

Orbitofrontal cortex, decision-making and drug addiction

Geoffrey Schoenbaum^{1,2}, Matthew R. Roesch¹ and Thomas A. Stalnaker¹

¹Department of Anatomy and Neurobiology, University of Maryland School of Medicine, 20 Penn Street, HSF-2 S251, Baltimore, MD 21201, USA

²Department of Psychiatry, University of Maryland School of Medicine, 20 Penn Street, HSF-2 S251, Baltimore, MD 21201, USA

The orbitofrontal cortex, as a part of prefrontal cortex, is implicated in executive function. However, within this broad region, the orbitofrontal cortex is distinguished by its unique pattern of connections with crucial subcortical associative learning nodes, such as basolateral amygdala and nucleus accumbens. By virtue of these connections, the orbitofrontal cortex is uniquely positioned to use associative information to project into the future, and to use the value of perceived or expected outcomes to guide decisions. This review will discuss recent evidence that supports this proposal and will examine evidence that loss of this signal, as the result of drug-induced changes in these brain circuits, might account for the maladaptive decision-making that characterizes drug addiction.

Introduction

Our ability to form expectations about the desirability or value of impending events underlies much of our emotion and behavior. In fact, two broad functions are crucially subserved by the formation of such expectations. On the one hand, expectations guide our immediate behavior, allowing us to pursue goals and avoid potential harm. On the other hand, expectations can be compared with actual outcomes to facilitate learning so that future behavior can become more adaptive. Both of these functions require that information about expected outcomes be maintained in memory so that it can be compared and integrated with information about internal state and current goals. Such an integrative process generates a signal that we will refer to as an outcome expectancy, a term long-used by learning theorists to refer to an internal representation of the consequences likely to follow a specific act [1]. The disruption of such a signal would be expected to create a myriad of difficulties, in the ability both to make adaptive decisions and to learn from negative consequences of decisions. In this review, we first describe recent evidence that the orbitofrontal cortex (OFC) plays a crucial role in the generation and use of outcome expectancies. Subsequently, we will discuss recent evidence that the maladaptive decisions that characterize drug addiction reflect, in part, a disruption of this signal as a result of drug-induced changes in the OFC and related brain areas.

Neural activity in the OFC and OFC-dependent behavior reflect a crucial role of the OFC in the generation of outcome expectancies

The ability to maintain information so that it can be manipulated, integrated with other information and then used to guide behavior has been variously described as working, scratchpad or representational memory, and it depends crucially on the prefrontal cortex [2]. Within the prefrontal cortex, the OFC, by its connections with limbic areas, is uniquely positioned to enable associative information regarding outcomes or consequences to access representational memory (Box 1). Indeed a growing number of studies suggest that a neural correlate of the expected value of outcomes is present and perhaps generated in the OFC. For example, human neuroimaging studies show that blood flow changes in the OFC during anticipation of expected outcomes and also when the value of an expected outcome is modified or not delivered [3–6]. This activation appears to reflect the incentive value of these items and is observed when that information is being used to guide decisions [7]. These results suggest that neurons in the OFC increase activity when such information is processed. Accordingly, neural activity in the OFC that precedes predicted rewards or punishments increases, typically reflecting the incentive values of these outcomes [8–11]. For example, when monkeys are presented with visual cues paired with differently preferred rewards, neurons in the OFC fire selectively according to whether the anticipated outcome is the preferred or non-preferred reward within that trial block [10]. Moreover, Roesch and Olson [11] have recently demonstrated that firing in the OFC tracks several other specific metrics of outcome value. For example, neurons fire differently for a reward depending on its expected size, the anticipated time required to obtain it and the possible aversive consequences associated with inappropriate behavior [11,12].

Such anticipatory activity appears to be a common feature of firing activity in the OFC across many tasks in which events occur in a sequential, and thus predictable, order (Box 2). Importantly, however, these selective responses can be observed in the absence of any signaling cues, and they are acquired as animals learn that particular cues predict a specific outcome. In other words, this selective activity represents the expectation of an animal, based on experience, of likely outcomes.

Corresponding author: Schoenbaum, G. (schoenbg@schoenbaumlab.org).

Available online 6 January 2006

Box 1. The anatomy of the orbitofrontal circuit in rats and primates

Rose and Woolsey [53] proposed that prefrontal cortex might be defined by the projections of the mediodorsal thalamus (MD) rather than by 'stratigraphic analogy' [54]. This definition provides a foundation on which to define prefrontal homologs across species. However, it is the functional and anatomical similarities that truly define homologous areas (Figure 1 of this box).

In the rat, the MD can be divided into three segments [55,56]. Projections from the medial and central segments of the MD define a region that includes the orbital areas and the ventral and dorsal agranular insular cortices [55–58]. These regions of the MD in rat receive direct afferents from the amygdala, medial temporal lobe, ventral pallidum and ventral tegmental area, and they receive olfactory input from the piriform cortex [55,56,59]. This pattern of connectivity is similar to that of the medially located, magnocellular division of primate MD, which defines the orbital prefrontal subdivision in primates [60–62]. Thus, a defined region in the orbital area of rat prefrontal cortex is likely to receive input from thalamus that is very similar to that reaching primate orbital prefrontal cortex. Based, in part, on this pattern of input, the projection fields of medial and central MD in the orbital and agranular insular areas of rat prefrontal cortex have been proposed as homologous to the primate orbitofrontal region [55,57,63–65]. These areas in rodents include the dorsal and ventral agranular insular cortex, and the lateral and

ventrolateral orbital regions. This conception of the rat orbitofrontal cortex (OFC) does not include the medial or ventromedial orbital cortex, which lie along the medial wall of the hemisphere. This region has patterns of connectivity with the MD and other areas that are more similar to other regions on the medial wall.

Other important connections highlight the similarity between the rat OFC and the primate OFC. Perhaps most notable are reciprocal connections with the basolateral complex of the amygdala (ABL), a region thought to be involved in affective or motivational aspects of learning [66–74]. In primate, these connections have been invoked to explain specific similarities in behavioral abnormalities resulting from damage to either the OFC or the ABL [14,17,75–77]. Reciprocal connections between basolateral amygdala and areas within rat OFC, particularly the agranular insular cortex [58,78–80], suggest that interactions between these structures might be similarly important for regulation of behavioral functions in rats. In addition, in both rats and primates, the OFC provides a strong efferent projection to the nucleus accumbens, overlapping with innervation from limbic structures such as the ABL and subiculum [81–84]. The specific circuitry connecting the OFC, limbic structures and nucleus accumbens presents a striking parallel across species that suggests possible similarities in functional interactions among these major components of the forebrain [81,84,85].

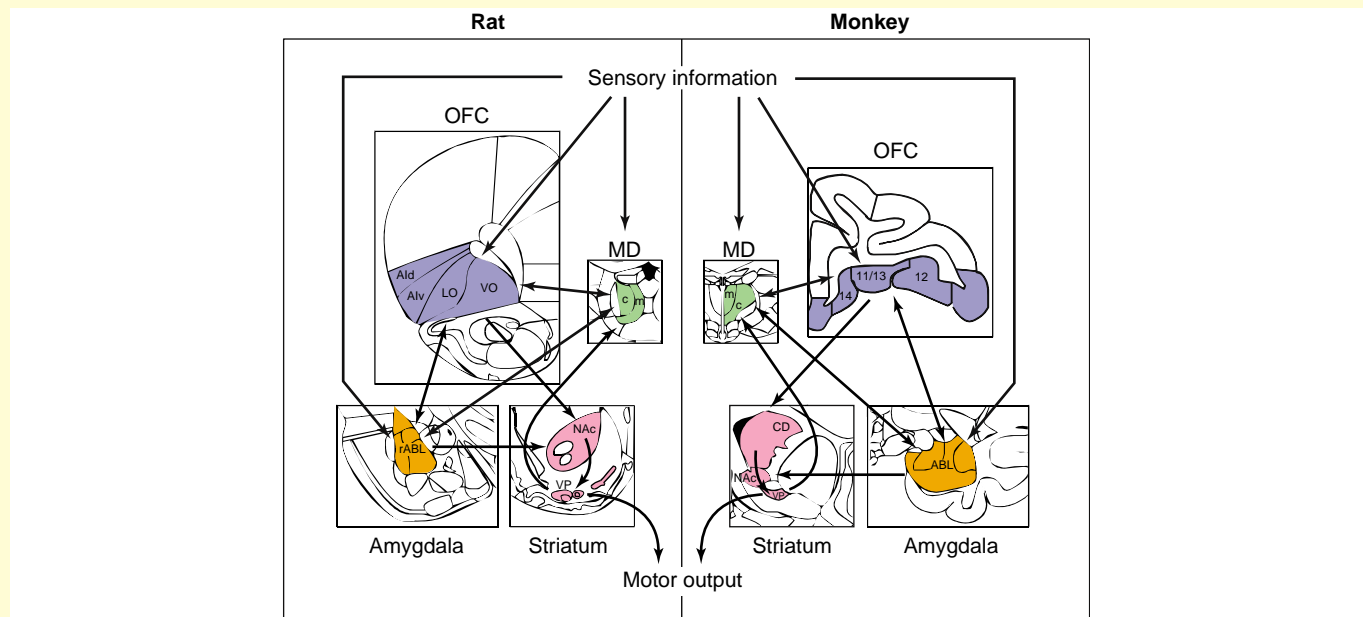


Figure 1. Anatomical relationships of the OFC (blue) in rats and monkeys. Based on their pattern of connectivity with the mediodorsal thalamus (MD, green), amygdala (orange) and striatum (pink), the orbital and agranular insular areas in rat prefrontal cortex are homologous to the primate OFC. In both species, the OFC receives robust input from sensory cortices and associative information from the amygdala, and sends outputs to the motor system through the striatum. Each box illustrates a representational coronal section. Additional abbreviations: Ald, dorsal agranular insula; Alv, ventral agranular insula; c, central; CD, caudate; LO, lateral orbital; m, medial; NAc, nucleus accumbens core; rABL, rostral basolateral amygdala; VO, ventral orbital, including ventrolateral and ventromedial orbital regions; VP, ventral pallidum.

These features are illustrated in Figure 1, which shows the population response of OFC neurons recorded in rats as they learn and reverse novel odor-discrimination problems [8,9,13]. In this simple task, the rat must learn that one odor predicts reward in a nearby fluid well, whereas the other odor predicts punishment. Early in learning, neurons in the OFC respond to one but not to the other outcome. At the same time, the neurons also begin to respond in anticipation of their preferred outcome. Over a number of studies, 15–20% of the neurons in the OFC developed such activity in this task, firing in anticipation

of either sucrose or quinine presentation [8,9,13]. The activity in this neural population reflects the value of the expected outcomes, maintained in what we have defined here as representational memory.

After learning, these neurons come to be activated by the cues that predict their preferred outcomes, thereby signaling the expected outcome even before a response is made. This is evident in the population response presented in Figure 1, which exhibits higher activity, after learning, in response to the odor cue that predicts the preferred outcome of the neuronal population.

Box 2. Orbitofrontal activity provides an ongoing signal of the value of impending events

The orbitofrontal cortex (OFC) is well positioned to use associative information to predict and then signal the value of future events. Although the main text of this review focuses on activity during delay periods before rewards to isolate this signal, the logical extension of this argument is that activity in the OFC encodes this signal throughout the performance of a task. Thus, the OFC provides a running commentary on the relative value of the current state and of possible courses of action under consideration.

This role is evident in the firing activity of OFC neurons during sampling of cues that are predictive of reward or punishment [86–88]. For example, in rats trained to perform an eight-odor discrimination task, in which four odors were associated with reward and four odors were associated with non-reward, OFC neurons were more strongly influenced by the associative significance of the odor cues than by the actual odor identities [87]. Indeed if odor identity is made irrelevant, OFC neurons will ignore this sensory feature of the cue. This was demonstrated by Ramus and Eichenbaum [89], who trained rats on an eight-odor continuous delayed non-match-to-sample task, in which the relevant construct associated with reward is not odor identity but rather the 'match' or 'non-match' comparison between the cue on the current and preceding trial. They found that 64% of the responsive neurons discriminated this match–non-match comparison, whereas only 16% fired selectively to one of the odors.

Although cue-selective firing has been interpreted as associative encoding, we suggest that this neuronal activity actually represents the ongoing evaluation of potential outcomes by the animal. Thus, the selective firing of these neurons does not simply reflect the fact that a specific cue has been reliably associated with a particular outcome in the past, but instead reflects the judgment of the animal given current circumstances that, acting on that associative information, will lead to that outcome in the future. This judgment is represented as the value of that specific outcome relative to internal goals or desires, and these expectancies are updated constantly. Thus, the firing in the OFC reflects in essence the expected value of the subsequent state that will be generated given a particular response, whether that state is a primary reinforcer or simply a step towards that ultimate goal. Consistent with this proposal, a review of the literature shows that encoding in the OFC reliably differentiates many events, even those removed from actual reward delivery, if they provide information about the likelihood of future reward (Figure 1 of this box). For example, in odor-discrimination training, OFC neurons fire in anticipation of the nose-poke that precedes odor sampling. The response of these neurons differs according to whether the sequence of recent trials [87,90] or the place [91] predicts a high probability of reward.

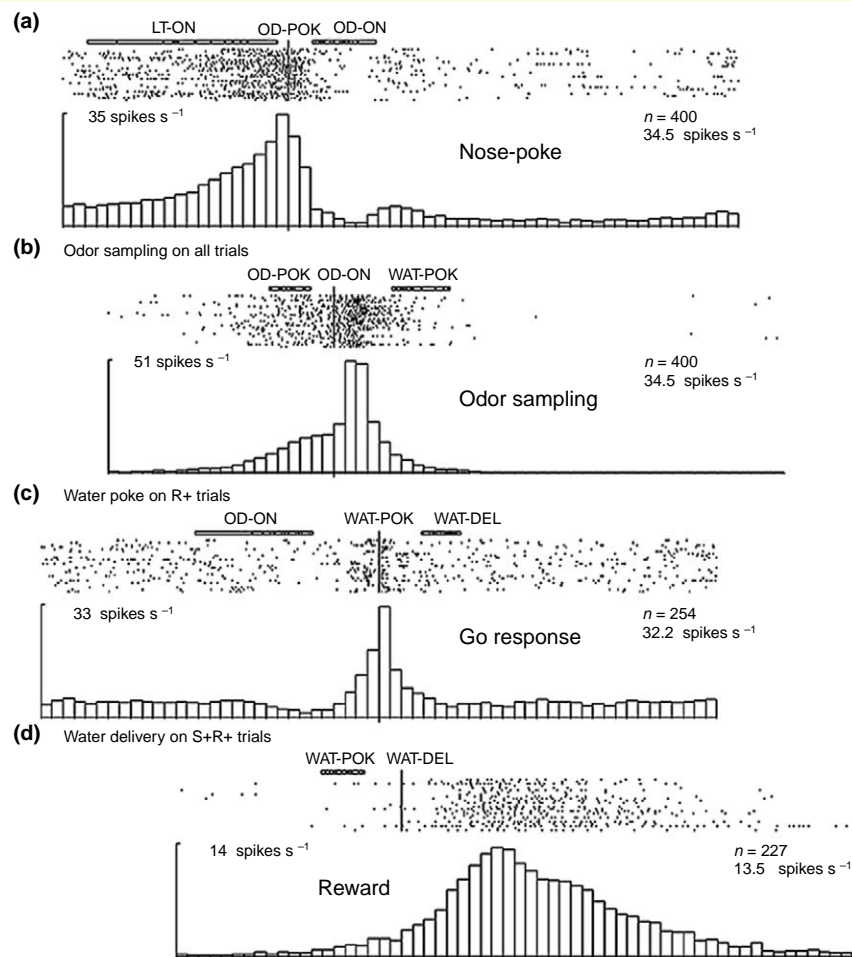


Figure 1. Neural activity in the OFC in anticipation of trial events. Neurons in the rat OFC were recorded during performance of an eight-odor, Go-NoGo odor-discrimination task. The activity in four different orbitofrontal neurons is shown, synchronized to four different task events (a–d). Activity is displayed in raster format at the top and as a peri-event time histogram at the bottom of each panel; labels over each figure indicate the synchronizing event and any events that occurred before or after light onset (LT-ON), odor poke (OD-POK), odor onset (OD-ON), water poke (WAT-POK) or water delivery (WAT-DEL). Numbers indicate number of trials (n) and number of spikes per second. The four neurons each fired in association with a different event, and the firing in each neuron increased in anticipation of that event. Adapted, with permission, from [87].

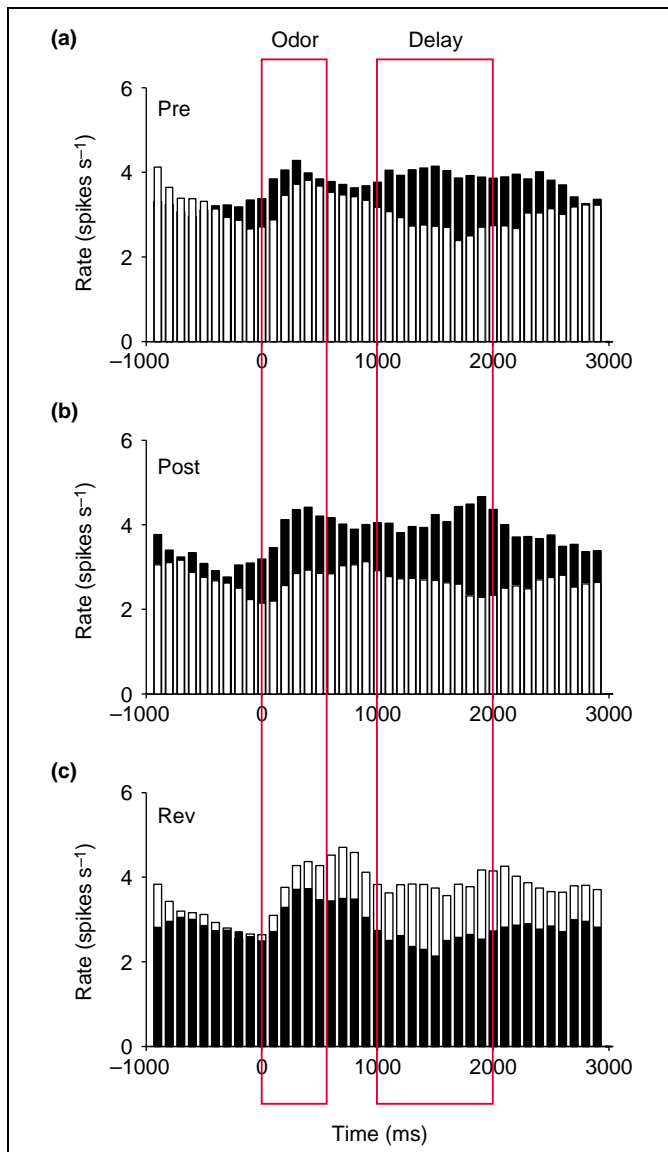


Figure 1. Signaling of outcome expectancies in the orbitofrontal cortex. Black bars show the response on trials involving the preferred outcome of the neurons in the post-criterion phase. White bars show the response to the non-preferred outcome. Activity is synchronized to odor onset and displayed in 100 ms bins; red boxes indicate the average duration of odor sampling and of the delay between responding and reinforcement when a response was made at the fluid well. (a) The average response of neurons that responded selectively during a delay period early in learning, after the rat went to the fluid well but before delivery of one of the outcomes (Pre). This activity signals the expected outcome after responding. (b) Later (Post), these neurons became activated by the odor cue that predicted the preferred outcomes of the neurons after learning, thereby signaling the expected outcome at the time the decision to respond is made. (c) This pattern switched after reversal of the odor–outcome associations (Rev). Average firing rate is shown on the y-axis in spikes per second. Adapted, with permission, from [13].

These signals would allow an animal to use expectations of likely outcomes to guide responses to cues and to facilitate learning when expectations are violated.

The notion that the OFC guides behavior by signaling outcome expectancies is consistent with the effects of OFC damage on behavior. These effects are typically evident when the appropriate response cannot be selected using simple associations, but instead requires outcome expectancies to be integrated over time or to be compared between alternative responses. For example, humans with damage to the OFC are unable to guide behavior

appropriately based on the consequences of their actions in the Iowa gambling task [14]. In this task, subjects must choose from decks of cards with varying rewards and penalties represented on the cards. To make advantageous choices, subjects must be able to integrate the value of these varying rewards and penalties over time. Individuals with OFC damage initially choose decks that yield higher rewards, indicating that they can use simple associations to direct behavior according to reward size; however, they fail to modify their responses to reflect occasional large penalties in those decks. Integrating information about the occasional, probabilistic penalties would be facilitated by an ability to maintain information about the value of the expected outcome in representational memory after a choice is made, so that violations of this expectation (occasional penalties) could be recognized. This deficit is analogous to the reversal deficits demonstrated in rats, monkeys and humans after damage to the OFC [15–21].

This ability to hold information about expected outcomes in representational memory has also been probed in a recent study in which subjects made choices between two stimuli that predicted punishment or reward at varying levels of probability [22]. In one part of this study, subjects were given feedback about the value of the outcome that they had not selected. Normal subjects were able to use this feedback to modulate their emotion about their choice and to learn to make better choices in future trials. For example, a small reward made them happier when they knew that they had avoided a large penalty. Individuals with OFC damage showed normal emotional responses to the rewards and punishments that they selected; however, feedback about the unselected outcome had no effect on either their emotions or on their subsequent performance. That is, they were happy when they received a reward, but they were no happier if they were informed that they had also avoided a large penalty. This impairment is consistent with a role for the OFC in maintaining associative information in representational memory to compare different outcome expectancies. Without this signal, individuals cannot compare the relative value of the selected and unselected outcomes and thus fail to use this comparative information to modulate emotional reactions and facilitate learning.

Although these examples are revealing, a more-direct demonstration of the crucial role of the OFC in generating outcome expectancies to guide decision-making comes from reinforcer devaluation tasks. These tasks assess the control of behavior by an internal representation of the value of an expected outcome. For example, in a Pavlovian version of this procedure (Figure 2), rats are first trained to associate a light cue with food. After conditioned responding is established to the light, the value of the food is reduced by pairing it with illness. Subsequently, in the probe test, the light cue is presented again in a non-rewarded extinction session. Animals that have received food-illness pairings respond less to the light cue than do non-devalued controls. Importantly, this decrease in responding is evident from the start of the session and is superimposed on the normal decreases in responding that result from extinction learning during the session.

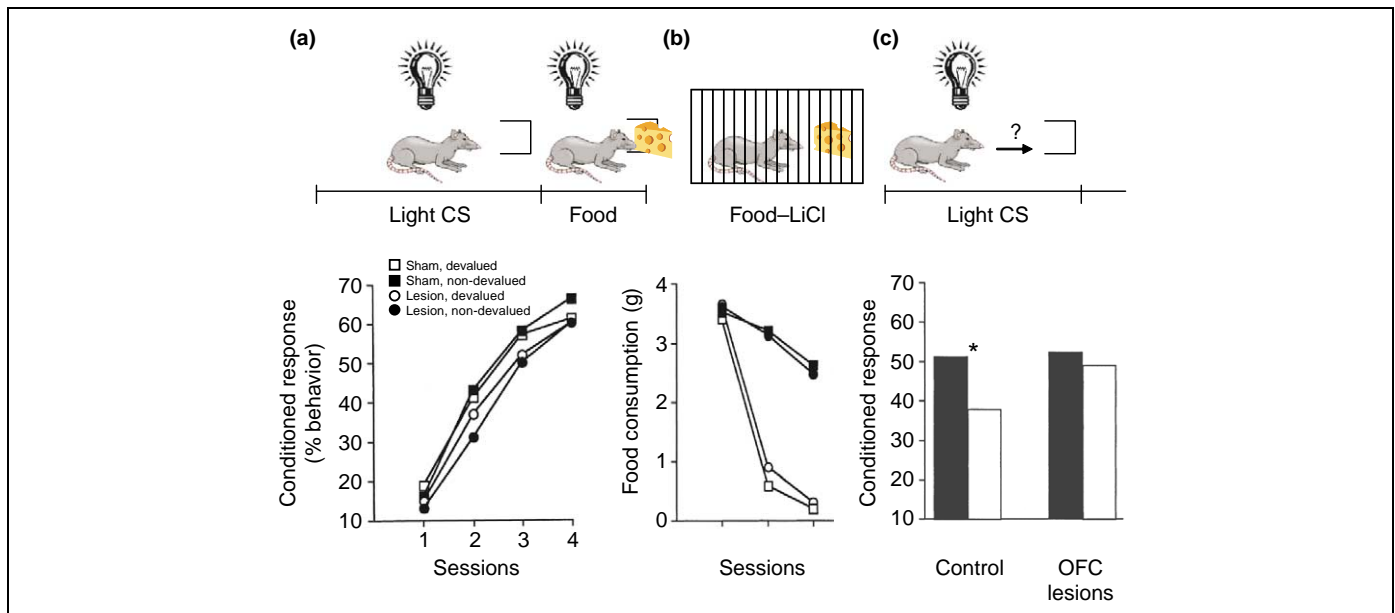


Figure 2. Effects of neurotoxic lesions of the orbitofrontal cortex (OFC) on performance in a reinforcer devaluation task. (a) Control rats and rats with bilateral neurotoxic lesions of the OFC were trained to associate a conditioned stimulus (CS, light) with an unconditioned stimulus (US, food). Over four sessions, both lesioned and control rats developed a conditioned responding at the food cup during light presentation. This food-cup response is represented as the percentage of total behavior. The lesion had no effect on the development of the food-cup response. (b) The rats then received presentations of the food item in their home cages followed by illness induced by lithium chloride (LiCl) injection. Some rats in each group received paired presentations of food and illness (black circles and squares); others received unpaired presentations (white circles and squares). Rats that received paired presentations stopped consuming the food item (i.e. the food stimulus was 'devalued'). Again, no effect of lesion was observed. (c) The following day, the rats were returned to the training environment, and conditioned responses to the light cue were measured. When exposed to the light CS, control rats that had received paired presentations of food and illness (for which the food had been devalued; white bars on the left) showed reduced conditioned responses to the food cup compared with unpaired controls (white bars on the right; asterisk indicates $P < 0.05$). Rats with OFC lesions did not show this decrease in conditioned response as a result of reinforcer devaluation. Lower panels adapted, with permission, from [23] © (1999) the Society for Neuroscience.

This initial decrease in responding must reflect the use of an internal representation of the current value of the food in combination with the original light-food association. Thus, reinforcer devaluation tasks provide a direct measure of the ability to manipulate and use outcome expectancies to guide behavior.

Rats with OFC lesions fail to show any effect of devaluation on conditioned responding in this paradigm, despite normal conditioning and devaluation of the outcome [23]. In other words, they continue to respond to the light cue and attempt to obtain the food, even though they will not consume it if it is presented (Figure 2). Importantly, OFC-lesioned rats display a normal ability to extinguish their responses within the test session, demonstrating that their deficit does not reflect a general inability to inhibit conditioned responses [24]. Rather, the OFC has a specific role in controlling conditioned responses according to internal representations of the new value of the expected outcome. Accordingly, OFC lesions made after learning continue to affect behavior in this task [25]. Similar results have been reported in monkeys trained to perform an instrumental version of this task [19].

Rats with OFC lesions also show neurophysiological changes in downstream regions that are consistent with the loss of outcome expectancies. In one study [26], responses were recorded from single units in the basolateral amygdala, an area that receives projections from OFC, in rats learning and reversing novel odor discriminations in the task described earlier. Under these conditions, OFC lesions disrupted outcome-expectant firing normally observed in the basolateral amygdala.

Furthermore, without OFC input, neurons of the basolateral amygdala became cue-selective much more slowly, particularly after cue-outcome associations were reversed. Slower associative encoding in the basolateral amygdala as a result of OFC lesions, particularly during reversal, is consistent with the idea that outcome expectancies facilitate learning in other structures, especially when expectations are violated as they are in reversals. Thus, OFC appears to generate and represent outcome expectancies that are critical not only to the guidance of behavior according to expectations about the future, but also to the ability to learn from violations of those expectations. Without this signal, animals engage in maladaptive behavior, driven by antecedent cues and stimulus-response habits, rather than by a cognitive representation of an outcome or goal.

Addictive behavior and outcome expectancies

Recent findings suggest that this conceptualization of OFC function has much to offer an understanding of drug addiction. According to the *Diagnostic and Statistical Manual of Mental Disorders* [27], a diagnosis of substance dependence requires that an individual display an inability to control his or her drug-seeking behavior, despite adverse consequences. Such addictive behavior is characterized variously as compulsive, impulsive, perseverative or under the control of drug-associated cues. Moreover, it is often observed despite a stated desire on the part of addicts to stop. Thus, a diagnosis of substance dependence requires a pattern of behavior similar to that of OFC-lesioned rats, monkeys and humans.

Accordingly, drug addiction is associated with changes in OFC structure and function. For example, imaging studies of addicts have consistently revealed abnormalities in blood flow in the OFC [28–33] (for an excellent review, see [34]). Alcohol and cocaine addicts display reductions in baseline measurements of OFC activation during acute withdrawal and even after long periods of abstinence. Conversely, during exposure to drug-related cues, addicts show an overactivation of OFC that correlates with the degree of craving that they experience. These changes are associated with impairments to OFC-dependent behaviors in drug addicts [35–39]. For example, alcohol and cocaine abusers display similar, although not as severe on average, impairments on the gambling task described earlier, as do individuals with lesions of the OFC. Similarly, other laboratory tests of decision-making have revealed that amphetamine abusers take longer and are less likely to choose the most rewarding option than are controls. But do these deficits reflect a pre-existing vulnerability to addiction in some people? Or are they a result of long-term drug-induced neuroadaptations? And if so, do they reflect changes in structure and/or function within OFC, or are they the result of changes elsewhere in corticolimbic networks that mimic the effects of OFC lesions?

To answer these questions, it is necessary to turn to animal models, in which addictive drugs can be delivered in a controlled manner against a relatively fixed genetic and environmental background. A growing number of such studies now demonstrate that prolonged exposure to addictive drugs – and particularly psychostimulants – results in relatively long-lasting brain and behavioral changes [40–50]. Importantly these effects are typically observed months after cessation of and in behavioral settings that are unrelated to drug exposure, consistent with the hypothesis that addictive drugs modify brain circuits that are crucial for the normal control of behavior. Recently, several studies have demonstrated effects on the OFC. For example, rats trained to self-administer amphetamine for several weeks have been reported to

show a reduction in dendritic spine density in the OFC one month later [46]. Furthermore, these drug-experienced rats exhibited less remodeling of their dendrites in response to appetitive instrumental training. These findings are particularly noteworthy in light of the increased spine density that has been previously reported in the medial prefrontal cortex, nucleus accumbens and elsewhere after treatment with psychostimulants [41]. Thus, among these corticolimbic regions, the OFC appears to be unique in showing evidence of decreased synaptic plasticity after drug exposure.

A decrease in plasticity in the OFC might be expected to impact on OFC-dependent functions. Consistent with this conjecture, rats given a two-week course of treatment with cocaine show long-lasting impairments in OFC-dependent behavior. Specifically, these animals are unable to use the value of predicted outcomes to guide their behavior. In one experiment [51], rats were given daily injections of cocaine for two weeks. Over one month later, these rats were tested in a Go–NoGo odor discrimination task. In this task, rats learn to go to a fluid port to obtain sucrose after smelling one odor and withhold going to the same fluid port to avoid quinine after smelling a second odor. Rats treated with cocaine learned these discriminations at the same rate as did saline-treated controls, but were unable to acquire reversals of the discriminations as rapidly as were the controls. Similar reversal deficits have also been demonstrated in primates that are given intermittent chronic access to cocaine [43]. Such reversal deficits are characteristic of OFC-lesioned animals and humans [15–21], where they are thought to reflect an inability to change established behaviors rapidly. We propose that the role of OFC in supporting this rapid flexibility relates to its importance in signaling outcome expectancies [26]. During reversal learning, the comparison of this signal with the actual, reversed outcome would generate error signals crucial to new learning [1]. Without this signal, OFC-lesioned rats would learn more slowly. As we have already discussed, a neurophysiological correlate of this slow learning has recently been demonstrated in the

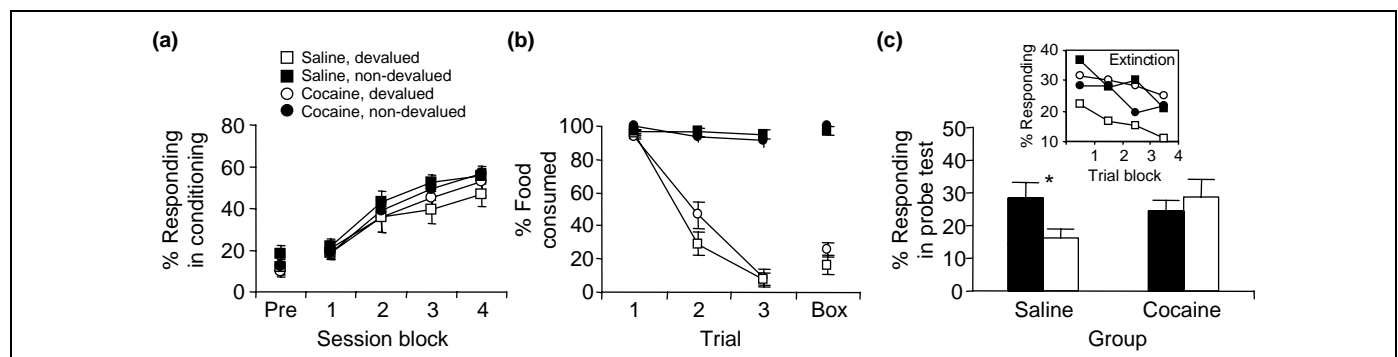


Figure 3. Effects of cocaine treatment on performance in the reinforcer devaluation task (Figure 2). Saline- and cocaine-treated rats were trained to associate a conditioned stimulus (CS, light) with an unconditioned stimulus (US, food). (a) Over four session blocks, both cocaine- and saline-treated rats developed a conditioned responding at the food cup during light presentation. This food-cup response is represented as the percentage of total behavior. There was no effect of the lesion on the development of the food-cup response. (b) The rats then received presentations of the food item in their home cages followed by illness induced by LiCl injection. Some rats in each group received paired presentations of food and illness (black circles and squares) while others received unpaired presentations (white circles and squares). Rats that received paired presentations stopped consuming the food item, both in their home cage and in the testing chamber ('Box'). Again, no effect of lesion was observed. (c) The following day, the rats were returned to the training environment, and conditioned responses to the light cue were measured. Control rats that had received paired presentations of food and illness showed reduced conditioned responses to the food cup compared with unpaired controls. Rats with orbitofrontal lesions did not show this decrease in conditioned response as a result of reinforcer devaluation. This effect was observed while all four groups showed normal extinction of responding during the probe test (inset). Adapted, with permission of Oxford University Press, from [24].

inflexible associative encoding of basolateral amygdala neurons in OFC-lesioned rats [26].

The loss of this signal is also evident in a second experiment in which rats were treated with cocaine for two weeks and then tested in the Pavlovian reinforcer devaluation task described earlier [24]. Again, testing was conducted about one month after the last cocaine treatment. These rats exhibited normal conditioning and devaluation, and also extinguished responding normally in the final test phase; however, devalued cocaine-treated rats did not show the normal spontaneous reduction in response to the predictive cue. This deficit (Figure 3) is identical to the deficit after OFC lesions in this task (Figure 2). These findings are consistent with an inability to signal the value of the expected outcome. Indeed, because in this task there is no ambiguity regarding the representations required to mediate normal performance, the deficits described here point unequivocally towards a loss of outcome expectancies in cocaine-treated rats.

Loss of this signaling mechanism would account for the propensity of addicts to continue to seek drugs, despite the almost inevitable negative consequences of such behavior, because it would render them unable to incorporate this predictive information into their decision-making and perhaps unable to learn from even repeated experience of these negative consequences. Although other brain systems might also be involved, drug-induced changes to this OFC-dependent signal would by themselves contribute powerfully to a transition from normal goal-directed behavior to compulsive habitual responding. This transition would reflect a change in the balance between these competing mechanisms of behavioral control. Such an explanation would hold for the drug-seeking behavior of addicts, and also for recent findings in several animal models of addiction in which rats are unable to withhold drug-seeking behavior, even when adverse outcomes are made contingent upon that behavior [45,47].

Concluding remarks

We have reviewed recent findings to support the proposal that the OFC is crucial for signaling the value of expected outcomes or consequences. We have also discussed how this idea might be important for understanding the pathology that underlies drug addiction. Of course these ideas raise many more questions. If the OFC generates signals regarding expected outcomes, it becomes crucial to understand how downstream areas use these signals – in normal animals, in addition to those exposed to addictive drugs. We have suggested how the basolateral amygdala might be involved [26]; however, understanding the role these signals have in the nucleus accumbens – and how they interact with other ‘limbic’ inputs – might be far more relevant for understanding addiction. Several laboratories are working hard to resolve these important issues. In addition, it will be important to demonstrate whether changes in OFC-dependent behavior after drug exposure actually reflect altered molecular or neurophysiological function in the OFC, as suggested by preliminary recording data [52], or alternatively whether they might reflect changes elsewhere in the circuit, such as in the nucleus accumbens, an area long implicated in addiction.

Box 3. Unanswered questions

- (i) How do downstream areas – particularly the nucleus accumbens – use signals regarding outcome expectancies from the OFC? How is this information integrated with other ‘limbic’ inputs to the accumbens?
- (ii) Can changes in OFC-dependent behaviors after drug exposure be linked to changes in molecular or neurophysiological targets within the OFC? Or do these behavioral deficits reflect changes elsewhere in learning circuits?
- (iii) Can drug-related changes in behavior or other markers be reversed by behavioral or pharmacological manipulations?
- (iv) Are functional changes in the OFC or related learning circuits different in animals given contingent versus non-contingent drug experiences? And if so, do the differences have a critical impact on behavior?
- (v) Do changes in the OFC underlie behavior in drug addiction models of compulsive drug seeking and relapse? And might they be particularly important early in the transition to addiction, promoting ongoing drug use before striatal changes, which are associated with more long-term access, become influential?

And, of course, any animal model of disease is only of value if it suggests a remedy for the pathological changes. This is difficult in the case of lesions but could be possible for deficits stemming from drug exposure. However, it remains to be seen whether manipulations might be undertaken to normalize the behavior and perhaps any molecular or neurophysiological correlates that are identified in drug-treated animals. We expect that these and many more issues will be addressed in the coming years (Box 3).

Acknowledgements

Our research was supported by grants from the NIDA (R01-DA015718 to G.S.), NINDS (T32-NS07375 to M.R.R.) and NIDCD (T32-DC00054 to T.A.S.).

References

- 1 Dickinson, A. (1989) Expectancy theory in animal conditioning. In *Contemporary Learning Theories: Pavlovian Conditioning and the Status of Traditional Learning Theory* (Klein, S.B. and Mowrer, R.R., eds), pp. 279–308, Erlbaum
- 2 Goldman-Rakic, P.S. (1987) Circuitry of primate prefrontal cortex and regulation of behavior by representational memory. In *Handbook of Physiology: The Nervous System* (Vol. V) (Mountcastle, V.B. et al., eds), pp. 373–417, American Physiology Society
- 3 Gottfried, J.A. et al. (2003) Encoding predictive reward value in human amygdala and orbitofrontal cortex. *Science* 301, 1104–1107
- 4 Gottfried, J.A. et al. (2002) Appetitive and aversive olfactory learning in humans studied using event-related functional magnetic resonance imaging. *J. Neurosci.* 22, 10829–10837
- 5 O’Doherty, J. et al. (2002) Neural responses during anticipation of a primary taste reward. *Neuron* 33, 815–826
- 6 Nobre, A.C. et al. (1999) Orbitofrontal cortex is activated during breaches of expectation in tasks of visual attention. *Nat. Neurosci.* 2, 11–12
- 7 Arana, F.S. et al. (2003) Dissociable contributions of the human amygdala and orbitofrontal cortex to incentive motivation and goal selection. *J. Neurosci.* 23, 9632–9638
- 8 Schoenbaum, G. et al. (2003) Encoding predicted outcome and acquired value in orbitofrontal cortex during cue sampling depends upon input from basolateral amygdala. *Neuron* 39, 855–867
- 9 Schoenbaum, G. et al. (1998) Orbitofrontal cortex and basolateral amygdala encode expected outcomes during learning. *Nat. Neurosci.* 1, 155–159
- 10 Tremblay, L. and Schultz, W. (1999) Relative reward preference in primate orbitofrontal cortex. *Nature* 398, 704–708

- 11 Roesch, M.R. and Olson, C.R. (2004) Neuronal activity related to reward value and motivation in primate frontal cortex. *Science* 304, 307–310
- 12 Roesch, M.R. and Olson, C.R. (2005) Neuronal activity in primate orbitofrontal cortex reflects the value of time. *J. Neurophysiol.* 94, 2457–2471
- 13 Schoenbaum, G. *et al.* Encoding changes in orbitofrontal cortex in reversal-impaired aged rats. *J. Neurophysiol.* (in press)
- 14 Bechara, A. *et al.* (1999) Different contributions of the human amygdala and ventromedial prefrontal cortex to decision-making. *J. Neurosci.* 19, 5473–5481
- 15 Schoenbaum, G. *et al.* (2003) Lesions of orbitofrontal cortex and basolateral amygdala complex disrupt acquisition of odor-guided discriminations and reversals. *Learn. Mem.* 10, 129–140
- 16 Rolls, E.T. *et al.* (1994) Emotion-related learning in patients with social and emotional changes associated with frontal lobe damage. *J. Neurol. Neurosurg. Psychiatry* 57, 1518–1524
- 17 Jones, B. and Mishkin, M. (1972) Limbic lesions and the problem of stimulus-reinforcement associations. *Exp. Neurol.* 36, 362–377
- 18 Chudasama, Y. and Robbins, T.W. (2003) Dissociable contributions of the orbitofrontal and infralimbic cortex to pavlovian autoshaping and discrimination reversal learning: further evidence for the functional heterogeneity of the rodent frontal cortex. *J. Neurosci.* 23, 8771–8780
- 19 Izquierdo, A. *et al.* (2004) Bilateral orbital prefrontal cortex lesions in rhesus monkeys disrupt choices guided by both reward value and reward contingency. *J. Neurosci.* 24, 7540–7548
- 20 Fellows, L.K. and Farah, M.J. (2003) Ventromedial frontal cortex mediates affective shifting in humans: evidence from a reversal learning paradigm. *Brain* 126, 1830–1837
- 21 Dias, R. *et al.* (1996) Dissociation in prefrontal cortex of affective and attentional shifts. *Nature* 380, 69–72
- 22 Camille, N. *et al.* (2004) The involvement of the orbitofrontal cortex in the experience of regret. *Science* 304, 1167–1170
- 23 Gallagher, M. *et al.* (1999) Orbitofrontal cortex and representation of incentive value in associative learning. *J. Neurosci.* 19, 6610–6614
- 24 Schoenbaum, G. and Setlow, B. (2005) Cocaine makes actions insensitive to outcomes but not extinction: implications for altered orbitofrontal-amygdalar function. *Cereb. Cortex* 15, 1162–1169
- 25 Pickens, C.L. *et al.* (2003) Different roles for orbitofrontal cortex and basolateral amygdala in a reinforcer devaluation task. *J. Neurosci.* 23, 11078–11084
- 26 Saddoris, M.P. *et al.* (2005) Rapid associative encoding in basolateral amygdala depends on connections with orbitofrontal cortex. *Neuron* 46, 321–331
- 27 American Psychiatric Association (2000) *Diagnostic and Statistical Manual of Mental Disorders (Text Revision)* (4th edn), American Psychiatric Association
- 28 London, E.D. *et al.* (2000) Orbitofrontal cortex and human drug abuse: functional imaging. *Cereb. Cortex* 10, 334–342
- 29 Rogers, R.D. *et al.* (1999) Dissociable deficits in the decision-making cognition of chronic amphetamine abusers, opiate abusers, patients with focal damage to prefrontal cortex, and tryptophan-depleted normal volunteers: evidence for monoaminergic mechanisms. *Neuropsychopharmacology* 20, 322–339
- 30 Maas, L.C. *et al.* (1998) Functional magnetic resonance imaging of human brain activation during cue-induced cocaine craving. *Am. J. Psychiatry* 155, 124–126
- 31 Breiter, H.C. *et al.* (1997) Acute effects of cocaine on human brain activity and emotion. *Neuron* 19, 591–611
- 32 Porrino, L.J. and Lyons, D. (2000) Orbital and medial prefrontal cortex and psychostimulant abuse: studies in animal models. *Cereb. Cortex* 10, 326–333
- 33 Volkow, N.D. and Fowler, J.S. (2000) Addiction, a disease of compulsion and drive: involvement of orbitofrontal cortex. *Cereb. Cortex* 10, 318–325
- 34 Dom, G. *et al.* (2005) Substance use disorders and the orbitofrontal cortex. *Br. J. Psychiatry* 187, 209–220
- 35 Bechara, A. *et al.* (2001) Decision-making deficits, linked to a dysfunctional ventromedial prefrontal cortex, revealed in alcohol and stimulant abusers. *Neuropsychologia* 39, 376–389
- 36 Coffey, S.F. *et al.* (2003) Impulsivity and rapid discounting of delayed hypothetical rewards in cocaine-dependent individuals. *Exp. Clin. Psychopharmacol.* 11, 18–25
- 37 Bechara, A. and 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
- 38 Bechara, A. *et al.* (2002) Decision-making and addiction (part II): myopia for the future or hypersensitivity to reward? *Neuropsychologia* 40, 1690–1705
- 39 Grant, S. *et al.* (2000) Drug abusers show impaired performance in a laboratory test of decision making. *Neuropsychologia* 38, 1180–1187
- 40 Harmer, C.J. and Phillips, G.D. (1998) Enhanced appetitive conditioning following repeated pretreatment with D-amphetamine. *Behav. Pharmacol.* 9, 299–308
- 41 Robinson, T.E. and Kolb, B. (1999) Alterations in the morphology of dendrites and dendritic spines in the nucleus accumbens and prefrontal cortex following repeated treatment with amphetamine or cocaine. *Eur. J. Neurosci.* 11, 1598–1604
- 42 Wyvell, C.L. and Berridge, K.C. (2001) Incentive sensitization by previous amphetamine exposure: increased cue-triggered ‘wanting’ for sucrose reward. *J. Neurosci.* 21, 7831–7840
- 43 Jentsch, J.D. *et al.* (2002) Impairments of reversal learning and response perseveration after repeated, intermittent cocaine administrations to monkeys. *Neuropsychopharmacology* 26, 183–190
- 44 Taylor, J.R. and Horger, B.A. (1999) Enhanced responding for conditioned reward produced by intra-accumbens amphetamine is potentiated after cocaine sensitization. *Psychopharmacology (Berl.)* 142, 31–40
- 45 Vanderschuren, L.J.M.J. and Everitt, B.J. (2004) Drug seeking becomes compulsive after prolonged cocaine self-administration. *Science* 305, 1017–1019
- 46 Crombag, H.S. *et al.* (2004) Opposite effects of amphetamine self-administration experience on dendritic spines in the medial and orbital prefrontal cortex. *Cereb. Cortex* 15, 341–348
- 47 Miles, F.J. *et al.* (2003) Oral cocaine seeking by rats: action or habit? *Behav. Neurosci.* 117, 927–938
- 48 Horger, B.A. *et al.* (1990) Preexposure sensitizes rats to the rewarding effects of cocaine. *Pharmacol. Biochem. Behav.* 37, 707–711
- 49 Phillips, G.D. *et al.* (2002) Blockade of sensitization-induced facilitation of appetitive conditioning by post-session intra-amygdaloid nafadotride. *Behav. Brain Res.* 134, 249–257
- 50 Taylor, J.R. and Jentsch, J.D. (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
- 51 Schoenbaum, G. *et al.* (2004) Cocaine-experienced rats exhibit learning deficits in a task sensitive to orbitofrontal cortex lesions. *Eur. J. Neurosci.* 19, 1997–2002
- 52 Stalnaker, T.A. *et al.* (2005) Orbitofrontal cortex fails to represent bad outcomes after cocaine exposure. Program number 112.2. In *Abstract Viewer and Itinerary Planner*, Society for Neuroscience Online (<http://sfn.scholarone.com/>)
- 53 Rose, J.E. and Woolsey, C.N. (1948) The orbitofrontal cortex and its connections with the mediodorsal nucleus in rabbit, sheep, and cat. *Res. Pub. Ass. Nerv. Ment. Dis.* 27, 210–232
- 54 Ramón y Cajal, S. (1988) Studies on the fine structure of the regional cortex of rodents 1: suboccipital cortex (retrosplenial cortex of Brodmann) [Trabajos del Laboratorio de Investigaciones Biológicas de la Universidad de Madrid, 20: 1–30, 1922]. In *Cajal on the Cerebral Cortex: An Annotated Translation of the Complete Writings* (Defelipe, J. and Jones, E.G., eds), pp. 524–546, Oxford University Press
- 55 Groenewegen, H.J. (1988) Organization of the afferent connections of the mediodorsal thalamic nucleus in the rat, related to the mediodorsal-prefrontal topography. *Neuroscience* 24, 379–431
- 56 Krettek, J.E. and Price, J.L. (1977) The cortical projections of the mediodorsal nucleus and adjacent thalamic nuclei in the rat. *J. Comp. Neurol.* 171, 157–192
- 57 Leonard, C.M. (1969) The prefrontal cortex of the rat. I. Cortical projections of the mediodorsal nucleus. II. Efferent connections. *Brain Res.* 12, 321–343
- 58 Kolb, B. (1984) Functions of the frontal cortex of the rat: a comparative review. *Brain Res.* 8, 65–98

- 59 Ray, J.P. and Price, J.L. (1992) The organization of the thalamocortical connections of the mediodorsal thalamic nucleus in the rat, related to the ventral forebrain – prefrontal cortex topography. *J. Comp. Neurol.* 323, 167–197
- 60 Goldman-Rakic, P.S. and Porrino, L.J. (1985) The primate mediodorsal (MD) nucleus and its projection to the frontal lobe. *J. Comp. Neurol.* 242, 535–560
- 61 Russchen, F.T. *et al.* (1987) The afferent input to the magnocellular division of the mediodorsal thalamic nucleus in the monkey, *Macaca fascicularis*. *J. Comp. Neurol.* 256, 175–210
- 62 Kievit, J. and Kuypers, H.G.J.M. (1977) Organization of the thalamocortical connections to the frontal lobe in the Rhesus monkey. *Exp. Brain Res.* 29, 299–322
- 63 Preuss, T.M. (1995) Do rats have prefrontal cortex? The Rose–Woolsey–Akert program reconsidered. *J. Comp. Neurol.* 7, 1–24
- 64 Ongur, D. and Price, J.L. (2000) The organization of networks within the orbital and medial prefrontal cortex of rats, monkeys and humans. *Cereb. Cortex* 10, 206–219
- 65 Schoenbaum, G. and Setlow, B. (2001) Integrating orbitofrontal cortex into prefrontal theory: common processing themes across species and subdivision. *Learn. Mem.* 8, 134–147
- 66 Baxter, M.G. and Murray, E.A. (2002) The amygdala and reward. *Nat. Rev. Neurosci.* 3, 563–573
- 67 Kluver, H. and Bucy, P.C. (1939) Preliminary analysis of the temporal lobes in monkeys. *Arch. Neurol. Psychiatry* 42, 979–1000
- 68 Brown, S. and Schafer, E.A. (1888) An investigation into the functions of the occipital and temporal lobes of the monkey's brain. *Philos. Trans. R. Soc. London Ser. B* 179, 303–327
- 69 LeDoux, J.E. (1996) *The Emotional Brain*, Simon and Schuster
- 70 Weiskrantz, L. (1956) Behavioral changes associated with ablations of the amygdaloid complex in monkeys. *J. Comp. Physiol. Psychol.* 9, 381–391
- 71 Holland, P.C. and Gallagher, M. (1999) Amygdala circuitry in attentional and representational processes. *Trends Cogn. Sci.* 3, 65–73
- 72 Gallagher, M. (2000) The amygdala and associative learning. In *The Amygdala: A Functional Analysis* (Aggleton, J.P., ed.), pp. 311–330, Oxford University Press
- 73 Davis, M. (2000) The role of the amygdala in conditioned and unconditioned fear and anxiety. In *The Amygdala: A Functional Analysis* (Aggleton, J.P., ed.), pp. 213–287, Oxford University Press
- 74 Everitt, B.J. and Robbins, T.W. (1992) Amygdala–ventral striatal interactions and reward-related processes. In *The Amygdala: Neurological Aspects of Emotion, Memory, and Mental Dysfunction* (Aggleton, J.P., ed.), pp. 401–429, John Wiley and Sons
- 75 Fuster, J.M. (1997) *The Prefrontal Cortex*, Lippin-Ravencott
- 76 Gaffan, D. and Murray, E.A. (1990) Amygdalar interaction with the mediodorsal nucleus of the thalamus and the ventromedial prefrontal cortex in stimulus-reward associative learning in the monkey. *J. Neurosci.* 10, 3479–3493
- 77 Baxter, M.G. *et al.* (2000) Control of response selection by reinforcer value requires interaction of amygdala and orbitofrontal cortex. *J. Neurosci.* 20, 4311–4319
- 78 Krettek, J.E. and Price, J.L. (1977) Projections from the amygdaloid complex to the cerebral cortex and thalamus in the rat and cat. *J. Comp. Neurol.* 172, 687–722
- 79 Kita, H. and Kitai, S.T. (1990) Amygdaloid projections to the frontal cortex and the striatum in the rat. *J. Comp. Neurol.* 298, 40–49
- 80 Shi, C.J. and Cassell, M.D. (1998) Cortical, thalamic, and amygdaloid connections of the anterior and posterior insular cortices. *J. Comp. Neurol.* 399, 440–468
- 81 Groenewegen, H.J. *et al.* (1990) The anatomical relationship of the prefrontal cortex with the striatopallidal system, the thalamus and the amygdala: evidence for a parallel organization. *Prog. Brain Res.* 85, 95–118
- 82 Groenewegen, H.J. *et al.* (1987) Organization of the projections from the subiculum to the ventral striatum in the rat. A study using anterograde transport of *Phaseolus vulgaris* leucoagglutinin. *Neuroscience* 23, 103–120
- 83 Haber, S.N. *et al.* (1995) The orbital and medial prefrontal circuit through the primate basal ganglia. *J. Neurosci.* 15, 4851–4867
- 84 McDonald, A.J. (1991) Organization of the amygdaloid projections to the prefrontal cortex and associated striatum in the rat. *Neuroscience* 44, 1–14
- 85 O'Donnell, P. (1999) Ensemble coding in the nucleus accumbens. *Psychobiology* 27, 187–197
- 86 Thorpe, S.J. *et al.* (1983) The orbitofrontal cortex: neuronal activity in the behaving monkey. *Exp. Brain Res.* 49, 93–115
- 87 Schoenbaum, G. and Eichenbaum, H. (1995) Information coding in the rodent prefrontal cortex. I. Single-neuron activity in orbitofrontal cortex compared with that in pyriform cortex. *J. Neurophysiol.* 74, 733–750
- 88 Schoenbaum, G. *et al.* (1999) Neural encoding in orbitofrontal cortex and basolateral amygdala during olfactory discrimination learning. *J. Neurosci.* 19, 1876–1884
- 89 Ramus, S.J. and Eichenbaum, H. (2000) Neural correlates of olfactory recognition memory in the rat orbitofrontal cortex. *J. Neurosci.* 20, 8199–8208
- 90 Schoenbaum, G. and Eichenbaum, H. (1995) Information coding in the rodent prefrontal cortex. II. Ensemble activity in orbitofrontal cortex. *J. Neurophysiol.* 74, 751–762
- 91 Lipton, P.A. *et al.* (1999) Crossmodal associative memory representations in rodent orbitofrontal cortex. *Neuron* 22, 349–359

Trends Editorial Policy

Trends journals are indispensable reading for anyone interested in the life-sciences. At the heart of the journal are authoritative overview articles which, through synthesis and discussion, present an integrated view of the latest research. TINS Reviews and Opinions are written by leading authors, and the majority are commissioned by the editor, but we occasionally consider proposals.

- **Review** articles provide clear, concise, well illustrated and balanced discussions of recent advances, synthesizing the primary literature and identifying important trends and key questions for ongoing research.
- **Opinion** articles are special reviews designed to stimulate debate and cover controversial and emerging areas of research, and to present new hypotheses relating to important outstanding questions in neuroscience.
- **Research Focus** articles provide a short but critical analysis of recent primary research papers, and are restricted to 1000 words plus one figure.

All submitted articles are thoroughly peer reviewed, and only articles that reach the required standard are published.

Authors wishing to contribute to TINS: please submit a point-by-point outline of your intended article, together with key references that illustrate both why you would be our first author of choice for such an article and the breadth of intended source material (to tins@elsevier.com).