinction as a result of over-expectation at the start of the probe test (Figure 5, right column). Importantly, these rats

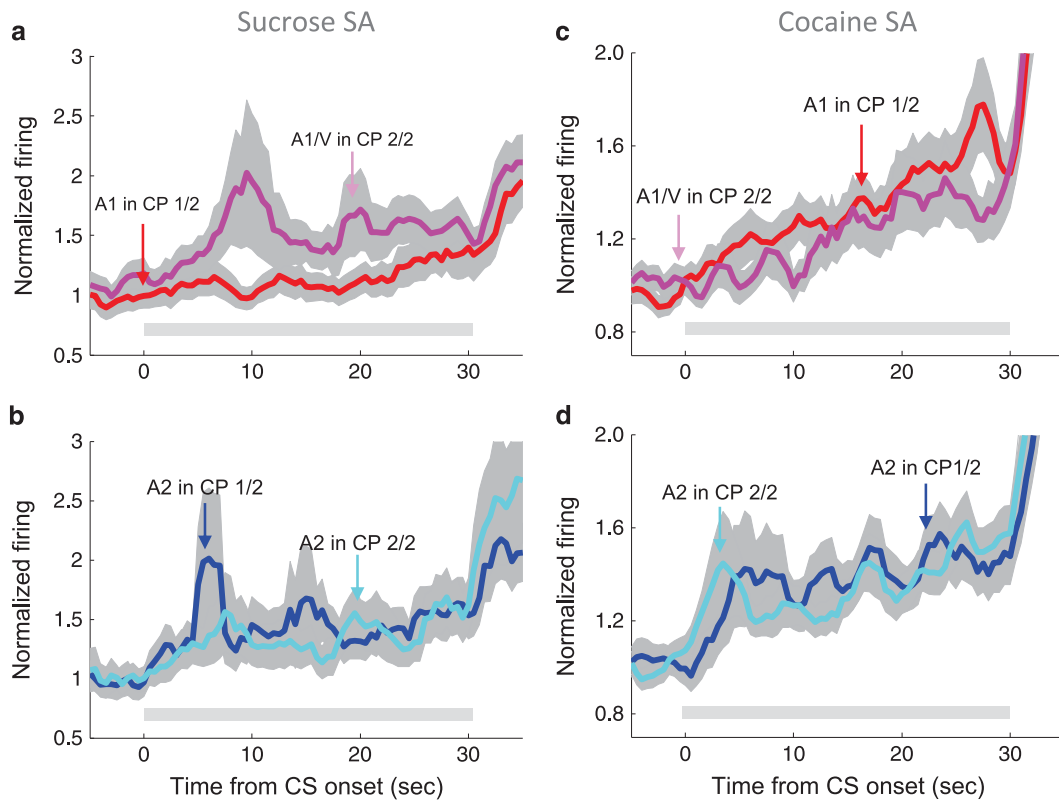


Figure 6 Effect of sucrose and cocaine self-administration on reward predictions when cues are compounded in orbitofrontal cortex (OFC). Line plots show population responses to A1 and A2 across the transition between conditioning and compound training in sucrose (c, d) and cocaine (a, b) groups. There is a sudden increase in neural activity to the compound cue (A1/V) in sucrose but not cocaine-trained rats. Conventions as in Figure 3. Adapted from Lucantonio *et al* (2014b).

conditioned normally (Figure 5, left column), maintained normal levels of conditioned responding during compound training (Figure 5, middle column), and extinguished responding in the face of reward omission *within* the probe test (Figure 5, right column). Thus, cocaine self-administration produced an effect on the behavior found to be orbitofrontal-dependent in our earlier work. We have also found similar effects following heroin use (Lucantonio *et al*, 2014a).

DRUG USE REDUCES ORBITOFRONTAL EXCITABILITY AND DISRUPTS NEURAL CORRELATES OF IMAGINED OUTCOMES

Self-administration of cocaine is neither a regionally nor temporally specific manipulation. Given the complexity of the Pavlovian over-expectation task and its dependence on other brain areas, the effects of cocaine self-administration, though similar to those caused by OFC manipulations, may actually reflect changes in other areas. To address this question, we have turned to electrophysiology to look more closely at the changes in synaptic function and information coding in orbitofrontal neurons occurring as a result of the self-administration.

Whole-cell recordings from orbitofrontal neurons in brain slices from rats characterized as impaired on the over-expectation task after self-administration of cocaine showed a specific reduction in synaptic efficacy relative to sucrose-

trained rats, as indexed the frequency of glutamate-mediated miniature excitatory postsynaptic currents (Lucantonio *et al*, 2014b). These changes in excitability are similar to those reported in mice in medial prefrontal areas (Chen *et al*, 2013).

Changes in excitability could reduce the ability of the orbitofrontal network to respond flexibly to sudden changes in afferent input, as is required to produce the heightened expectations upon which over-expectation is founded. To test this question, we again trained rats to self-administer either cocaine or sucrose and then tested them on Pavlovian over-expectation, this time while recording single unit activity in the OFC (Lucantonio *et al*, 2014b). We recorded activity from 58 neurons in sucrose-trained rats across the transition between conditioning and compound training. As a group, these rats showed successful learning in the task, and their neural activity across the transition was similar to what we observed in our prior study. Specifically, firing to A1 and A2 was similar at the end of conditioning and then increased significantly to A1 when it was compounded without changing to A2 (Figure 6, Sucrose SA).

This pattern contrasted sharply with the pattern observed in the activity of 131 neurons recorded in cocaine-trained rats. These rats failed to show any evidence of extinction as a result of over-expectation in the probe test, and neural activity across the transition point exhibited a corresponding failure to increase to the compound cues (Figure 6, Cocaine SA). This change in firing occurred even though the increase in firing from baseline during A1 and A2 at the end of conditioning was similar to that in controls. Indeed the

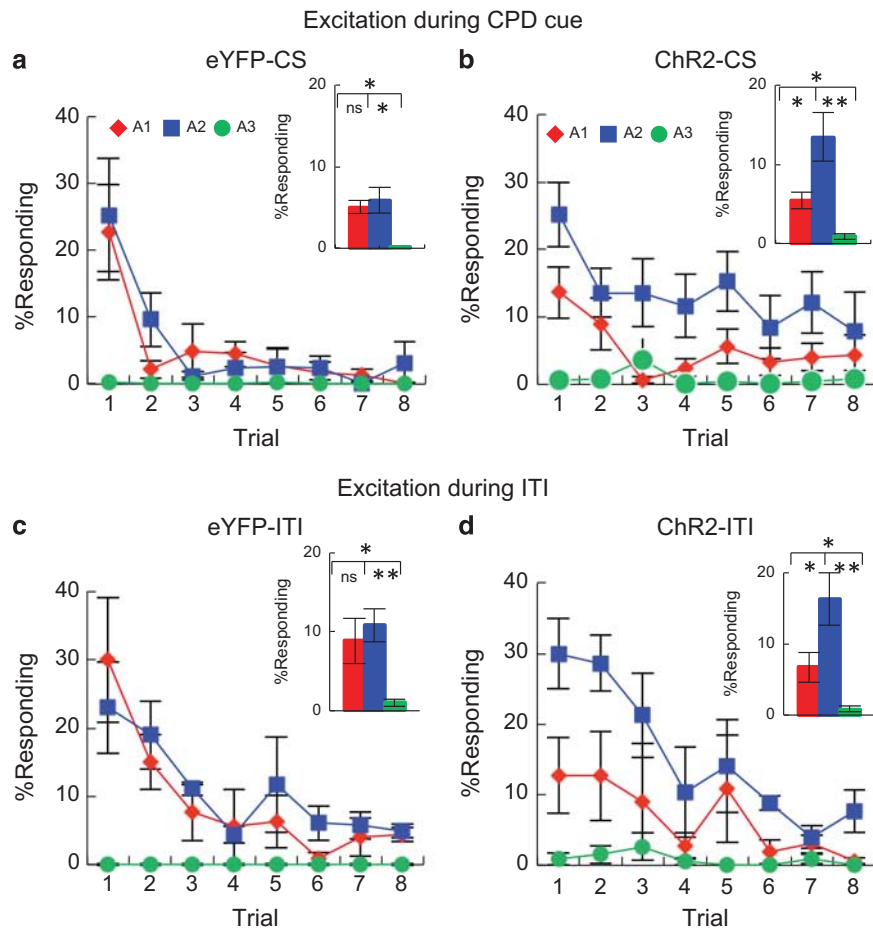


Figure 7 Excitation of OFC during presentation of the compound cues fixes cocaine-induced deficits in Pavlovian over-expectation. (a–d) Conditioned responding to the three auditory cues (A1, A2, and A3) during the probe test at the end of the over-expectation in control (eYFP) (a, c) and experimental (ChR2) (b, d) groups. Top panels show responding after excitation during the compound (CPD) cue. Bottom panels show responding after retraining and excitation during the inter-trial intervals (ITI). The line plots show responding across eight trials, and bar graphs show average responding of eight trials. Red, blue, and green indicate A1, A2, and A3, respectively. All rats had prior experience self-administering cocaine. Controls (eYFP) showed the same learning deficit observed in prior cocaine-experienced rats, whereas the experimental group (ChR2) showed evidence of restored learning after excitation during the cue or during the ITI. * $p < 0.05$. ** $p < 0.01$. Error bars = SEM. ChR2, channelrhodopsin; eYFP, enhanced yellow fluorescent protein; NS, not significant; OFC, orbitofrontal cortex. Adapted from Lucantonio *et al* (2014b).

proportions of cue-selective neurons (A1 or A2 selective neurons) and the increase in the proportions of these neurons across conditioning were not affected by cocaine self-administration. Thus orbitofrontal neurons recorded in rats that had self-administered cocaine were able to change firing properties to the cues as a result of pairing with reward normally. However the same neurons were unable to react on the fly, to increase that firing when the cues were presented together in the compound phase. This inability to integrate the meaning of the previously conditioned cues is precisely the critical function that underlies behavioral summation and over-expectation.

CAN WE FIX IT?

If cocaine use causes long-lasting changes in orbitofrontal function because it has a general effect on diminishing the efficacy of synaptic interactions within the OFC, then it might be possible to overcome these effects by transiently activating the orbitofrontal network. To test this, we

conducted a final experiment in which we used optogenetics to manipulate the activity of neurons in the OFC in a temporally specific manner. Cocaine-experienced rats received infusions into OFC of a viral construct containing either eYFP (control group) or ChR2 (a light-gated cation channel, experimental group) linked to a CAMKII promoter, in order to mainly target pyramidal neurons. Fiber probes were implanted so that we could deliver light over the infusion sites. After several weeks recovery time to allow viral expression, these rats were then trained in the Pavlovian over-expectation task. Blue light was delivered into OFC specifically during the compound cue to activate pyramidal neurons, when our recording results suggested the normal increase in activity was missing in the cocaine-trained rats. Consistent with the proposal that this loss in activity is in fact responsible for the cocaine-induced behavioral deficit, ChR2 (but not eYFP) rats recovered normal learning (Figure 7).

Though restoration of learning by direct stimulation of OFC may seem somewhat trivial, this effect is important for at least two reasons. First it shows that the downstream mechanism where by orbitofrontal output leads to

learning—perhaps the VTA dopamine neurons (Chang *et al*, 2016; Takahashi *et al*, 2009,2011)—is in fact largely intact, since if it was not, then artificially correcting the problem in the OFC would not fix the problem. Second it suggests that the input to the OFC—firing reflecting the identity and possibly the meaning of the cues—is also largely functional. Randomly activating a subset of orbitofrontal neurons would not be expected to cause extinction of a particular cue. Although we cannot rule out some compensatory mechanism, the restoration of this very specific behavioral effect is most likely to occur only if our manipulation somehow reinstates a reasonable facsimile of the normal firing pattern across the population. This can only really happen if the afferent input is largely intact and what we are doing by stimulating is simply to augment this normal pattern.

But more interesting is what happened when we retrained these rats and repeated the experiment, this time activating the OFC transiently during the inter-trial intervals. Remarkably this again resulted in normal learning in the ChR2 (but not eYFP) rats (Figure 7). Whether normal learning persisted as a result of the prior stimulation, given over a week earlier, or was restored due to the stimulation in the inter-trial intervals, this result shows that the recovery of normal orbitofrontal function persists beyond the period of stimulation. This prolonged effect is critical for any practical application of this manipulation, since we are unlikely to be able to predict the need for orbitofrontal function in any real-world situation. Happily, this final result holds out the possibility that, if addicted individuals suffer from OFC dysfunction as several studies suggest, it might be possible to restore it to some degree by briefly activating this region (using transcranial magnetic stimulation; Terraneo *et al*, 2016), daily or weekly or maybe only once, and thereby allow the subject to bring back on line the sort of judgment and imagination of outcomes that is normally mediated by the OFC. In this regard, it is worth noting that there are reasonable anatomical and functional homologies between the lateral orbitofrontal area that is the target of our manipulations and the lateral orbital network that has been identified in primates (Ongur and Price, 2000; Stalnaker *et al*, 2015; Wallis, 2012); this lateral network might be specifically targeted.

This idea could of course be easily tested. The first step would be to develop a human version of the over-expectation task, one that captures the inherent Pavlovian nature of the task we have used in our studies. Some labs use procedures already that would be a good model for this (Johnsrude *et al*, 2000). The second step would be to test whether addicts show impairments on this task. There is good evidence for impairments in orbitofrontal-dependent functions in addicts, but knowing that they show a deficit in this specific task would be useful, since this task provides a direct assessment of the ability to mentally simulate or imagine novel outcomes to both guide behavior and influence learning. We would predict that any laboratory impairment in the ability to imagine novel outcomes would be predictive of their subsequent vulnerability to the poor judgment (abnormal learning?) that leads to relapse to drug seeking. The third step would be to then test the effects of stimulation, first on task performance and then on clinical outcomes. Again we would predict that to the extent stimulation can correct the laboratory impairment, it will also lead to a reduced

vulnerability to relapse, since it essentially would reflect the restoration of the normal ability to imagine and mentally simulate consequences when considering whether or not to engage in drug seeking and drug use.

FUNDING AND DISCLOSURE

This work was supported by the Intramural Research Program at the National Institute on Drug Abuse. The authors declare no conflict of interest.

ACKNOWLEDGMENTS

The opinions expressed in this article are the authors' own and do not reflect the view of the NIH/DHHS.

REFERENCES

- American Psychiatric Association. Diagnostic and Statistical Manual of Mental Disorders, 5th edn. American Psychiatric Association: Arlington, VA, 2013.
- Bechara A, Damasio H (2002). Decision-making and addiction (part I): impaired activation of somatic states in substance dependent individuals when pondering decisions with negative future consequences. *Neuropsychologia* **40**: 1675–1689.
- Bechara A, Damasio H, Tranel D, Damasio AR (1997). Deciding advantageously before knowing the advantageous strategy. *Science* **275**: 1293–1294.
- Bechara A, Dolan S, Denburg N, Hindes A, Andersen SW, Nathan PE (2001). Decision-making deficits, linked to a dysfunctional ventromedial prefrontal cortex, revealed in alcohol and stimulant abusers. *Neuropsychologia* **39**: 376–389.
- Calu DJ, Stalnaker TA, Franz TM, Singh T, Shaham Y, Schoenbaum G (2007). Withdrawal from cocaine self-administration produces long-lasting deficits in orbitofrontal-dependent reversal learning in rats. *Learn Mem* **14**: 325–328.
- Camille N, Coricelli G, Sallet J, Pradat-Diehl P, Duhamel J-R, Sirigu A (2004). The involvement of the orbitofrontal cortex in the experience of regret. *Science* **304**: 1168–1170.
- Camille N, Griffiths CA, Vo K, Fellows LK, Kable JW (2011). Ventromedial frontal lobe damage disrupts value maximization in humans. *J Neurosci* **31**: 7527–7532.
- Chang CY, Esber GR, Marrero-Garcia Y, Yau H-J, Bonci A, Schoenbaum G (2016). Brief optogenetic inhibition of VTA dopamine neurons mimics the effects of endogenous negative prediction errors during Pavlovian over-expectation. *Nat Neurosci* **19**: 111–116.
- Chen BT, Yau H-J, Hatch C, Kusumoto-Yoshida I, Cho SL, Hopf FW *et al* (2013). Rescuing cocaine-induced prefrontal cortex hypoactivity prevents compulsive cocaine seeking. *Nature* **496**: 359–362.
- Daw ND, Niv Y, Dayan P (2005). Uncertainty-based competition between prefrontal and dorsolateral striatal systems for behavioral control. *Nat Neurosci* **8**: 1704–1711.
- Dickinson A, Balleine BW (1994). Motivational control of goal-directed action. *Anim Learn Behav* **22**: 1–18.
- Ersche KD, Roiser JP, Robbins TW, Sahakian BJ (2008). Chronic cocaine but not chronic amphetamine use is associated with perseverative responding in humans. *Psychopharmacology* **197**: 421–431.
- Fellows LK, Farah MJ (2003). Ventromedial frontal cortex mediates affective shifting in humans: evidence from a reversal learning paradigm. *Brain* **126**: 1830–1837.

- Gallagher M, McMahan RW, Schoenbaum G (1999). Orbitofrontal cortex and representation of incentive value in associative learning. *J Neurosci* **19**: 6610–6614.
- Gremel CM, Costa RM (2013). Orbitofrontal and striatal circuits dynamically encode the shift between goal-directed and habitual actions. *Nat Commun* **4**: 2264.
- Hornak J, O'Doherty J, Bramham J, Rolls ET, Morris RG, Bullock PR et al (2004). Reward-related reversal learning after surgical excisions in orbito-frontal or dorsolateral prefrontal cortex in humans. *J Cogn Neurosci* **16**: 463–478.
- Izquierdo A, Belcher AM, Scott L, Cazares VA, Chen J, O'Dell SJ et al (2009). Reversal-specific learning impairments after a binge regimen of methamphetamine in rats: possible involvement of striatal dopamine. *Neuropsychopharmacology* **35**: 505–514.
- Izquierdo AD, Suda RK, Murray EA (2004). Bilateral orbital prefrontal cortex lesions in rhesus monkeys disrupt choices guided by both reward value and reward contingency. *Journal of Neuroscience* **24**: 7540–7548.
- Jentsch JD, Taylor JR (1999). Impulsivity resulting from frontostriatal dysfunction in drug abuse: implications for the control of behavior by reward-related stimuli. *Psychopharmacology* **146**: 373–390.
- Johnsrude IS, Owen AM, White NM, Zhao WV, Bohbot V (2000). Impaired preference conditioning after anterior temporal lobe resection in humans. *J Neurosci* **20**: 2649–2656.
- Jones B, Mishkin M (1972). Limbic lesions and the problem of stimulus-reinforcement associations. *Exp Neurol* **36**: 362–377.
- Jones JL, Esber GR, McDannald MA, Gruber AJ, Hernandez G, Mirenzi A et al (2012). Orbitofrontal cortex supports behavior and learning using inferred but not cached values. *Science* **338**: 953–956.
- Kennerley SW, Behrens TE, Wallis JD (2011). Double dissociation of value computations in orbitofrontal and anterior cingulate neurons. *Nat Neurosci* **14**: 1581–1589.
- Krueger DD, Howell JL, Oo H, Olausson P, Taylor JR, Nairn AC (2009). Prior chronic cocaine exposure in mice induces persistent alterations in cognitive function. *Behav Pharmacol* **20**: 695–704.
- LeBlanc KH, Maidment NT, Ostlund SB (2013). Repeated cocaine exposure facilitates the expression of incentive motivation and induces habitual control in rats. *PLoS One* **8**: e61355.
- LeBlanc KH, Ostlund SB, Maidment NT (2012). Pavlovian-to-instrumental transfer in cocaine seeking rats. *Behav Neurosci* **126**: 681–689.
- Lucantonio F, Kambhampati S, Haney RZ, Atalayer D, Rowland NE, Shaham Y et al (2014a). Effects of prior cocaine versus morphine or heroin self-administration on extinction learning driven by overexpectation versus omission of reward. *Biol Psychiatry* **77**: 912–920.
- Lucantonio F, Stalnaker TA, Shaham Y, Niv Y, Schoenbaum G (2012). The impact of orbitofrontal dysfunction on cocaine addiction. *Nat Neurosci* **15**: 358–366.
- Lucantonio F, Takahashi YK, Hoffman AF, Chang CY, Bali-Chaudhary S, Shaham Y et al (2014b). Orbitofrontal activation restores insight lost after cocaine use. *Nat Neurosci* **17**: 1092–1099.
- Machado CJ, Bachevalier J (2007). The effects of selective amygdala, orbital frontal cortex or hippocampal formation lesions on reward assessment in nonhuman primates. *Eur J Neurosci* **25**: 2885–2904.
- McDannald MA, Esber GR, Wegener MA, Wied HM, Tzu-Lan L, Stalnaker TA et al (2014). Orbitofrontal neurons acquire responses to 'valueless' Pavlovian cues during unblocking. *eLIFE* **3**: e02653.
- McDannald MA, Lucantonio F, Burke KA, Niv Y, Schoenbaum G (2011). Ventral striatum and orbitofrontal cortex are both required for model-based, but not model-free, reinforcement learning. *J Neurosci* **31**: 2700–2705.
- Mobini S, Body S, Ho M-Y, Bradshaw CM, Szabadi E, Deakin JFW et al (2002). Effects of lesions of the orbitofrontal cortex on sensitivity to delayed and probabilistic reinforcement. *Psychopharmacology* **160**: 290–298.
- Murray EA, O'Doherty J, Schoenbaum G (2007). What we know and do not know about the functions of the orbitofrontal cortex after 20 years of cross-species studies. *J Neurosci* **27**: 8166–8169.
- Nelson A, Killcross S (2006). Amphetamine exposure enhances habit formation. *J Neurosci* **26**: 3805–3812.
- Nobre AC, Coull JT, Frith CD, Mesulam MM (1999). Orbitofrontal cortex is activated during breaches of expectation in tasks of visual attention. *Nat Neurosci* **2**: 11–12.
- Ongur D, Price JL (2000). The organization of networks within the orbital and medial prefrontal cortex of rats, monkeys and humans. *Cerebr Cortex* **10**: 206–219.
- Ostlund SB, Balleine BW (2007). Orbitofrontal cortex mediates outcome encoding in Pavlovian but not instrumental learning. *J Neurosci* **27**: 4819–4825.
- Otchy TM, Wolff SBE, Rhee JY, Pehlevan C, Kawai R, Kempf A et al (2015). Acute off-target effects of neural circuit manipulations. *Nature* **528**: 358–363.
- Padoa-Schioppa C (2011). Neurobiology of economic choice: a goods-based model. *Annu Rev Neurosci* **34**: 333–359.
- Padoa-Schioppa C, Assad JA (2006). Neurons in orbitofrontal cortex encode economic value. *Nature* **441**: 223–226.
- Pickens CL, Saddoris MP, Gallagher M, Holland PC (2005). Orbitofrontal lesions impair use of cue-outcome associations in a devaluation task. *Behav Neurosci* **119**: 317–322.
- Pickens CL, Setlow B, Saddoris MP, Gallagher M, Holland PC, Schoenbaum G (2003). Different roles for orbitofrontal cortex and basolateral amygdala in a reinforcer devaluation task. *J Neurosci* **23**: 11078–11084.
- Porter JN, Olsen AS, Gurnsey K, Dugan BP, Jedema HP, Bradberry CW (2011). Chronic cocaine self-administration in Rhesus monkeys: impact on associative learning, cognitive control, and working memory. *J Neurosci* **31**: 4926–4934.
- Rescorla RA (1970). Reduction in effectiveness of reinforcement after prior extinction conditioning. *Learn Motiv* **1**: 372–381.
- Rescorla RA (2006). Spontaneous recovery from overexpectation. *Learn Behav* **34**: 13–20.
- Rescorla RA (2007). Renewal from overexpectation. *Learn Behav* **35**: 19–26.
- Rescorla RA, Wagner AR. A theory of Pavlovian conditioning: Variations in the effectiveness of reinforcement and non-reinforcement. In Black AH, Prokasy WF (eds). *Classical conditioning II. Current research and theory*. Appleton-Century-Crofts: New York, 1972, pp 64–99.
- Rolls ET (1996). The orbitofrontal cortex. *Philos Trans R Soc London B* **351**: 1433–1443.
- Rudebeck PH, Murray EA (2014). The orbitofrontal oracle: cortical mechanisms for the prediction and evaluation of specific behavioral outcomes. *Neuron* **84**: 1143–1156.
- Schoenbaum G, Roesch MR, Stalnaker TA, Takahashi YK (2009). A new perspective on the role of the orbitofrontal cortex in adaptive behaviour. *Nat Rev Neurosci* **10**: 885–892.
- Schoenbaum G, Saddoris MP, Ramus SJ, Shaham Y, Setlow B (2004). Cocaine-experienced rats exhibit learning deficits in a task sensitive to orbitofrontal cortex lesions. *Eur J Neurosci* **19**(7): 1997–2002.
- Schoenbaum G, Setlow B (2005). Cocaine makes actions insensitive to outcomes but not extinction: implications for altered orbitofrontal-amygdalar function. *Cerebr Cortex* **15**: 1162–1169.
- Schoenbaum G, Shaham Y (2007). The role of orbitofrontal cortex in drug addiction: a review of preclinical studies. *Biol Psychiatry* **63**: 256–262.
- Schultz W, Dickinson A (2000). Neuronal coding of prediction errors. *Annu Rev Neurosci* **23**: 473–500.
- Stalnaker TA, Cooch NK, Schoenbaum G (2015). What the orbitofrontal cortex does not do. *Nat Neurosci* **18**: 620–627.

- Sul JH, Kim H, Huh N, Lee D, Jung MW (2010). Distinct roles of rodent orbitofrontal and medial prefrontal cortex in decision making. *Neuron* **66**: 449–460.
- Takahashi Y, Roesch MR, Stalnaker TA, Haney RZ, Calu DJ, Taylor AR et al (2009). The orbitofrontal cortex and ventral tegmental area are necessary for learning from unexpected outcomes. *Neuron* **62**: 269–280.
- Takahashi YK, Chang CY, Lucantonio F, Haney RZ, Berg BA, Yau H-J et al (2013). Neural estimates of imagined outcomes in the orbitofrontal cortex drive behavior and learning. *Neuron* **80**: 507–518.
- Takahashi YK, Roesch MR, Wilson RC, Toreson K, O'Donnell P, Niv Y et al (2011). Expectancy-related changes in firing of dopamine neurons depend on orbitofrontal cortex. *Nat Neurosci* **14**: 1590–1597.
- Taylor JR, Jentsch JD (2001). Repeated intermittent administration of psychomotor stimulant drugs alters the acquisition of Pavlovian approach behavior in rats: differential effects of cocaine, d-amphetamine and 3,4-methylenedioxymethamphetamine ("ecstasy"). *Biol Psychiatry* **50**: 137–143.
- Terraneo A, Leggio L, Saladini M, Ermani M, Bonci A, Gallimberti L (2016). Transcranial magnetic stimulation of dorsolateral prefrontal cortex reduces cocaine use: a pilot study. *Eur J Neuropsychopharmacol* **26**: 37–44.
- Thorpe SJ, Rolls ET, Maddison S (1983). The orbitofrontal cortex: neuronal activity in the behaving monkey. *Exp Brain Res* **49**: 93–115.
- Tobler PN, O'Doherty J, Dolan RJ, Schultz W (2006). Human neural learning depends on reward prediction errors in the blocking paradigm. *J Neurophysiol* **95**: 301–310.
- Tsuchida A, Doll BB, Fellows LK (2010). Beyond reversal: a critical role for human orbitofrontal cortex in flexible learning from probabilistic feedback. *J Neurosci* **30**: 16868–16875.
- Volkow ND, Fowler JS (2000). Addiction, a disease of compulsion and drive: involvement of orbitofrontal cortex. *Cerebr Cortex* **10**: 318–325.
- Wallis JD (2012). Cross-species studies of orbitofrontal cortex and value-based decision-making. *Nat Neurosci* **15**: 13–19.
- Walton ME, Behrens TE, Noonan MP, Rushworth MF (2011). Giving credit where credit is due: orbitofrontal cortex and valuation in an uncertain world. *Ann NY Acad Sci* **1239**: 14–24.
- Walton ME, Behrens TEJ, Buckley MJ, Rudebeck PH, Rushworth MF (2010). Separable learning systems in the macaque brain and the role of the orbitofrontal cortex in contingent learning. *Neuron* **65**: 927–939.
- West EA, DesJardin JT, Gale K, Malkova L (2011). Transient inactivation of orbitofrontal cortex blocks reinforcer devaluation in macaques. *J Neurosci* **31**: 15128–15135.
- Wied HM, Jones JL, Cooch NK, Berg BA, Schoenbaum G (2013). Disruption of model-based behavior and learning by cocaine self-administration in rats. *Psychopharmacology* **229**: 493–501.
- Wilson RC, Takahashi YK, Schoenbaum G, Niv Y (2014). Orbitofrontal cortex as a cognitive map of task space. *Neuron* **81**: 267–279.
- Winstanley CA, Theobald DEH, Cardinal RN, Robbins TW (2004). Contrasting roles of basolateral amygdala and orbitofrontal cortex in impulsive choice. *J Neurosci* **24**: 4718–4722.
- Wyvell CL, Berridge KC (2001). Incentive sensitization by previous amphetamine exposure: increased cue-triggered "wanting" for sucrose reward. *J Neurosci* **21**: 7831–7840.