Available online at www.sciencedirect.com
SCIENCE DIRECT®

Behavioural Brain Research 146 (2003) 19–29

BEHAVIOURAL BRAIN RESEARCH
www.elsevier.com/locate/bbr

Review

A systems approach to orbitofrontal cortex function: recordings in rat orbitofrontal cortex reveal interactions with different learning systems

Geoffrey Schoenbaum[a,*], Barry Setlow[b], Seth J. Ramus[c]

[a] *Department of Anatomy and Neurobiology, University of Maryland School of Medicine,*
685 West Baltimore Street, Baltimore, MD 21201, USA
[b] *Department of Psychology, Texas A&M University, College Station, TX 77843, USA*
[c] *Departments of Psychology and Neuroscience, Bowdoin College, 6900 College Station, Brunswick, ME 04011, USA*

Abstract

The recognition that certain aspects of prefrontal function can be effectively modeled in rats has led to a slow expansion of interest in rat prefrontal cortex over the past decade. One of the most promising of these model systems is the orbitofrontal cortex of the rat. Rat orbitofrontal cortex is anatomically similar to the orbital prefrontal region in primates, and this similarity is borne out by behavioral and neurophysiological findings. Here we will present data on orbitofrontal cortex function from a number of parallel studies from our laboratories that employed single unit recording techniques to probe neural encoding in rat orbitofrontal cortex and related parts of the amygdala and the hippocampal memory systems. Together, these reports and associated behavioral studies suggest that the orbitofrontal region, in both rats and primates, is specialized to integrate concrete and abstract sensory constructs with information regarding the incentive value of associated outcomes to guide or modulate behavior. To the extent that monkey prefrontal function can model certain aspects of human prefrontal function, we argue that this model can now be extended to the rat orbitofrontal cortex. In addition, we argue that the function of orbitofrontal cortex needs to be considered in terms of its interactions with other brain systems.

> A common functional identity is determined by the same type of structure and connections, whatever the mammal examined. . . . stratiographic analogy, grossly appreciated in Nissl or Weigert preparations, constitutes a valuable data point, but it is not completely decisive or infallible.
> –Cajal (p. 524, 1922)

Not long ago, the ‘functional identity’ of broad cortical areas between different species was defined by cytoarchitectonic criteria. These criteria largely eliminated rats as an effective model for the study of prefrontal function because rats lack the signature granular cell layer that distinguishes much of prefrontal cortex in both human and non-human primate species. Yet, as Cajal knew [1], such features are only markers for function. Rose and Woolsey [2] recognized this fact in 1948 when they suggested that prefrontal cortex be defined by projections from mediodorsal thalamus rather than by these cytoarchitectonic markers. Subsequently, Leonard [3] suggested that the cortical projections from the medial and lateral subdivisions of the mediodorsal thalamus in the rat might define prefrontal areas functionally similar to the orbital prefrontal region and the ‘frontal convexity’ in primates. These observations have generated a slow expansion of interest in rat models of prefrontal function. Importantly, the development of such models has followed the principle elucidated by Cajal and coworkers [1,3–5], that functional identity should be based upon convergent evidence. As a result, there are now useful rat models of prefrontal function from which conclusions can be reasonably expected to apply to primates. Indeed, convergent evidence from anatomical, behavioral, and neurophysiological approaches appears to define a region of rat prefrontal cortex that is remarkably similar to primate orbitofrontal cortex [6]. Here we will describe a systems approach to the analysis of a circuit of structures related to this orbitofrontal region from which general conclusions regarding orbitofrontal function across species may be drawn.

Orbitofrontal cortex has been described variously as important for affective learning, response inhibition and rule acquisition [7]. Yet the orbitofrontal cortex is linked to other structures also involved in aspects of these functions. In particular, the orbitofrontal cortex has reciprocal connections with basolateral amygdala [8–11] and the parahippocam-

---

[*] Corresponding author. Tel.: +1-410-706-3814; fax: +1-410-706-2512.
E-mail address: schoenbg@schoenbaumlab.org (G. Schoenbaum).

0166-4328/$ – see front matter © 2003 Elsevier B.V. All rights reserved.
doi:10.1016/j.bbr.2003.09.013

pal region [12,13]. In addition, this area of prefrontal cortex receives processed sensory input from olfactory, taste, and somatosensory areas, and has connections with motor systems (see Schoenbaum et al. [6], for review). Thus, the orbitofrontal region forms a nexus in which sensory information can be integrated with abstract relational processing regarding associations between cues and with affective information regarding the incentive value of likely outcomes. The orbitofrontal cortex clearly uses this convergence of information to guide behavioral output. To explore this integrative function, we will first present data to suggest that during a simple discrimination task in which odors are associated with appetitive or aversive outcomes, networks in orbitofrontal cortex and basolateral amygdala cooperate to mediate the process by which the odor cues acquire incentive value and guide subsequent behavior. These findings suggest that orbitofrontal regions apply information regarding incentive value acquired or cued via input from amygdala to guide responding. Thereafter, we will consider data from orbitofrontal cortex and the parahippocampal region in a more complex discrimination task in which abstract representations are associated with outcomes. In this task, orbitofrontal cortex appears to integrate information from the parahippocampal region to form more abstract representations of the cues to guide task performance. Finally, we will speculate on how these functions of orbitofrontal cortex are employed in normal behavior involving both simple and abstract concepts.

1. Incentive value and neural activity in the orbitofrontal cortex and the basolateral amygdala

The orbitofrontal cortex and amygdala have strong reciprocal connections. This pattern of connectivity has been invoked to explain the apparent behavioral similarities in the effects of lesions of these two structures. These similarities were first evident in initial reports made many years ago [14–16], which noted that damage to the frontal part of the brain or within the temporal lobe appeared to cause deficits in socially appropriate behavior. More recently, Weiskrantz [17] proposed that the amygdala was particularly important for forming associations between cues and primary reinforcers, and this function has since been localized to basolateral amygdala (consisting of the lateral, basal, and accessory basal nuclei). Rats with lesions of the basolateral amygdala are impaired on a variety of appetitively and aversively motivated tasks [18–24], and these deficits reflect in part a loss of stimulus–outcome associations [25,26]. Notably, basolateral amygdala is the region within amygdala that is most strongly connected to the orbitofrontal areas, and recently the orbitofrontal region has also been implicated in supporting behavior mediated by representations of outcome [27,28].

These reports are consistent with findings from neuronal recording studies across species, showing that cells in orbitofrontal cortex [29–37] and basolateral amygdala [30,36,38–41] fire to cues based on their significance or past association with reward. Animals that have been trained in a variety of tasks in which associations are made between cues and outcomes exhibit such cue-selective firing in both regions. In order to better define the relative roles of the two populations of neurons, Schoenbaum et al. [30] examined encoding in both structures in a single paradigm designed to ask how selective firing developed with learning and was affected by changes in the associations with outcome in the task. Cells were recorded in the orbitofrontal cortex and basolateral amygdala of rats during acquisition and reversal of new odor discrimination problems, and the pattern of firing in each cell was characterized across phases related to learning and reversal.

Cells in both the orbitofrontal cortex and basolateral amygdala exhibited firing selective for one or the other odor cues (i.e. fired more during sampling of one of the cues than the other), and the firing during accurate performance was in approximately similar proportions for the two brain regions. However, a detailed examination across learning and reversal revealed substantial differences between these two regions in the encoding properties of this cue-selective neuronal firing. Cue-selective firing in orbitofrontal cortex closely reflected the conjunctions between particular odor cues and the outcome with which they were associated (Fig. 1). Selectivity emerged with the development of accurate go/no-go choice behavior, and as a group these cells became non-selective when the outcomes of the odor cues were switched after reversal. Moreover, a large proportion of the previously non-selective neurons became cue-selective when presented with new odor–outcome combinations during reversal training. This encoding, which reflected the conjunctions between a particular odor cue and outcome, was also closely dependent on whether that information was used to guide accurate choice performance in the discrimination task.

By contrast, cells in the basolateral amygdala exhibited encoding more closely related to outcome than choice performance (Fig. 2). Cue-selective neurons in basolateral amygdala developed responses only a few trials into pre-criterion training. This training phase was characterized by changes in response latency and chance or near-chance choice performance. Thus, odor-selectivity in basolateral amygdala neurons emerged much more quickly than in comparable cells in orbitofrontal cortex and independently from the use of that information to guide discriminative responding. After reversal, the majority of the cue-selective neurons in the amygdala reversed their cue-selectivity (i.e. became selective for the opposite cue), highlighting the greater dependence of firing in basolateral amygdala on associated outcome rather than on the conjunctions between particular odors and outcomes that are reconfigured after reversal.

Overall, this pattern of findings is entirely consistent with the behavioral findings implicating both the orbitofrontal

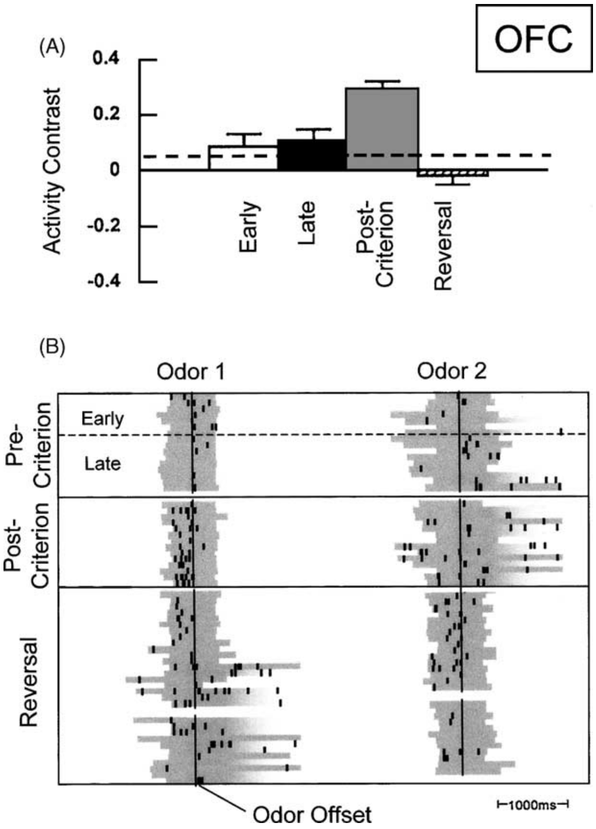

Fig. 1. Cue-selective neurons in orbitofrontal cortex fire in relation to conjunctions of cues and outcomes. Neurons in rat orbitofrontal cortex were recorded during acquisition and reversal of new 2-odor discrimination problems in a go, no-go paradigm. On each trial, a single odor was presented. Responses after sampling of the positive odor resulted in delivery of a sucrose solution; responses after sampling of the negative odor resulted in delivery of a quinine solution. Rats were presented with a novel odor pair in each session. Neural activity was recorded as the rats learned to withhold responses to the negative cue to avoid the quinine solution, during accurate performance (post-criterion), and during reversal training in which the reinforcers paired with the odor cues were switched. (A) Contrast in population activity during the odor sampling window (200 ms before to 150 ms after odor offset) for the 96 of 328 orbitofrontal neurons that exhibited selective firing to odor cues during accurate performance. Activity contrast was calculated as the difference in firing to positive and negative odors, referenced to the polarity of this difference during post-criterion trials, and normalized by the sum of those rates in each training phase. The dotted line represents a baseline value from shuffled trials. Selectivity in this population of neurons differed significantly from baseline only during the post-criterion phase, indicating that differential activity in these cells (illustrated in B) only developed in the post-criterion phase of pre-reversal training and disappeared after reversal. (B) An example of a neuron with differential firing during odor sampling that developed during accurate performance of the discrimination and disappeared following reversal. Neural activity is shown for representative trials in raster format. Trials are shown sequentially for each odor. Activity on each trial begins with odor onset, is synchronized to odor offset, and ends with a response or after 1500 ms for no-go trials (faded). This cell develops selective firing to odor 1 in the post-criterion trial block, during accurate performance on the discrimination. This selective response diminishes rapidly after reversal and, in fact, disappears. Note that selectivity does not reemerge when the criterion is met on the reversal (rasters following break in reversal section) and thus does not simply reflect the discriminative motor response or reinforcer identity. Adapted from Schoenbaum et al. [30].

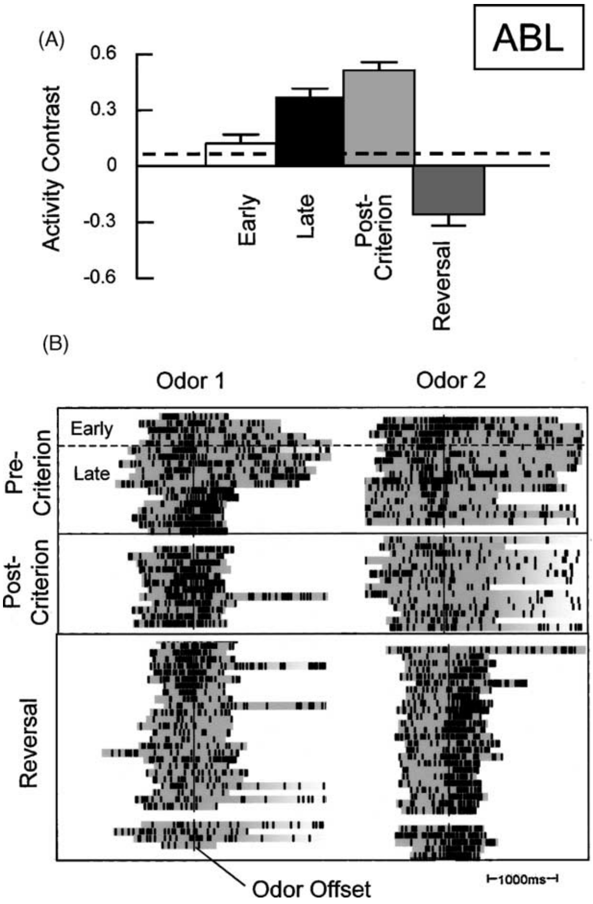

Fig. 2. Cue-selective neurons in the basolateral amygdala fire in relation to outcomes. Neurons in rat basolateral amygdala were recorded during acquisition and reversal of new 2-odor discrimination problems in a go, no-go paradigm identical to that described in Fig. 1. (A) Contrast in population activity during the odor sampling window (200 ms before to 150 ms after odor offset) for the 60 of 229 basolateral amygdala neurons that exhibited selective firing to odor cues during accurate performance. Activity contrast was calculated as described in Fig. 1. Selectivity in this population of neurons differed significantly from baseline during both the late pre-criterion and post-criterion phases, indicating that differential activity in these cells (illustrated in B) developed within the pre-criterion phase of pre-reversal training. Notably choice performance was at near-chance levels during this phase, thus selectivity in these cells developed independent of changes in behavior. In addition, these cells reversed selectivity during reversal training. (B) An example of a neuron with differential firing during odor sampling for the positive odor cue, which developed rapidly during initial acquisition and again after reversal. Neural activity is shown for representative trials in raster format. Trials are shown sequentially for each odor. Activity on each trial begins with odor onset, is synchronized to odor offset, and ends with a response or after 1500 ms for no-go trials (faded). This cell develops selective firing to odor 1 in the late pre-criterion trial block (after the dotted line). This selective response reverses rapidly after reversal of the outcomes associated with the two odors and well before the criterion is met on the reversal (rasters following break in reversal section) and thus does not simply reflect the discriminative motor response. Adapted from Schoenbaum et al. [30].

cortex and basolateral amygdala in behaviors based on the acquired value of cues [21,23,24,27,28,42]. However, these findings also suggest that these two brain regions perform subtly different roles within a system of structures involved in this function. In this system, the basolateral amygdala may be proportionately more involved in mediating the acquisition of incentive value by the cues, and the orbitofrontal cortex may access or use these representations subsequently to guide responses.

A recent study tested this hypothesis—again applying precisely the same behavioral paradigm as described above [43]. Recordings were made in the orbitofrontal cortex of rats with bilateral neurotoxic lesions of the basolateral amygdala. In this study, neuronal representations in orbitofrontal cortex during odor sampling were compared between lesioned and intact control rats during learning and reversal training. The results of this study indicate that without input from basolateral amygdala, representations related to the identities of the odor cues become less associative and more closely bound to the sensory features of the odor cues. In addition, neurons with outcome-related firing early in training failed to become activated by the associated odor cues after learning. These findings indicate that representations that reflect associations between the cues and outcomes in orbitofrontal cortex depend on input from basolateral amygdala for their formation during learning. After learning, the role of basolateral amygdala may be more limited, in that in some circumstances, it is not needed for expression/retention of this information once it is already learned [44,45].

2. Interactions between orbitofrontal cortex and the medial temporal lobe

If contributions from basolateral amygdala are important for molding sensory representations in orbitofrontal cortex to reflect associations between cues and outcomes, then connections with the parahippocampal region may be crucial for supporting more complex representations of the cues themselves. In particular, the parahippocampal region (including the perirhinal and entorhinal cortices) is crucial to so-called declarative, or episodic memory, in which representations of relationships between items (episodes) are maintained by other structures (see Ramus and Eichenbaum [46] for review). Such representations are not critical in a simple odor discrimination task like that described above [47,48]; however, they are crucial to tasks that require the rat to rapidly create flexible abstractions of the cues, as is required in the delayed non-matching to sample task [49,50].

In one version of the delayed non-matching to sample task, rats must use odor cues to make responses as in the simple discrimination paradigm. Unlike in the odor discrimination task described above, however, reward is predicted not by the identity of the odor but rather by whether or not the odor on the present trial matches the one on the previous trial. A reward is only given for a response if the odor is a "non-match." Thus, rats must form abstract representations indicating whether an odor is a "match" or "non-match" in order to perform the task. Lesions of either orbitofrontal cortex or the parahippocampal region impair performance in this task [49], and cells in both regions appear to encode the critical "match" or "non-match" judgment during odor sampling [32,51]. These data show that processing in both the parahippocampal region and the orbitofrontal cortex are equally critical to the task and suggest that, like orbitofrontal cortex and basolateral amygdala, these regions may form a system for integrating a particular kind of information—in this case declarative representations—with other information to guide behavior.

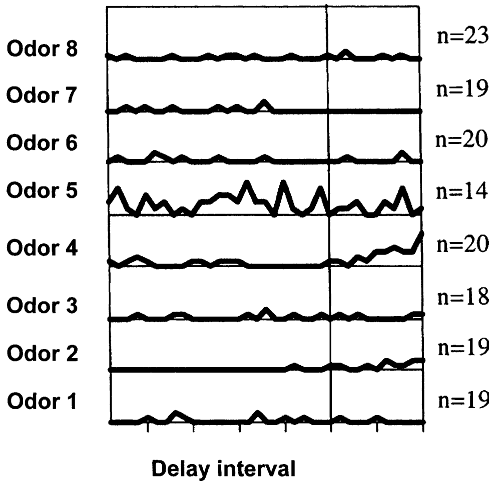

Fig. 3. Example of a parahippocmapal cortex neuron with a persistent representation of an odor stimulus. This cell was recorded from the entorhinal cortex (part of the parahippocampal region) of a rat performing an odor version of the delayed non-matching to sample task, and shows firing during the delay interval after the presentation of odor 5, but not following the presentation of any of the other seven odors. This sustained odor-specific firing seems to represent an active maintenance of the sample odor, a representation that is critical to building the match/non-match representations. Adapted from Young et al. [51].

A direct comparison of these two reports [32,51] indicates that, in addition to the similarities noted above, there are also key differences between the firing of neurons in orbitofrontal cortex and the parahippocampal region in this paradigm. First, more cells in the parahippocampal region than in orbitofrontal cortex demonstrated sustained cue-selective firing during the delay interval (Figs. 3 and 5B). This sustained firing seems to represent an active maintenance of the sample odor, a representation that is critical to building the match/non-match representations. Second, more cells in orbitofrontal cortex fired differently to the odor cues depending on whether the odor was a match or non-match to the previous odor (Figs. 4 and 5A). Such activity appears to reflect the significance of this abstract sensory construct and its association with reward. Third, there was a correlation between the performance of rats in a given session and the proportion of cells recorded in orbitofrontal cortex that fired differently to a given odor on match and non-match trials (Fig. 8A). In other words, orbitofrontal neurons code the critical match/non-match judgment. These data suggest that orbitofrontal cortex may play a key role in the abstrac-

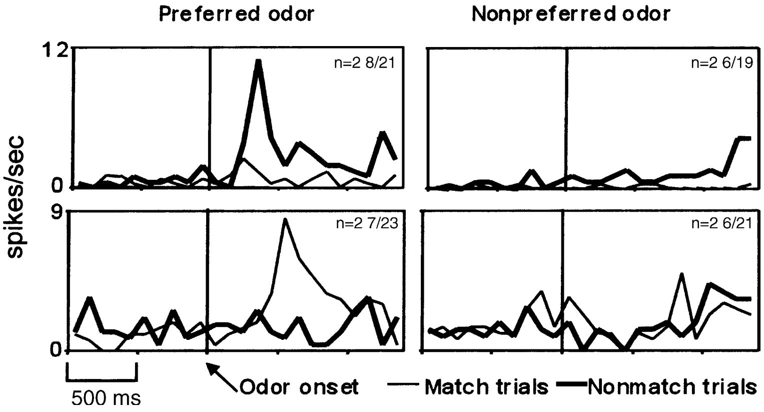

Fig. 4. Examples of two orbitofrontal cortex neurons with cue-selective increases in firing on match trials (match enhancement: bottom row panels) or non-match trials (match suppression: top row panels). These cells were recorded during high levels of performance of the odor version of the delayed non-matching to sample task. Left panel indicates firing for the preferred odor for the cell, and right panels the non-preferred odor. Note, that for both of these cells the match suppression or enhancement effect was cue-selective. *Dark lines* indicate average firing rates for non-match trials (number of recorded neurons indicated in dark numbers in upper right corner of panel), *light lines* indicate average firing rates for match trials (*n* indicated in light numbers). Firing rate is in spikes per second. Adapted from Ramus and Eichenbaum [32].

tion and representation of the non-matching rule and/or in applying that representation to guide responding.

Further supporting this idea, cells in orbitofrontal cortex typically fired in relation to more task events (i.e. the responses were more complex) than cells recorded in the parahippocampal region [32]. In fact, 88.4% of the neurons (244 of 276 cells) recorded in orbitofrontal cortex during performance of the delayed non-matching to sample task fired in relation to more than one task event. Indeed, neurons in orbitofrontal cortex are responsive to a variety of task events in other paradigms [32,35,37,52]. Together, these findings from orbitofrontal neurons recorded in different behavioral paradigms are consistent with the idea that orbitofrontal cortex is a nexus for the integration of sensory signals, originating from piriform cortex and elsewhere, and memory signals, arising in the parahippocampal region. They also suggest that orbitofrontal cortex participates in both the memory representations for specific stimuli and the acquisition and application of task rules. By this view, the parahippocampal region helps the orbitofrontal cortex to maintain representations of specific odor cues or episodes, while the orbitofrontal cortex uses this mnemonic information to abstract and apply task-specific rules. Similar findings have been reported in the monkey lateral prefrontal cortex [53] and perirhinal cortex [54]. Although task and species differences make direct comparison to the rodent work difficult, Miller has come to similar conclusions about the role of prefrontal cortex in recognition memory performance [55].

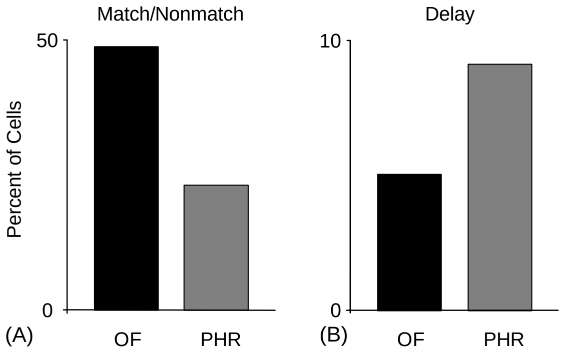

Fig. 5. Differences between the neural coding of cells in orbitofrontal cortex (OF; Ramus and Eichenbaum [32]) and the parahippocampal region (PHR; Young et al. [51]) of rats performing the same odor-guided non-matching to sample task. A greater proportion of cells in the orbitofrontal cortex demonstrated cue-selective changes in firing on match vs. non-match trials (match enhancement or suppression, see Fig. 4), while a greater proportion of cells in the parahippocampal region demonstrated sustained stimulus-selective firing during the delay interval (see Fig. 3).

3. Integration across systems

If the orbitofrontal cortex cooperates with the basolateral amygdala in some tasks to represent the value of cues, and with the parahippocampal region in other tasks to represent abstract properties such as the match/non-match characteristics of cues, the question remains as to how these two systems can be merged within a single model of orbitofrontal cortex function. Neurophysiological evidence from the studies reviewed here and in other reports suggests that in each case, representations in orbitofrontal cortex are driven by what is important for task performance. In other words, the cue-selective activity of orbitofrontal neurons discriminates attributes or features of the cues according to their associations with outcomes, and does so when these associations are reflected in accurate task performance. In this way, input from the two systems we have highlighted, as well as from other sources, is integrated in orbitofrontal cortex.

Indeed, initial recordings in rat orbitofrontal cortex showed that task variables in an 8-odor discrimination task were represented by the firing of the population according to their importance in the task [37,56]. In this study, neu-

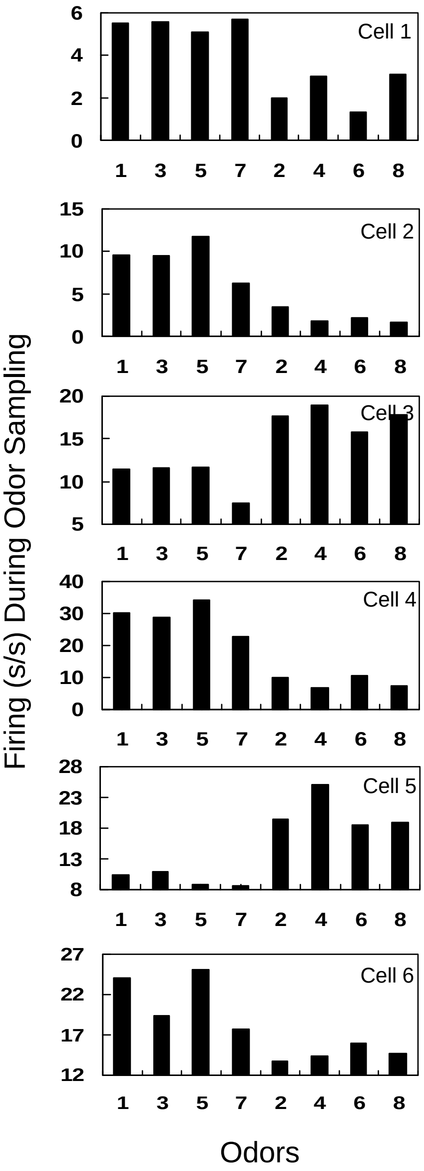

Fig. 6. Neurons in orbitofrontal cortex that exhibited differential firing during odor sampling. Neurons in rat orbitofrontal cortex were recorded during performance of an 8-odor go, no-go odor discrimination task. On each trial, one odor was presented. After odor sampling, a response could be made at a nearby fluid well for a water reward. Four odors were positive, indicating that a response would be rewarded, and four odors were negative, indicating that no reward would be given for responding. Rats were well-trained on the discrimination prior to recording, always responding to positive odors and rarely to negative odors. Activity is shown in spikes/second during odor sampling for six different neurons recorded in orbitofrontal cortex. Each panel shows activity for a different neuron to each of the eight different odors presented in each session. Positive odors 1^+, 3^+, 5^+, and 7^+ are on the left of each panel, and negative odors 2^-, 4^-, 6^-, and 8^- are on the right of each panel. Adapted from Schoenbaum and Eichenbaum [56].

rons in orbitofrontal cortex of well-trained rats exhibited differential activity during odor sampling in a go, no-go discrimination task (Fig. 6). The odor sequence across trials in this experiment included elements of a simple discrimination task as well as elements of the delayed non-matching to sample task we have just discussed. For example, each odor was directly associated with reward or non-reward (valence), but there were also odor pairs embedded in the sequence, such that the odor on one trial could often be used to predict reward on the next trial. Such an arrangement likely recruited information from both basolateral amygdala and the medial temporal lobe.

Consistent with this idea, activity in some orbitofrontal cells reflected the associations between particular odors or sets of odors and the reward, but activity in other cells reflected the fixed sequence of odors across multiple trials. These latter correlates were particularly interesting because such fixed and predictable sequences were used by the rats to facilitate their performance on the task. Subsequent ensemble analyses (Fig. 7A) confirmed that activity in orbitofrontal neurons represented odor identity and valence (Fig. 7B), features that were useful in performance in the task. Reoccurring sequences of specific odors were also well represented (Fig. 7D), but random features of the odor sequence, such as the valence or identity of the odor on the previous trial, were not (Fig. 7C). Moreover, information regarding the repeated sequences was well encoded by the ensembles *prospectively* (Fig. 7D, left panel) when it could be used to guide behavior but not *retrospectively* when it would not have been useful (Fig. 7D, right panel). In other words, information conveyed by the odors was represented in orbitofrontal cortex according to its utility in the task [55].

Similarly, in the two sets of studies reviewed here, encoding in orbitofrontal cortex appeared to be particularly sensitive to task requirements and choice performance. In the odor discrimination experiments in which rats learned new odor problems in each recording session [30], cue-selective firing in orbitofrontal neurons represented the associations between particular odors and outcomes during periods of accurate choice performance (Fig. 1). By contrast, selectivity in basolateral amygdala neurons that essentially represented the same information developed independent of choice performance (Fig. 2). Similarly, in the study examining encoding in orbitofrontal cortex in an odor-guided non-matching to sample task [32], match/non-match correlates were most prevalent in sessions characterized by accurate performance (Fig. 8A). During sessions when the rats performed poorly, few neurons exhibited firing that reflected this comparison during odor sampling. Notably, the prevalence of selective firing in orbitofrontal neurons related to the identity of the odor cues showed no relationship to the rats' performance in this task (Fig. 8B). Thus, neurons in orbitofrontal cortex appear to parse or manipulate afferent input from the two systems we have highlighted—basolateral amygdala and the parahippocampal region—in a manner that extracts the representations that are most useful to guiding responses in a particular circumstance.

This function is particularly evident in a direct comparison of firing activity in orbitofrontal cortex during accurate performance on the non-matching to sample task [32] and

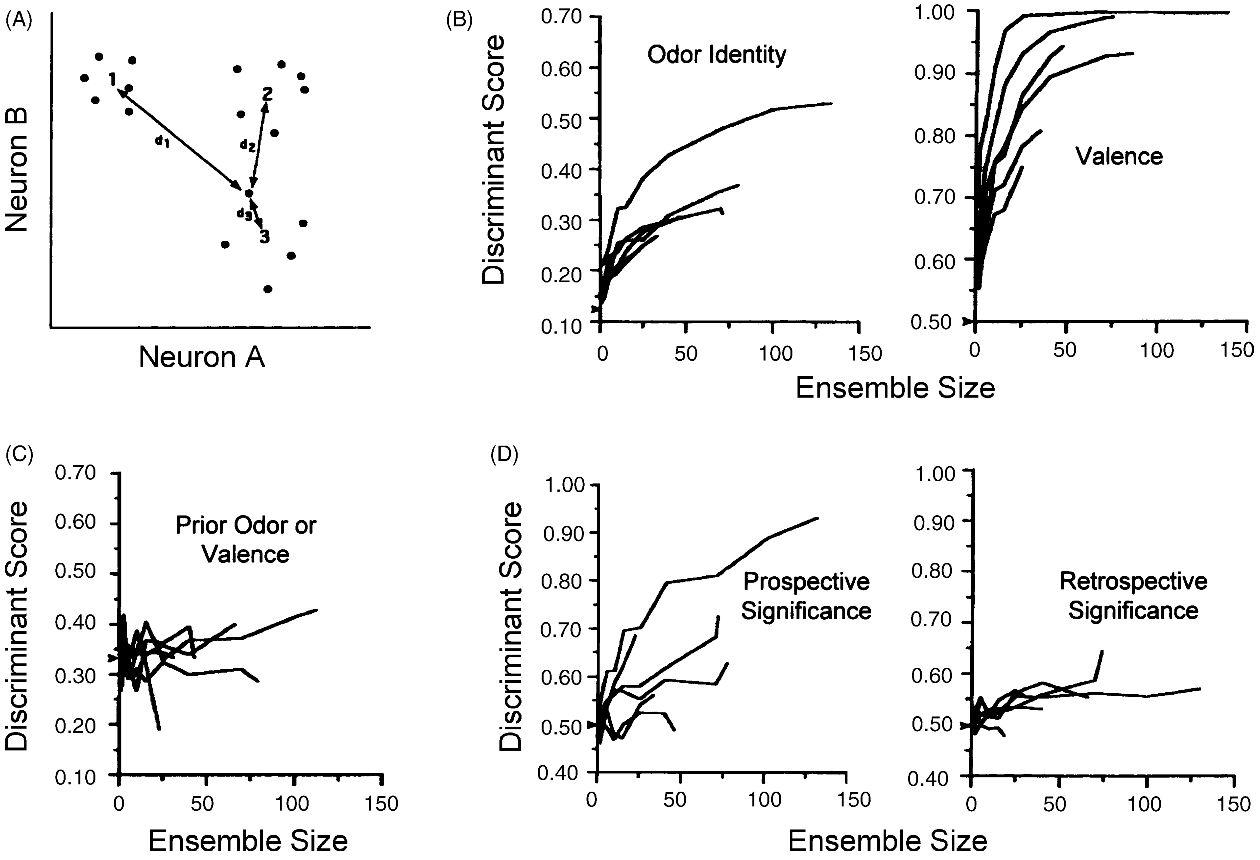

Fig. 7. Encoding in orbitofrontal cortex during odor sampling in an 8-odor go, no-go discrimination task (described in Fig. 6). A modified linear discriminant analysis [77] was performed to determine how well activity in populations of odor responsive orbitofrontal neurons in each rat could identify attributes of the odor cue on each trial and of the sequence of odor cues on preceding or subsequent trials. (A) Illustration of ensemble analysis. This analysis creates a space using the activity of each neuron as one of *n* dimensions. The population response on each trial is plotted in this space, and then the average population vector (1, 2, 3 or positive vs. negative, etc.) for each item in a discrimination (odor, valence, predictedness, etc.) is determined. The population vector for the response on each trial is then classified as belonging to the nearest cluster in this space by calculating the distance to each average vector ($\boldsymbol{d}_1$, $\boldsymbol{d}_2$, $\boldsymbol{d}_3$). The discriminant score is calculated by comparing the classification of each trial to the actual identity. (B) Discriminant score for ensembles of orbitofrontal neurons during discrimination performance. Populations were composed of odor responsive neurons. Individual neurons responded to both odor identity and valence, and ensembles performed well at correctly classifying odors according to these attributes. (C) Other neurons fired differentially depending on the odor presented on the preceding trial or on whether the prior trial had been rewarded or not. Such incidental information was irrelevant to task performance, and ensembles performed at chance at identifying these attributes. (D) Neurons also fired differently according to fixed pairs of odors that were always presented in sequence. This information was reflected in the latency to initiate the next (predicted) trial, and the ensembles performed well at discriminating between predictive and non-predictive odors (prospective encoding, left panel in D). However, ensembles performed poorly at classifying whether an odor had been predicted (retrospective encoding, right panel in D). Adapted from Gochin et al. [77] and Schoenbaum and Eichenbaum [56].

another odor discrimination study that was conducted earlier [37]. Both studies used a go, no-go paradigm and a set of eight odor cues in each session. The same set of odor cues was presented with the same contingencies in each session, and thus the studies differed only in the way in which the rats had to use the odor cues to guide performance. Both reports found high levels of task-relevant activity in orbitofrontal neurons associated with events throughout the trials, particularly during sampling of the odor cues. However, a comparison of cue-selective firing obtained during recording sessions in which rats performed accurately on the each task indicates that encoding in these cells differed according to task requirements in the two studies.

In the non-matching to sample task, in which a given odor could be rewarded on one trial but not on the next, selective firing to any of the eight different odors was observed in only about 16% of the orbitofrontal neurons (43 of 276 cells). By contrast, many more (64%) of the orbitofrontal neurons recorded during accurate performance in this task (175 of 276 cells) fired differently on match and non-match trials with the same group of odors. Further analysis of the odor-selective neurons revealed that about half (21 of 43) also fired differentially to a particular odor based on whether the odor was similar to (match) or different from (non-match) the previously sampled odor, and that only 13 of 43 odor-selective neurons were selective strictly based on the identity of the odor cues during odor sampling.

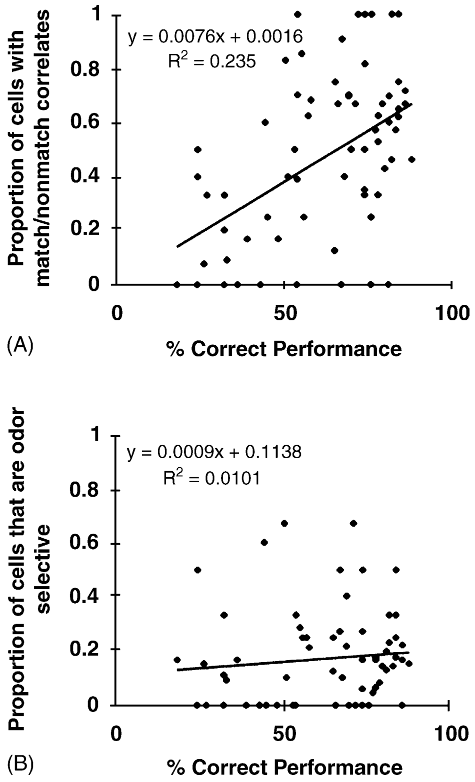

Fig. 8. The proportion of cells in orbitofrontal cortex that demonstrate cue-selective match enhancement or suppression varies with performance on the delayed non-matching to sample task. (A) This figure represents proportion of cue-selective match/non-match cells recorded in each session as a function of the percent correct performance on the same session. The number of match/non-match cells recorded in orbitofrontal cortex during a given session increased as the rats' performance increased. (B) By contrast, there was no similar relationship between the proportion of cells with cue-selective responses and the performance of the rats. Adapted from Ramus and Eichenbaum [32].

By comparison, in the odor discrimination task, odor-selectivity was observed in 77% or 311 of 404 orbitofrontal neurons in the odor discrimination task (in which a given odor was always associated with reward), as compared with 16% in the delayed non-matching task, and only a small proportion (∼17%) of cells fired differently during odor sampling based on random sequences of odors across trials. Indeed, *retrospective* encoding of odor sequence—a correlate that would be conceptually similar to the match/non-match correlate—was essentially not represented in ensembles composed of orbitofrontal cells recorded during discrimination performance (Fig. 7C and D, right panel), whereas odor identity was well represented (Fig. 7B). Thus, representations in orbitofrontal cortex during odor sampling in the discrimination and non-matching tasks encoded the attributes or features of the odor cues (odor identity or the match/non-match property) according to their associations with outcomes when this information was reflected in accurate task performance.

4. Critical contributions to behavior

The neurophysiological studies we have reviewed suggest that the orbitofrontal region, in both rats and primates, is specialized to integrate concrete and abstract sensory constructs with information regarding the incentive value of associated outcomes to guide or modulate behavior. Can this function account for the diversity of deficits and symptoms that are reported to result from orbitofrontal damage?

We can begin by considering the classic syndrome resulting from orbitofrontal damage, composed of behaviors that are not appropriate to current circumstances. These behaviors are variously described as perseverative, disinhibited, and impulsive, [57] and have often been attributed to an inability to withhold responding, thus assigning a role of response inhibition to the orbitofrontal cortex. Yet in some ways, this is a symptom rather than a function. Another interpretation of this deficit is that the subjects are unable to use incentive or motivational information to modify responding [58–60]. Such goal-neglect [61] might be manifested as perseverative, disinhibited, or impulsive responding, depending on the circumstances and whether the existing response tendency could be effectively modified using other associative mechanisms (e.g. "habit" or stimulus–response learning [62]). If this were possible in a particular situation, then no deficit would be observed. If not, then one would expect orbitofrontal damage to result in impairments in the ability to modulate previously learned responses and to acquire new responses based on information about the outcomes associated with cues. This variable pattern of deficits is precisely what is observed after orbitofrontal lesions. In humans, non-human primates, and rats, orbitofrontal lesions commonly result in reversal impairments [59,63–68], when an existing learned response tendency must be modified because the motivational value of the associated outcome has changed. These impairments result, in some cases, from perseverative errors and, in other cases, from errors in new learning after the reversal. These deficits are often observed despite an apparently normal ability to inhibit the same response when learning the original discrimination [59,64,65]. We have found in such a setting (the go, no-go odor discrimination task used in the recording experiments) that in animals with orbitofrontal lesions, apparently "normal" go, no-go learning can take place in the apparent absence of information about the outcomes associated with the odors [65]. Evidence for the absence of such information in lesioned animals is provided by the abolishment of normally observed differences in latency to respond at the fluid well following sampling of the positive and negative odors [65].

A similar function can account for the deficits observed after orbitofrontal lesions in Pavlovian and instrumental devaluation paradigms, which are explicitly designed to test this hypothesis. For example, orbitofrontal-lesioned monkeys trained to discriminate visual objects show apparently normal performance on the discriminations but then fail to modify their choice behavior when the value of one of the outcomes is decreased by overfeeding [27,69]. Similarly, orbitofrontal-lesioned rats trained in a Pavlovian conditioning task acquire apparently normal responses to a conditioned stimulus paired with food, but fail to modulate their responses to the cue after the value of the food outcome is decreased by pairing it with illness [28]. Note that original performance in both paradigms could be acquired using stimulus–response mechanisms, but to modify behavior after devaluation of the outcomes requires the ability to form and use associations between the cues and outcomes.

Even some of the more complex impairments now reported to occur after orbitofrontal damage may be understood in this fashion. In addition to representing the value of an associated outcome during cue sampling, many neurons in orbitofrontal cortex maintain representations of the value of the pending outcome across delays imposed between cue sampling and outcome delivery. Such working memory representations of value have been observed both in our experiments [30] and in primates [33,70]. These results suggest that like other prefrontal areas, the orbitofrontal regions have the ability to maintain representations across delays in the absence of sensory input. Orbitofrontal cortex may be particularly crucial when incentive information must be summed or integrated over time to appreciate the value of outcomes or consequences. Such is the case in the so-called gambling task employed to test human patients with orbital lesions [71]. In this task, a subject must make choices from decks of cards to obtain money reward. Choosing from “bad” decks usually delivers a high reward but occasionally results in a large penalty, leading to a net “loss” for the subject. Choosing from “good” decks usually delivers a low reward but seldom delivers a significant penalty, leading to a net “gain.” In these testing situations, subjects with orbitofrontal damage are able to make choices between consistently high- and low-rewarding decks of cards—thus, they perform like controls initially, choosing more from the high-reward decks. However, these subjects seem to have difficulty integrating the occasional large penalties into the decision-making process. As a result, they fail to bias their responding away from the “bad” decks to account for these penalties as quickly as controls. This step requires the ability to maintain, manipulate, and update information about the value of the cues and outcomes in memory, that appears to be dependent on the orbital prefrontal region [72].

Similar capabilities are required to perform normally in tasks when there is probabilistic or delayed reward [73,74]. For example, rats can be trained to discriminate between levers that deliver rewards of different magnitudes. Once trained, delays are introduced such that the rats must choose between an immediate small reward and a much larger but delayed reward. Normal rats calibrate their choices to reflect the difference in reward magnitude and the length of the delay. In another version of the task, delivery of the larger reward becomes probabilistic. In both circumstances, the animal must be able to associate cues with rewards, maintain those reward values in working memory, and integrate them across time. Rats with orbitofrontal damage perform abnormally in these tasks and variants of them [73,74]. Importantly, there seems to be little or no effect of lesions at zero delay or when the probabilities are similar, but an increasing effect as the delay increases or the probability decreases for obtaining the larger reward. Such delay-dependent effects indicate that orbitofrontal lesions disrupted the rats’ ability to integrate information about incentive value over time. In contrast, lesions of connected regions amygdala [75] and accumbens [76] appear to cause non-delay-dependent deficits in such tasks, possibly suggesting a deficit in discriminating reward value even in the absence of a delay.

In conclusion, the studies reviewed in this paper are consistent with the idea that the orbitofrontal regions integrate concrete and abstract sensory constructs with information regarding the incentive value of associated outcomes to guide or modulate behavior. Importantly, our work suggests strongly that this aspect of orbitofrontal function can be effectively studied in rats as well as primates. Our work further points out the importance of considering the function of the orbitofrontal cortex in relation to other brain systems, since the orbitofrontal cortex forms a nexus in which sensory, mnemonic and affective information are integrated to modulate behavioral outcomes.

Acknowledgements

This work was supported by K08-AG00882 and R01-DA015718 to G.S. and F32-MH12699 to B.S. and R01-MN51570 and R01-MH61492 to S.R.

References

[1] Cajal RY. Studies on the fine structure of the regional cortex of rodents 1: suboccipital cortex (retrosplenial cortex of Brodmann) [Trabajos del Laboratorio de Investigaciones Biologicas de la Universidad de Madrid 1922;20:1–30]. In: Jones EG, editor. Cajal on the cerebral cortex: an annotated translation of the complete writings. New York: Oxford University Press; 1988.

[2] Rose JE, Woolsey CN. The orbitofrontal cortex and its connections with the mediodorsal nucleus in rabbit, sheep, and cat. Res Pub Ass Nerv Ment Dis 1948;27:210–32.

[3] Leonard CM. The prefrontal cortex of the rat. I. Cortical projections of the mediodorsal nucleus. II. Efferent connections. Brain Res 1969;12:321–43.

[4] Markowitsch HJ, Pritzel M. The prefrontal cortex: projection area of the thalamic mediodorsal nucleus. Physiol Psychol 1979;7:1–6.

[5] Kolb B. Functions of the frontal cortex of the rat: a comparative review. Brain Res Rev 1984;8:65–98.

[6] Schoenbaum G, Setlow B, Gallagher M. Orbitofrontal cortex: modeling prefrontal function in rats. In: Squire L, Schacter D, editors. The neuropsychology of memory. New York: Guilford Press; 2002.

[7] Wise SP, Murray EA, Gerfen CR. The frontal cortex—basal ganglia system in primates. Crit Rev Neurobiol 1996;10:317–56.

[8] Kita H, Kitai ST. Amygdaloid projections to the frontal cortex and the striatum in the rat. J Comp Neurol 1990;298:40–9.

[9] Krettek JE, Price JL. Projections from the amygdaloid complex to the cerebral cortex and thalamus in the rat and cat. J Comp Neurol 1977;172:225–54.

[10] Shi CJ, Cassell MD. Cortical, thalamic, and amygdaloid connections of the anterior and posterior insular cortices. J Comp Neurol 1998;399:440–68.

[11] Ghashghaei HT, Barbas H. Pathways for emotion: interactions of prefrontal and anterior temporal pathways in the amygdala of the rhesus monkey. Neuroscience 2002;115:1261–79.

[12] Deacon TW, Eichenbaum H, Rosenberg P, Eckmann KW. Afferent connections of the perirhinal cortex in the rat. J Comp Neurol 1983;220:168–90.

[13] Price JL, Carmichael ST, Carnes KM, Clugnet M-C, Kuroda M, Ray JP. Olfactory input to the prefrontal cortex. In: Davis J, Eichenbaum H, editors. Olfaction: a model system for computational neuroscience. Cambridge, MA: MIT Press; 1991.

[14] Brown S, Schafer EA. An investigation into the functions of the occipital and temporal lobes of the monkey's brain. Phil Trans R Soc Lond B 1888;179:303–27.

[15] Harlow JM. Recovery after passage of an iron bar through the head. Pub Massachusetts Med Soc 1868;2:329–46.

[16] Kluver H, Bucy PC. Preliminary analysis of the temporal lobes in monkeys. Arch Neurol Psychiatry 1939;42:979–1000.

[17] Weiskrantz L. Behavioral changes associated with ablations of the amygdaloid complex in monkeys. J Comp Physiol Psychol 1956;9:381–91.

[18] Cador M, Robbins TW, Everitt BJ. Involvement of the amygdala in stimulus–reward associations: interactions with the ventral striatum. Neuroscience 1989;30:77–86.

[19] LeDoux JE, Cicchetti P, Xagoraris A, Romanski LM. The lateral amygdaloid nucleus: sensory interface of the amygdala in fear conditioning. J Neurosci 1990;10:1062–9.

[20] McDonald RJ, White NM. A triple dissociation of memory systems: hippocampus, amygdala, and dorsal striatum. Behav Neurosci 1993;107:3–22.

[21] Hatfield T, Han JS, Conley M, Gallagher M, Holland P. Neurotoxic lesions of basolateral, but not central, amygdala interfere with Pavlovian second-order conditioning and reinforcer devaluation effects. J Neurosci 1996;16:5256–65.

[22] Gewirtz JC, Davis M. Second-order fear conditioning prevented by blocking NMDA receptors in amygdala. Nature 1997;388:471–4.

[23] Malkova L, Gaffan D, Murray EA. Excitotoxic lesions of the amygdala fail to produce impairment in visual learning for auditory secondary reinforcement but interfere with reinforcer devaluation effects in rhesus monkeys. J Neurosci 1997;17:6011–20.

[24] Parkinson JA, Crofts HS, McGuigan M, Tomic DL, Everitt BJ, Roberts AC. The role of the primate amygdala in conditioned reinforcement. J Neurosci 2001;21:7770–80.

[25] Gallagher M. The amygdala and associative learning. In: Aggleton JP, editor. The amygdala: a functional analysis. New York: Oxford University Press; 2000.

[26] Everitt BJ, Cardinal RN, Hall J, Parkinson JA, Robbins TW. Differential involvement of amygdala subsystems in appetitive conditioning and drug addiction. In: Aggleton JP, editor. The amygdala: a functional analysis. New York: Oxford University Press; 2000.

[27] Baxter MG, Parker A, Lindner CCC, Izquierdo AD, Murray EA. Control of response selection by reinforcer value requires interaction of amygdala and orbitofrontal cortex. J Neurosci 2000;20:4311–9.

[28] Gallagher M, McMahan RW, Schoenbaum G. Orbitofrontal cortex and representation of incentive value in associative learning. J Neurosci 1999;19:6610–4.

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

[30] Schoenbaum G, Chiba AA, Gallagher M. Neural encoding in orbitofrontal cortex and basolateral amygdala during olfactory discrimination learning. J Neurosci 1999;19:1876–84.

[31] Rolls ET, Critchley HD, Mason R, Wakeman EA. Orbitofrontal cortex neurons: role in olfactory and visual association learning. J Neurophysiol 1996;75:1970–81.

[32] Ramus SJ, Eichenbaum H. Neural correlates of olfactory recognition memory in the rat orbitofrontal cortex. J Neurosci 2000;20:8199–208.

[33] Tremblay L, Schultz W. Relative reward preference in primate orbitofrontal cortex. Nature 1999;398:704–8.

[34] Wallis JD, Anderson KC, Miller EK. Single neurons in prefrontal cortex encode abstract rules. Nature 2001;411:953–6.

[35] Lipton PA, Alvarez P, Eichenbaum H. Crossmodal associative memory representations in rodent orbitofrontal cortex. Neuron 1999;22:349–59.

[36] Schoenbaum G, Chiba AA, Gallagher M. Orbitofrontal cortex and basolateral amygdala encode expected outcomes during learning. Nat Neurosci 1998;1:155–9.

[37] Schoenbaum G, Eichenbaum H. Information coding in the rodent prefrontal cortex. I. Single-neuron activity in orbitofrontal cortex compared with that in pyriform cortex. J Neurophysiol 1995;74:733–50.

[38] Maren S. Auditory fear conditioning increases CS-elicited spike firing in lateral amygdala neurons even after extensive overtraining. Eur J Neurosci 2000;12:4047–54.

[39] Nishijo H, Ono T, Nishino H. Single neuron responses in alert monkey during complex sensory stimulation with affective significance. J Neurosci 1988;8:3570–83.

[40] Quirk GJ, Repa JC, LeDoux JE. Fear conditioning enhances short-latency auditory responses of lateral amygdala neurons: parallel recordings in the freely behaving rat. Neuron 1995;15:1029–39.

[41] Muramoto K, Ono T, Nishijo H, Fukuda M. Rat amygdaloid neuron responses during auditory discrimination. Neuroscience 1993;52:621–36.

[42] Pears A, Parkinson JA, Everitt BJ, Roberts AC. Effects of orbitofrontal cortex lesions on responding with conditioned reinforcement. Brain Cogn 2001;47:44–6.

[43] Schoenbaum G, Setlow B, Saddoris MP, Gallagher M. Encoding predicted outcome and acquired value in orbitofrontal cortex during cue sampling depends upon input from basolateral amygdala. Neuron 2003;39:855–67.

[44] Setlow B, Gallagher M, Holland P. The basolateral complex of the amygdala is necessary for acquisition but not expression of CS motivational value in appetitive Pavlovian second-order conditioning. Eur J Neurosci 2002;15:1841–53.

[45] Pickens, CL, Saddoris, MP, Setlow, B, Gallagher, M, Holland PC, Schoenbaum G. Different roles for orbitofrontal cortex and basolateral amygdala in a reinforcer devaluation task. Journal of Neuroscience, in press.

[46] Ramus SJ, Eichenbaum H. A brain system for declarative memory. In: Pomerantz J, editor. Topics in integrative neuroscience: from cells to cognition. Cambridge: Cambridge University Press, in press.

[47] Otto T, Schottler R, Staubli U, Eichenbaum H, Lynch G. Hippocampus and olfactory discrimination learning: effects of entorhinal cortex lesions on olfactory learning and memory in a successive-cue, go-no-go task. Behav Neurosci 1991;105:111–9.

[48] Eichenbaum H, Fagan A, Cohen NJ. Normal olfactory discrimination learning set and facilitation of reversal learning after medial-temporal damage in rats: implications for an account of preserved learning abilities in amnesia. J Neurosci 1986;6:1876–84.

[49] Otto T, Eichenbaum H. Complementary roles of the orbital prefrontal cortex and the perirhinal–entorhinal cortices in an odor-guided delayed-nonmatching-to-sample task. Behav Neurosci 1992;106:762–75.

[50] Eichenbaum H, Alvarez P, Ramus SJ. Animal models of amnesia. In: Cermak L, editor. Handbook of neuropsychology: memory disorders. Amsterdam: Elsevier Science; 2000.

[51] Young BJ, Otto T, Fox GD, Eichenbaum H. Memory representation within the parahippocampal region. J Neurosci 1997;17:5183–95.

[52] Alvarez P, Eichenbaum H. Representations of odors in the rat orbitofrontal cortex change during and after learning. Behav Neurosci 2002;116:421–33.

[53] Miller EK, Erickson CA, Desimone R. Neural mechanisms of visual working memory in prefrontal cortex of the Macaque. J Neurosci 1996;16:5154–67.

[54] Suzuki WA, Miller EK, Desimone R. Object and place memory in the Macaque entorhinal cortex. J Neurophysiol 1997;78:1062–81.

[55] Miller EK. The prefrontal cortex and cognitive control. Nat Neurosci Rev 2000;1:59–65.

[56] Schoenbaum G, Eichenbaum H. Information coding in the rodent prefrontal cortex. II. Ensemble activity in orbitofrontal cortex. J Neurophysiol 1995;74:751–62.

[57] Fuster JM. The prefrontal cortex. New York: Lippin-Ravencott; 1997.

[58] Hauser MD. Perseveration, inhibition and the prefrontal cortex: a new look. Curr Opin Neurobiol 1999;9:214–22.

[59] Dias R, Robbins TW, Roberts AC. Dissociable forms of inhibitory control within prefrontal cortex with an analog of the Wisconsin card sort test: restriction to novel situations and independence from on-line processing. J Neurosci 1997;17:9285–97.

[60] Schoenbaum G, Setlow B. Integrating orbitofrontal cortex into prefrontal theory: common processing themes across species and subdivision. Learn Mem 2001;8:134–47.

[61] Duncan J, Emslie H, Williams P. Intelligence and the frontal lobe: the organization of goal-directed behavior. Cogn Psychol 1996;30:257–303.

[62] Thorndike EL. Animal intelligence: an experimental study of the associative processes in animals. Psychol Rev 1898;2:1–107.

[63] Rolls ET, Hornak J, Wade D, McGrath J. Emotion-related learning in patients with social and emotional changes associated with frontal lobe damage. J Neurol Neurosurg Psychiatry 1994;57:1518–24.

[64] Brown VJ, Bowman EM. Rodent models of prefrontal cortical function. Trends Neurosci 2002;25:340–3.

[65] Schoenbaum G, Setlow B, Nugent SL, Saddoris MP, Gallagher M. Lesions of orbitofrontal cortex and basolateral amygdala complex disrupt acquisition of odor-guided discriminations and reversals. Learn Mem 2003;10:129–40.

[66] Ferry AT, Lu XC, Price JL. Effects of excitotoxic lesions in the ventral striatopallidal–thalmocortical pathway on odor reversal learning: inability to extinguish an incorrect response. Experimental Brain Research 2000;131:320–35.

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

[68] Jones B, Mishkin M. Limbic lesions and the problem of stimulus–reinforcement associations. Exp Neurol 1972;36:362–77.

[69] Izquierdo AD, Murray EA. Bilateral orbital prefrontal cortex lesions disrupt reinforcer devaluation effects in rhesus monkeys. Soc Neurosci Abstr 2000;26:978.

[70] Hikosaka K, Watanabe M. Delay activity of orbital and lateral prefrontal neurons of the monkey varying with different rewards. Cereb Cortex 2000;10:263–71.

[71] Bechara A, Damasio H, Tranel D, Damasio AR. Deciding advantageously before knowing the advantageous strategy. Science 1997;275:1293–4.

[72] Bechara A, Damasio H, Tranel D, Andersen SW. Dissociation of working memory from decision making within the human prefrontal cortex. J Neurosci 1998;18:428–37.

[73] Mobini S, Body S, Ho MY, Bradshaw CM, Szabadi E, Deakin JFW, et al. Effects of lesions of the orbitofrontal cortex on sensitivity to delayed and probabilistic reinforcement. Psychopharmacology 2002;160:290–8.

[74] DeCoteau WE, Kesner RP, Williams JM. Short-term memory for food reward magnitude: the role of the prefrontal cortex. Behav Brain Res 1997;88:239–49.

[75] Kesner RP, Williams JM. Memory for magnitude of reinforcement: dissociation between amygdala and hippocampus. Neurobio Learn Mem 1995;64:237–44.

[76] Cardinal RN, Pennicott DR, Sugathapala CL, Robbins TW, Everitt BJ. Impulsive choice induced in rats by lesions of the nucleus accumbens core. Science 2001;292:2499–501.

[77] Gochin PM, Colombo M, Dorfman GA, Gerstein GL, Gross CG. Neural ensemble coding in inferior temporal cortex. J Neurophysiol 1994;71:2325–37.